Spelling and Reading Comprehension: Investigating Unique Relations in Third Grade

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Kimberly A. Murphy

Graduate Program in Education: Teaching and Learning

The Ohio State University
2016

Dissertation Committee:
Laura M. Justice, Chair
Richard G. Lomax
Monique T. Mills
Abstract

Considerable research effort has focused on understanding reading comprehension and reading comprehension difficulties. The purpose of the present study was to add to the small but growing body of literature on the role that spelling may play. It was the first to investigate the full range of lexical-level literacy skills and the unique contribution that spelling may make to reading comprehension when controlling for word recognition and vocabulary. The study also examined whether the relation between spelling and reading comprehension varied with spelling scoring metric. Data were collected from 63 children in grade 3 who were participating in a larger study on language and reading comprehension. In addition to measuring reading comprehension, word recognition, and vocabulary, four spelling scoring metrics were examined: the number of words spelled correctly (WSC), the number of correct letter sequences (CLS), and Spelling Sensitivity Scores (Masterson & Apel, 2010) for elements (SSE) and words (SSW). Results of hierarchical regressions showed that spelling was a significant, unique predictor of reading comprehension when the CLS metric was used. The scoring metrics were differentially related to reading comprehension. Metrics that gave credit based on orthographic precision only (WSC, CLS) were more highly related to reading comprehension than metrics that scored across multiple linguistic domains (SSE, SSW) and therefore gave credit even for imprecise spellings (i.e., those that demonstrated
phonological accuracy only). These results indicate that spelling may provide unique information about children’s reading comprehension and suggest promising future research directions to investigate the predictive ability of spelling and its potential to aid in the identification of children with, or at risk for, reading comprehension difficulties.
Acknowledgments

I would like to thank my mentors, friends, and family who were there for me along this journey, from the initial conceptualization of my research to the writing of the final words. This work is possible because of your guidance, support, and love.

I am deeply grateful to my advisor and mentor, Laura Justice, for inviting me on this path, providing me with a multitude of opportunities, and always believing in me. You gave me the skills and confidence I need to go forth into the world of academia, knowing I will do good work and make a difference. To my wonderful committee members, Richard Lomax and Monique Mills, your support has been invaluable. You encouraged me, mentored me, and helped me through the tough spots. I truly appreciated and benefitted from your expertise and passion.

I would like to acknowledge funding sources that made my doctoral work possible. I was extremely fortunate to work as a GRA on the LARRC project, funded by Grant R305F100002 from the U.S. Department of Education, Institute of Education Sciences. I also received GRA support from the Crane Center for Early Childhood Research and Policy. These assistantships provided not only funding but invaluable research experiences as well.

To all the fabulous LARRC staff, especially Jennifer Bostic, Lisa Baldwin-Skinner, and the OSU field assessors, thank you for making my data collection plans
possible. You jumped in without hesitation and made it happen. To those who assisted with coding of my data, Susie Mauck, Rachel Wade, and Tricia Biancone, your willingness to help was much appreciated. Susie, you went above and beyond your role at the Research Methodology Center, for which I am also grateful.

A special thank you to Shayne Piasta, for helping me realize that I could do what I imagined for my dissertation, giving me the courage to do so, and endlessly talking through ideas with me. You have a gift for helping people see things more clearly and I’m grateful to have been a recipient of that gift.

Last but not least, I am forever grateful for the support of my friends and family throughout this journey. I will simply say thank you, from the bottom of my heart. To my ‘old’ friends in St. John’s and Montreal who encouraged and supported me from the beginning and kept nudging me all the way along. To my ‘new’ friends in Columbus who welcomed me and then did the same. You listened, you offered advice, you understood when I struggled and you pulled me back up. You cheered my successes. Finally, and most importantly, I thank my family for every little thing they’ve done to keep me going. Mom and Dad, you knew when to spoil me with extra TLC (which was most of the time) and that was exactly what I needed. Your love and support mean the world to me.
LARRC Acknowledgements

This dissertation used data from the Language and Reading Research Consortium (LARRC). LARRC project sites and investigators are as follows:

Ohio State University (Columbus, OH): Laura M. Justice (Site PI), Richard Lomax, Ann O’Connell, Jill Pentimonti¹, Stephen A. Petrill², Shayne B. Piasta

Arizona State University (Tempe, AZ): Shelley Gray (Site PI), Maria Adelaida Restrepo.

Lancaster University (Lancaster, UK): Kate Cain (Site PI).

University of Kansas (Lawrence, KS): Hugh Catts³ (Site PI), Mindy Bridges, Diane Nielsen.

University of Nebraska-Lincoln (Lincoln, NE): Tiffany Hogan (Site PI), Jim Bovaird, J. Ron Nelson.⁴

MGH Institute of Health Professions (Boston, MA): Tiffany Hogan (Site PI)

1. Jill Pentimonti is now at American Institutes for Research (AIR).

2. Stephen A. Petrill was a LARRC co-investigator from 2010-2013.

3. Hugh Catts is now at Florida State University.

4. J. Ron Nelson was a LARRC co-investigator from 2010-2012.

This work was supported by grant # R305F100002 of the Institute of Education Sciences’ Reading for Understanding Initiative. We are deeply grateful to the numerous staff, research associates, school administrators, teachers, children, and families who participated. Key personnel at study sites include: Crystle Alonzo, Lisa Baldwin-Skinner, Lauren Barnes, Garey Berry, Beau Bevens, Jennifer Bostic, Tracy Centanni, Beth
Chandler, Lori Chleborad, Emmanuel Cortez, Dawn Davis, Kelsey Dickerhoof, Jaclyn Dynia, Michel Eltschinger, Yvonne Fraser, Sara Gilliam, Alexandria Hamilton, Cindy Honnens, Hui Jiang, Chi Luu, Junko Maekawa, Carol Mesa, Denise Meyer, Maria Moratto, Kimberly Murphy, Trevor Rey, Lizeth Sanchez-Verduzco, Amber Sherman, Shannon Tierney, and Gloria Yeomans-Maldonado.

The views presented in this work do not represent those of the federal government, nor do they endorse any products or findings presented herein. Correspondence concerning this work should be sent to Kimberly Murphy.
Vita

1987 ......................................................... O’Donel High School

1991 .......................................................... B.Sc. Psychology, Memorial University of Newfoundland

1993 .......................................................... M.Sc.(A) Speech-Language Pathology, McGill University

1993 to 2012 ................................................. Speech-Language Pathologist

2012 to present ............................................... Graduate Research Associate, Department of Teaching and Learning, The Ohio State University
Publications


Gingras, A. & Murphy, K.A. (2009). Emerging literacy in a multicultural community. CASLPA Communiqué, 23 (3).
In Press


Field of Study

Major Field: Education: Teaching and Learning

Area of Specialization: Reading and Literacy in Early and Middle Childhood
# Table of Contents

Abstract ......................................................................................................................... ii

Acknowledgments ........................................................................................................ iv

Vita ................................................................................................................................. viii

Publications .................................................................................................................... ix

Field of Study ................................................................................................................ x

Table of Contents .......................................................................................................... xi

List of Tables ................................................................................................................ xiii

List of Figures ................................................................................................................. xiv

Chapter 1: Introduction ................................................................................................. 1

Chapter 2: Review of Literature ................................................................................... 6

Chapter 3: Methods ....................................................................................................... 41

Chapter 4: Results ......................................................................................................... 56

Chapter 5: Discussion ................................................................................................. 66

References ..................................................................................................................... 87

Appendix A: Tables ..................................................................................................... 109
Appendix B: Figures ................................................................. 117

Appendix C: Assessor Training Materials - Spelling Assessment Fidelity Checklist.... 127

Appendix D: Spelling Administration and Instructions............................................. 129
List of Tables

Table 1. Descriptives .......................................................................................................................... 109
Table 2. Correlations between the Predictors and Reading Comprehension...................... 110
Table 3. Loadings for Principal Components Analysis of Lexical Literacy Skill .......... 111
Table 4. Hierarchical Regression Results for the Words Spelled Correctly Metric ........... 112
Table 5. Hierarchical Regression Results for the Correct Letter Sequences Metric ...... 113
Table 6. Hierarchical Regression Results for the Spelling Sensitivity-Words Metric ...... 114
Table 7. Hierarchical Regression Results for the Spelling Sensitivity-Elements Metric 115
Table 8. Pratt Index for Models with each Spelling Metric......................................................... 116
List of Figures

Figure 1. Scatterplot of word recognition with reading comprehension .................. 117
Figure 2. Scatterplot of vocabulary with reading comprehension .......................... 118
Figure 3. Scatterplot of words spelled correctly with reading comprehension .......... 119
Figure 4. Scatterplot of correct letter sequences with reading comprehension ........ 120
Figure 5. Scatterplot of spelling sensitivity-words with reading comprehension ...... 121
Figure 6. Scatterplot of spelling sensitivity-elements with reading comprehension .... 122
Figure 7. Scatterplot of words spelled correctly with word recognition ................... 123
Figure 8. Scatterplot of correct letter sequences with word recognition ................. 124
Figure 9. Scatterplot of spelling sensitivity-words with word recognition ............... 125
Figure 10. Scatterplot of spelling sensitivity-elements with word recognition ........ 126
Chapter 1: Introduction

Reading is a complex skill that defies a single theory (Perfetti & Stafura, 2014). To read for understanding is the ultimate goal of reading yet, in the United States in 2015, 64% of fourth graders and 66% of eighth graders performed at basic or below basic levels in reading comprehension (National Center for Education Statistics, 2015). Early reading difficulties often persist through adolescence (Ferrer et al., 2015) and about 16% of children who do not read proficiently in third grade will fail to graduate from high school (The Annie E. Casey Foundation, 2012). Thus, identifying children who may struggle with reading comprehension is an important endeavor and has been the subject of a large body of research.

The present study sought to examine the relation between spelling and reading comprehension. Processes involved in reading comprehension occur across linguistic levels, from words (lexical level) to sentences and text (supralexical level). The present study is concerned specifically with lexical-level skills, which include spelling, word recognition, and vocabulary. The contribution of word recognition and vocabulary to reading comprehension has received much attention in the literature, but strikingly little research has been conducted on spelling, particularly with regards to its association with reading comprehension. The primary aim of the present study, therefore, is to investigate
the unique contribution of spelling to reading comprehension in the presence of word recognition and vocabulary.

Spelling is a literacy construct of interest for several reasons. Once considered primarily a visual skill that required rote memorization to learn, spelling is now understood to be a complex, developmental, and language-based skill (Ouellette & Sénéchal, 2008; Treiman & Bourassa, 2000). Researchers have argued for the importance of spelling to literacy development (Foorman & Petscher, 2010; Pinto, Bigozzi, Tarchi, Accorti Gamannossi, & Canneti, 2015). For instance, Pinto and colleagues (Pinto et al., 2015) stated that “early progress in spelling becomes a resource for later reading acquisition” (p. 1233). Spelling, like word recognition, depends upon the integration of core linguistic skills including phonology, orthography, and morphology, and therefore provides insight into developing competencies in children’s linguistic systems. Spelling can be thought of as the ‘flip side of the coin’ (Ehri, 2000) from word recognition. Whereas word recognition involves the decoding or translation of print to speech, spelling involves the encoding of speech to print. Although they have much in common and are highly correlated (Abbott, Berninger, & Fayol, 2010), spelling and word recognition are not perfectly overlapping skills. Spelling is more difficult than word recognition, in part because it is a production task. Accurate spelling requires a perfectly precise orthographic image of a word, whereas a less precise orthographic image will sometimes suffice for word recognition, especially where contextual cues can facilitate recognition.
Spelling is difficult also in part due to the complexity of English orthography. English is an opaque orthography, meaning there is a lack of one-to-one correspondence between letters and phonemes. To overcome the mismatch between the number of letters and the number of phonemes (40+) in the language, some phonemes are represented by letter combinations (e.g., /<s> and <h> are combined to form /ʃ/), and some letters can represent more than one phoneme (e.g., /<c> can sound like /k/ or /s/). Furthermore, English is a morphophonemic language (N. Chomsky & Halle, 1968) in which the morphological structure of words is maintained even in cases where pronunciation shifts (e.g., the pronunciation of <sign> is different in the words sign and signal, but the spelling is maintained for the sake of meaning). English also has many homophones, words with the same pronunciation but different meanings (e.g., disc-disk, feet-feat), which generally are spelled differently to distinguish them. Furthermore, many words in English originated in other languages and maintain the spelling associated with that language (e.g., the spelling <ch> for the sound /ʃ/ in chef is the result of its French origin). These examples illustrate the complexity and therefore challenges of learning English orthography and becoming a good speller.

Given the complexity of English orthography and that spelling is a multidimensional linguistic skill, it is not surprising that there are several methods of scoring children’s spellings. Scoring metrics vary in the depth in which they attend to the accuracy of the spelling (e.g., whole word versus phonemes, graphemes, and morphemes) and whether they give credit for partially correct spellings. Some research suggests that metrics that give partial credit may indicate a stronger relation between spelling and word
recognition (Clemens, Oslund, Simmons, & Simmons, 2014), but this is a question that remains unanswered with respect to reading comprehension. The present study therefore aimed to compare the relation between spelling and reading comprehension across a variety of scoring metrics.

The theoretical rationale for the present study is grounded in the lexical quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2002). The lexical quality hypothesis states that high quality lexical representations, in phonology, orthography, morphology, and semantics, support efficient word recognition and therefore play an important role in reading comprehension. Spelling is an indicator of underlying lexical quality and may, in fact, be a stronger indicator than word recognition since it requires highly specified orthographic representations that are well connected to phonology and semantics (Ehri, 2000; Perfetti, 1997). It is therefore hypothesized that spelling will relate to reading comprehension and may contribute unique variance beyond the contribution of word recognition.

To summarize, the present study uses a lexical quality framework to examine the relation between spelling and reading comprehension. This study makes a unique contribution to the literature by investigating the role of spelling amongst the full set of lexical-level literacy skills and exploring whether the strength of the relation varies with spelling scoring metric. This work has the potential to impact researchers’ and practitioners’ use of spelling assessment as a tool for gaining insight into children’s literacy skills, especially as related to reading comprehension. Furthermore, this work may inform future efforts to examine the longitudinal prediction of reading
comprehension from spelling, and to develop a spelling screener that can be used in the early identification of reading comprehension difficulties. Other researchers (Robbins, Hosp, Hosp, & Flynn, 2010) have recommended spelling assessment as a screener for decoding difficulties but the present study may offer support for examining the utility of spelling as a screener for reading comprehension as well. Additionally, this study adds support to the growing appreciation of spelling as an essential literacy construct and target for literacy instruction.
Chapter 2: Literature Review

The purpose of chapter two is to discuss the literature relevant to reading comprehension and spelling. Within the following literature review, I first provide a broad overview of reading comprehension theory, followed by a discussion of the word-level components of reading comprehension. Next, I present an overview of the spelling process, including what spelling is, how it develops, and its core linguistic components. I then provide a theoretical rationale for the relation between spelling and reading and a review of existing research that has explored the relation between spelling and reading, with a particular focus on reading comprehension. Finally, I turn to a description of a variety of spelling scoring metrics and the research that has compared them.

**Reading Comprehension Theory**

The ability to read for understanding, that is, to comprehend what one reads, is necessary not only for skilled literacy itself but also for academic achievement, since reading is a primary mechanism through which content knowledge is acquired in school. A variety of theoretical frameworks that describe the processes and/or products of reading comprehension exist, two of which are particularly relevant to the present study.

**The simple view of reading.** A particularly influential framework of reading, the simple view of reading (Gough & Tunmer, 1986; Hoover & Gough, 1990) states that reading comprehension is the product of decoding (referred to in the present study as
word recognition) and linguistic comprehension. These components have also been referred to as print-dependent and print-independent, respectively (Protopapas, Simos, Sideridis, & Mouzaki, 2012). As defined by Hoover and Gough (1990), word recognition is the rapid retrieval of a word from print. Linguistic comprehension involves using the semantic information accessed during decoding to derive sentence- and discourse-level meaning of a text. For the purpose of the present study, the term word recognition is used to refer both to words read by phonological decoding (i.e., using knowledge of letter-sound correspondences to ‘sound out’ a word) and words read automatically by sight (i.e., quick and accurate recognition that is attained after repeated experiences with a word; Language and Reading Research Consortium, 2015).

To extract and construct meaning from a text requires skill in both word recognition and linguistic comprehension, such that the relation between these and reading comprehension is multiplicative. Interestingly, the reliance on each component changes across development. In the early grades, when children are learning how to decode, reading comprehension depends heavily on, and is constrained by, word recognition. As word recognition becomes automated and more fluent, linguistic comprehension skills become more important for comprehension (Catts, Hogan, & Adlof, 2005; García & Cain, 2014; Hoover & Gough, 1990; Language and Reading Research Consortium, 2015). There is strong evidence of these developmental changes in the relation between the components of the simple view and reading comprehension. In a meta-analysis involving 42,000 readers aged 5 to 53, García and Cain (2014) found a decreasing correlation between word recognition and reading comprehension over time.
Using multiple regression analysis in a longitudinal study, Catts et al. (2005) found that the percent of unique variance in reading comprehension explained by word recognition decreased across grades 2, 4, and 8 (from 27% to 13% to 2%, respectively), whereas the unique contribution of linguistic comprehension increased across these grades (9%, 21%, 36%). The amount of shared variance in reading comprehension explained by these variables remained relatively stable (40%, 39%, 36%). The Language and Reading Research Consortium (2015) used latent variables and structural equation modeling to model the simple view of reading in a cross-sectional study of children in grades 1, 2 and 3. The use of latent variables decreases measurement error and therefore provides a better assessment of the construct(s) of interest. The structure coefficient for word recognition decreased from .81 in first grade to .48 in second and third grade. The structure coefficient for listening comprehension increased from .22 in first grade to .57 in second and .60 in third grade. An important finding of this study was that the shift from word recognition to listening comprehension within the simple view occurred at second grade (Language and Reading Research Consortium, 2015).

There is a great deal of empirical support for the simple view of reading and the contribution of its components to reading comprehension (e.g., de Jong & van der Leij, 2002; Kendeou, van den Broek, White, & Lynch, 2009; Storch & Whitehurst, 2002; Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009), although some critics have suggested that it overly simplifies the process of comprehending text (e.g., Goswami, 2008; John R. Kirby & Savage, 2008; Tunmer & Chapman, 2012). For example, researchers have investigated whether fluency and vocabulary should be included in the
simple view (Adlof, Catts, & Little, 2006; Language and Reading Research Consortium, 2015; Ouellette & Beers, 2010). Furthermore, a fundamental tenet of the simple view of reading, that word recognition and listening comprehension contribute to reading comprehension independently, has been challenged by recent studies showing evidence of indirect effects of vocabulary, a component of linguistic comprehension, on word recognition (Language and Reading Research Consortium, 2015; Protopapas et al., 2012; Reed, Petscher, & Foorman, 2016; Tunmer & Chapman, 2012). The results of Tunmer and Chapman’s (2012) structural equation model, after a misspecification was corrected by Wagner and colleagues (2015), showed that a model which included an indirect path from linguistic comprehension (vocabulary) to word recognition provided an equally good fit to the data as the simple view model. This provides support for an alternative interpretation of the simple view in which word recognition and linguistic comprehension are not assumed to be independent. The role of vocabulary within the simple view will be further discussed below.

**Lexical quality and the reading systems framework.** The lexical quality hypothesis provides a framework for understanding word recognition and its role in reading comprehension (Perfetti, 2007; Perfetti & Hart, 2002). The lexical quality hypothesis posits that efficiency in word recognition is dependent upon high-quality lexical representations. The degree to which representations in phonology, orthography, morphosyntax, and semantics are established and integrated for a word determines how quickly and accurately that word will be read. High quality representations are well specified and connected, leading to efficient (i.e., rapid and reliable) word recognition
and, subsequently, better comprehension. The lexical quality hypothesis is consistent with connectionist triangle models which highlight the interplay among phonology, orthography, and meaning in word recognition (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989).

The reading systems framework, recently proposed by Perfetti and Stafura (2014), draws from and expands on the lexical quality hypothesis by situating word knowledge within a broad view of reading comprehension. Generally speaking, the reading systems framework consists of knowledge sources (linguistic, orthographic, and general knowledge) and the cognitive and linguistic processes that act upon them. Processes of reading include decoding, word identification, meaning retrieval, sentence parsing, inferencing, and comprehension monitoring. Within the reading systems framework, the word identification system and the word comprehension system support text comprehension, with the lexicon (word meanings) serving as the connection point between the systems. Once a word is identified during reading, its meaning is retrieved and then integrated into the reader’s mental model of the text. These word-to-text integration processes allow the reader to refine and update their understanding as they read through a text, with word meanings being integrated into sentences, sentences into paragraphs, and so on. As Perfetti and Stafura state, word comprehension is both “the output of the word identification system and the input to the comprehension systems” (Perfetti & Stafura, 2014, p. 32). Thus, word knowledge plays a central role in reading comprehension and is considered a ‘pressure point’ for reading comprehension.

Individual differences in word knowledge are a major source of individual differences in
comprehension, such that children who have poor word knowledge are likely to also experience reading comprehension difficulties (Perfetti & Stafura, 2014).

Whereas the simple view of reading asserts dissociation between word recognition and linguistic comprehension, the lexical quality hypothesis focuses on connectivity between these components due to underlying lexical skill (phonology, orthography, and semantics) that relates to both. Consistent with evidence that word recognition and linguistic comprehension are not independent (e.g., Reed et al., 2016; Tunmer & Chapman, 2012), some researchers have argued in favor of a lexical quality view of reading comprehension. For example, based on their results of a study of the role of vocabulary within the simple view of reading in grades 3 to 6, Protopapas, Mouzaki, Sideridis, Kotsolakou, and Simos (2013) proposed that “it may be preferable to conceptualize a lexical skill domain, emphasizing interrelations rather than distinctions among components” (p. 198). Similarly, support for a lexical knowledge factor was found in an examination of the dimensionality of language and reading in grades 4 to 10 (Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015).

A central theoretical argument of the present study is that lexical quality may be better indexed by spelling than by word recognition (Perfetti, 2007). Spelling requires highly specified orthographic representations that are well connected to phonological and semantic representations. Word recognition, on the other hand, can occur in the absence of precise lexical representations. The reader can rely on partial decoding (i.e., sounding out part of the word) and semantic cues (i.e., top-down contextual support) to identify words but this strategy is insufficient for accurate spelling (Frith, 1980). Thus, some
children who appear to be adequate readers are poor spellers because of underlying weakness in lexical representations. Similar results have been found in adults (Andrews & Bond, 2009; Hersch & Andrews, 2012). Given that spelling appears to be a better indicator of lexical quality, it follows that spelling may also be a more sensitive indicator of existing or future reading comprehension difficulties.

Lexical-Level Components of Reading Comprehension

Consistent with a lexical quality framework, the focus of the present study is on the role of lexical-level literacy skills in reading comprehension. Lexical-level literacy skills include word recognition and vocabulary, as highlighted in the theories presented above, as well as spelling (Bowers, Kirby, & Deacon, 2010). Empirical support for the role of word recognition and vocabulary in reading comprehension is presented next. An in-depth discussion of spelling and its role will be presented in subsequent sections.

Word recognition. The definition of word recognition used in the present study includes both decoding and sight word reading. Decoding refers to the process of “sounding out” a word, that is, using knowledge of phoneme-grapheme correspondences to retrieve and pronounce a sound for each grapheme in the word and then blending the sounds together to pronounce the word. Both pseudowords (pronounceable nonwords) and unfamiliar real words require decoding, whereas familiar words can be read from memory by sight. Sight word reading is an automatic process whereby the pronunciation of a word is rapidly retrieved from the visual input (Ehri, 2005). This occurs as a result of the ability to distinguish phonemes within words and to map phonemes onto graphemes (i.e., phonemic awareness and alphabetic knowledge). With repeated exposures, the
spelling of a word is secured in memory and linked to its pronunciation and meaning.
Ehri (2014) refers to this process as orthographic mapping.

Many studies support the essential role of word recognition in reading comprehension (Catts, Nielsen, Bridges, & Liu, 2014; García & Cain, 2014; Kendeou et al., 2009; Verhoeven & van Leeuwe, 2008). Results of meta-analyses have shown correlations ranging from .40 (in young children; National Early Literacy Panel, 2008) to .74 (across ages 5-53; García & Cain, 2014). In a longitudinal study of early identification of reading comprehension difficulties, Catts et al. (2014) found that kindergarten word recognition predictors (letter knowledge, phonological awareness, rapid naming, and nonword repetition) and second grade word recognition significantly predicted second grade reading comprehension, assessed with two measures, the Reading Comprehension Measure (RCM) and the Measures of Academic Progress: Reading (MAP). Specifically, for the RCM, the kindergarten logistic regression model accounted for 43.9% (Nagelkerke pseudo-$R^2$) of the variance on the RCM and demonstrated a high level of predictive accuracy based on the area under the curve (AUC) statistic of .873. For the MAP, the model accounted for 25% of variance and had an AUC of .774. The second grade model performed less well for the RCM (accounting for 21.7% of the variance with an AUC of .756), but was more accurate for the MAP (accounting for 42% of the variance with an AUC of .871).

As children become more experienced with reading and the orthographic system, the ability to decode and identify words becomes increasingly automatic, which is commonly referred to as reading fluency. Fluent word reading is characterized by both
accurate and rapid recognition of words. Whereas a typical first grader might read a text slowly and with effort, by third grade many children can read texts fluently and with ease and efficiency. Being able to read fluently is important to reading comprehension, as fluent reading frees up cognitive resources for comprehension processes (Adams, 1990; Perfetti, 2007). There is theoretical support for the notion that word recognition fluency is important for reading comprehension. An influential model of information processing by LaBerge and Samuels (1974) proposed that learning occurs at two levels of performance: accuracy and automaticity. Importantly, attentional resources are consumed at the accuracy level but not once automaticity has been attained. Perfetti’s verbal efficiency theory also states that slow and effortful decoding is a major factor in reading comprehension difficulties (Perfetti, 1985). An association between word recognition fluency and reading comprehension is also supported empirically (e.g., Foorman, Herrera, Petscher, Mitchell, & Truckenmiller, 2015; Silverman, Speece, Harring, & Ritchey, 2013).

There has been uncertainty in the literature regarding the manner in which word recognition is best indexed, particularly within the simple view of reading (Protopapas et al., 2012). Questions concern the content of the materials (real words vs. pseudowords) and the nature of the assessment (accuracy vs. fluency measures). Although some studies have found real word and pseudoword reading to have similar correlations with reading comprehension (Protopapas, Sideridis, Mouzaki, & Simos, 2007), other research suggests that real word reading is more predictive than pseudoword reading (García & Cain, 2014). In a study examining the factor structure of word recognition, Protopapas and
colleagues (2012) found that real and pseudoword reading represent a single construct whereas accuracy and fluency are separable constructs. Particularly relevant to the present study, the Language and Reading Research Consortium (2015) found that word recognition fluency (encompassing both real and pseudowords) predicted reading comprehension in third graders, but word recognition accuracy did not.

**Vocabulary.** There is a large body of evidence on the importance of vocabulary in reading comprehension (Muter, Hulme, Snowling, & Stevenson, 2004; NICHD Early Child Care Research Network [ECCRN], 2005; Oakhill & Cain, 2012; Ouellette, 2006). Moderate longitudinal correlations have been reported between vocabulary and reading comprehension (Muter et al., 2004), and results of multiple regression analyses have found that vocabulary contributes unique variance to reading comprehension when controlling for other known predictors such as phonological awareness (Roth, Speece, & Cooper, 2002; Sénéchal, Ouellette, & Rodney, 2006). In a study of early identification of reading comprehension difficulties, Catts et al. (2014) reported that a second grade screening tool was a more accurate predictor of third grade reading comprehension when a measure of vocabulary was included. In an effort to answer questions about directionality and causality in the relation between vocabulary and reading comprehension, Quinn, Wagner, Petscher, and Lopez (2015) conducted a study using latent change score modeling to examine the codevelopment of these skills from first through fourth grade. This analytic procedure allows for modeling of dynamic development relations, providing insight into causal influences between constructs.
Results showed that vocabulary had a leading influence on reading comprehension, supporting a causal role for vocabulary in the development of reading comprehension.

Vocabulary can be conceptualized as knowledge across three dimensions. Breadth refers to the size of the lexicon – the sheer number of words known. Depth refers to how well words are known; for example, the ability to define words and knowledge of relations amongst words. Fluency refers to the speed of retrieval of word meanings. Each of these dimensions of vocabulary appears to be associated with reading comprehension. Roth et al. (2002) reported moderate to strong correlations between kindergarten measures on each of the dimensions and reading comprehension in first and second grade. In a multiple regression with kindergarten predictors, the ability to provide word definitions accounted for 3% unique variance in first grade reading comprehension and 9% of unique variance in second grade reading comprehension. Additionally, word retrieval accounted for 6% of unique variance in second grade reading comprehension (Roth et al., 2002). Ouellette (2006) examined the contribution of vocabulary breadth and depth to reading comprehension in a sample of fourth grade students. Results indicated that receptive vocabulary breadth and vocabulary depth each explained unique variance (6% and 8%, respectively) in reading comprehension after accounting for age, nonverbal intelligence, pseudoword decoding and visual word recognition. To determine the factor structure of the vocabulary dimensions, Tannenbaum and colleagues (Tannenbaum, Torgesen, & Wagner, 2006) conducted a factor analysis and found that a two-factor model, breadth and depth/fluency, best fit the data. These factors were highly correlated with each other and together explained 50% of the variance in reading comprehension.
among third graders. Breadth accounted for more unique variance (19%) than depth/fluency (2%). Together these studies suggest that vocabulary consists of multiple dimensions and breadth appears to be most strongly associated with reading comprehension. In the present study, vocabulary is assessed using a measure of receptive vocabulary breadth.

Although the importance of vocabulary to reading comprehension is generally well-accepted, there is debate in the literature about its role in the simple view of reading; specifically, whether it has a direct influence on reading comprehension, independent from word recognition and linguistic comprehension, or whether it plays an indirect role through word recognition and/or linguistic comprehension (Language and Reading Research Consortium, 2015; Ouellette & Beers, 2010; Protopapas et al., 2013; Protopapas et al., 2012; Tunmer & Chapman, 2012). Studies employing regression analyses have found that vocabulary explains unique variance in reading comprehension after accounting for word recognition and linguistic comprehension (Braze, Tabor, Shankweiler, & Mencl, 2007; Ouellette & Beers, 2010; Protopapas et al., 2013; Tunmer & Chapman, 2012; Verhoeven & van Leeuwe, 2008). For example, Ouellette and Beers (2010) found that vocabulary explained unique variance in reading comprehension beyond what was explained by phonological awareness, word recognition, and listening comprehension for children in grade 6. Vocabulary was not a unique predictor in grade 1, however, suggesting that its role may increase across grades. Extending this research by using multiple indicators of each construct, Braze et al. (2015) found similar results with adults.
Recent studies have employed a latent variable approach and structural modeling to further examine the role of vocabulary in reading comprehension (Braze et al., 2015; Protopapas et al., 2013; Tunmer & Chapman, 2012). Results suggest that vocabulary is not a separate factor within the simple view of reading; rather, it is part of linguistic comprehension. In a study of 7-year olds \( (n = 122) \), Tunmer and Chapman (2012) tested the simple view of reading model by first conducting an exploratory factor analysis with measures of pseudoword decoding, context-free real word recognition, letter-sound knowledge, vocabulary, listening comprehension, and reading comprehension. The pseudoword decoding and word recognition measures loaded onto a decoding factor and vocabulary and listening comprehension loaded onto a linguistic comprehension factor. Results of structural equation modeling found that, in addition to direct paths from decoding and linguistic comprehension to reading comprehension, a path from decoding to linguistic comprehension was also significant. The authors concluded that linguistic comprehension, of which vocabulary is a component, influenced reading comprehension both directly and indirectly through decoding.

The Language and Reading Research Consortium (2015) conducted a study in which reading comprehension, listening comprehension, vocabulary, and word recognition were assessed with multiple measures. Structural equation modeling was used to test direct and indirect effects of vocabulary on reading comprehension in a typically developing sample of children in grades 1 to 3 \( (n = 371) \). Across each grade, vocabulary had significant indirect effects on reading comprehension through both word recognition and listening comprehension, with slightly stronger effects on listening.
comprehension than on word recognition. There was no direct effect of vocabulary on reading comprehension. Together, these studies suggest that vocabulary influences reading comprehension indirectly through both linguistic comprehension and word recognition.

To summarize, reading comprehension is a complex process that can be conceptualized through frameworks such as the simple view of reading, which focuses on independent contributions of print-dependent and print-independent components, and the lexical quality hypothesis, which highlights word knowledge characterized by integrated phonology, orthography, and semantics. Drawing on the lexical quality framework, the present study is focused on lexical-level skills and their role in reading comprehension, with a primary aim of investigating the unique role that spelling may play. As described above, there is a substantial body of literature on the role of word recognition and vocabulary, but the relation between spelling and reading comprehension has received little attention in the literature. Before turning to this research on spelling and reading, the next section will first provide an overview of spelling process and theory.

**The Spelling Process**

Whereas the language basis of reading has been established for quite some time, it is relatively recent that the developmental and linguistic nature of spelling has been widely recognized. Seminal work conducted by researchers such as N. Chomsky and Halle (1968), C. Chomsky (1970), and Venezky (1967, 1970) revealed that the orthographic system of English is highly regular, more so than had previously been recognized. Venezky (1999) documented 200 regularities in English spelling, including,
for example, that the <ow> spelling can occur at the beginning of words before the consonants <d, l, n>, or that <ck> occurs only after single vowel letters. Around the same time, researchers began investigating children’s invented spellings and found them to be consistent, evolving over time, and aligned with children’s knowledge of language (Beers, Beers, & Grant, 1977; Beers & Henderson, 1977; Bissex, 1980; Read, 1971). The pioneering descriptive work by Read (1971) led to two important shifts in our understanding about spelling; first, that spelling is a developmental process, and second, that children’s spellings provide insight into their knowledge of the writing system. The errors that children make in their spellings are logical, not random as previously thought, and they reflect children’s developing concepts about how sounds, patterns, and meaning (morphology) are represented in writing (Invernizzi, Abouzeid, & Gill, 1994). Recent quantitative research added further evidence that invented spelling is developmental and indicative of underlying linguistic skills such as phonemic awareness and orthographic knowledge (Ouellette & Sénéchal, 2008). A meta-analysis conducted by the National Early Literacy Panel (National Early Literacy Panel, 2008) found that children’s early invented spellings were highly correlated with decoding in kindergarten, first, and second grade (r ranging from .49 to .63).

Learning to spell. Broadly speaking, children progress from spelling based on meaning and the representation of sounds (phonology) to applying orthographic and morphological knowledge. Treiman and Bourassa (2000) reviewed the research on spelling development and described a progression from pre-alphabetic to alphabetic, orthographic, and morphological. Children’s earliest spelling attempts represent semantic
characteristics, for example the word _snake_ would be longer than the word _butterfly_. As children learn the alphabetic principle, that letters represent the sounds of spoken words, their spellings reflect the phonology of words. Their spellings at this stage are often plausible but unconventional; for example they may spell the word _sick_ as _<sik>_.

Children’s knowledge of the orthographic system becomes more sophisticated as they gain experience with print and orthographic patterns. For example, they apply knowledge of the <ck> pattern in words like _sick_. Finally, children learn and apply knowledge of the morphological structure of words. They develop an awareness of the morphological elements in a spoken word (e.g., a base plus suffix), they learn the conventional spellings of prefixes and suffixes, and they apply this knowledge in their spelling. For example, a child who was aware of the morphological structure of the word _jumped_ would spell the suffix correctly even though it sounds like /ʃ/.

**Theory of Spelling Development**

Spelling development has been framed within several theoretical views. Stage theory proposes that spelling development is sequential. Children pass through discrete developmental stages of learning the writing system, relying on certain types of linguistic knowledge in each stage (e.g., Adams, 1990; Ehri, 2000; Gentry, 1982). Treiman and Bourassa (2000) concluded that, although stage models provide a general overview, they are unable to capture the complexity of the phonological, orthographic, and morphological systems in spelling. More recent theories have moved away from the view that children use one source of linguistic knowledge to the exclusion of others. Instead, overlapping waves or repertoire theory, also called triple word form theory, proposes that
children concurrently draw upon the full repertoire of linguistic skills available to them across all points in spelling development (Masterson & Apel, 2007; Rittle-Johnson & Siegler, 1999). Whereas stage theory proposes that orthographic and morphological knowledge are used only in later stages of development, research shows that children do in fact apply these linguistic skills very early in development (Apel, Wilson-Fowler, Brimo, & Perrin, 2012; Kim, Apel, & Al Otaiba, 2013; Treiman & Cassar, 1996).

Connectionist triangle models (Foorman, 1994; Houghton & Zorzi, 2003; Seidenberg & McClelland, 1989) align with repertoire theory but expand on the interplay between phonology, orthography, and meaning by also taking word and pattern frequency into account. Connectionist theories are typically implemented as computer models in which the connections between phonology, orthography and meaning are weighted according to frequency and regularity and it is these weighted connections that drive the system. The spelling that a child associates with a pronunciation depends on the frequency with which the orthographic pattern has been encountered; thus spelling is also viewed as statistical learning (Treiman & Kessler, 2006). Within this view, it is proposed that children take the surrounding context, such as adjacent letters, into account when choosing between alternative spellings for a sound.

**Linguistic Components of Spelling**

Spelling is a complex linguistic process that, like word recognition, draws on multiple sources of linguistic knowledge, including phonological awareness, orthographic knowledge, and morphological knowledge (Deacon, 2012; Kim, Apel, et al., 2013). Spelling a word involves identifying the phonemes and, if applicable, morphemes in a
word, and representing those phonemes and morphemes using orthographic conventions. Often, spelling also requires choosing the correct orthographic pattern from a number of possibilities. Analyzing the types of errors that children make in their spelling can provide insight into their developing competency in these underlying linguistic systems. Spelling assessments that take a linguistic error analysis approach have been developed for this purpose and may therefore provide a more sensitive measure of spelling development than traditional correct/incorrect spelling scoring metrics (Apel & Masterson, 2001; Masterson & Apel, 2010).

Phonological awareness. Phonological awareness is the awareness of and ability to manipulate the sound structures of language. It includes skills such as rhyming, alliteration, and syllable blending and segmenting. The term phonemic awareness refers specifically to manipulations at the phoneme level. Children use their phonemic awareness to identify the sounds that need to be represented in spelling a word. The ability to segment words into sounds is a good predictor of early spelling (Nation & Hulme, 1997; National Early Literacy Panel, 2008). Phonological spelling errors include omitting sounds, e.g., spelling <bt> for boat and <srap> for strap, and inserting sounds, e.g., spelling <selam> for slam.

Orthographic knowledge. Orthography refers to the conventional writing system of a language. Knowledge of how this system works is necessary for accurate decoding and encoding of written language (e.g., Apel, Wolter, & Masterson, 2006; Conrad, Harris, & Williams, 2013; Ehri, 2005). There are two types of orthographic knowledge (Apel, 2011; Conrad et al., 2013). The first is general knowledge of orthographic
patterns, or rules, that apply to all words. This includes knowledge of basic phoneme-grapheme correspondences, e.g., the sound /f/ can be spelled with <f, ff, ph>; allowable letter combinations, e.g., <ck, shr> are allowed in English but <nr, fp> are not; and orthotactic rules, which are contextual and positional constraints for which letters can be used where, e.g., <ck, ld> are allowable combinations but never at the beginning of a word. This type of orthographic knowledge relates to the concept of *legality* in spelling. Spellings are considered legal if they are orthographically plausible. For example, *rich* spelled as *ritch* is incorrect but legal because the <tch> grapheme is an allowable spelling at the end of words following a short vowel. Phonetic plausibility may also be considered. For example, the word *drink* spelled as <jrink> is phonetically plausible since the sound of <j> is articulated when saying *drink*. However, <jr> is not an allowable letter combination in English, so this spelling is considered illegal based on orthographic conventions.

The second type of orthographic knowledge is word specific. It requires knowledge not only of legal spelling patterns but also of which patterns are correct in specific words; for example, knowing that the word *rain* is spelled <rain>, not <rane>. The spelling <rane> is orthographically legal but is not the correct representation for that word. This knowledge has been referred to using a variety of terms (Apel, 2011) such as mental graphemic representations (MGRs; Apel, 2009) and sight words (Ehri, 2005). In this paper, the term MGR will be used.

**Morphological knowledge.** Morphemes are the smallest units of language that convey meaning. Morphology, therefore, refers to the structure of words. Words are
comprised of either a single morpheme, referred to as the base element, or a base with an affix or affixes. Affixes include prefixes and suffixes and are classified as inflectional or derivational. Inflectional affixes are suffixes that relate to the grammatical function of a word and always have a regular meaning, such as past tense -ed and comparative -er. Derivational affixes are prefixes or suffixes that change the grammatical class of a word, such as the suffix -able which changes a verb to an adjective (e.g., walk-walkable). Words that are derived from the same base generally retain the sense of the base, albeit modified by the affix (e.g., appear-disappear, shape-misshapen).

Morphological knowledge includes awareness of morphemic elements in words, e.g., that vacation consists of the base vacate and the suffix -ion; knowledge of the meanings of the morphemic elements, e.g., that the prefix -un means not; an understanding of the meaning relations between words and their inflected and derived forms, e.g., the words tractor, traction, and extract are all derived from the bound base tract and therefore share a relation to the meaning “to pull or draw”; and knowledge of spelling rules for adding suffixes. Morphological spelling rules apply at the juncture of the base and the suffix and either modify the spelling of the final letter in the base, as in the rules for dropping a silent -e or changing <y> to <i>, or require a doubling of the final letter before adding the suffix. Correct spelling of multimorphemic words, then, requires knowledge of affixes and correct application of suffixing rules. Examples of morphological errors include: spelling jumped as <jumpt>, which indicates a lack of awareness of the past tense suffix in the word and/or knowledge that the past tense suffix is spelled with -ed even when it sounds like /t/; happiness spelled as <happyness>, which
indicates incorrect use of a suffixing rule; and *magician* spelled as <magishan>, which indicates lack of awareness that the base is *magic* and/or knowledge that the spelling of the base is retained when a suffix is added, despite the pronunciation shift.

Morphological knowledge may be particularly important, however, as children progress through the grades and encounter a greater number of multimorphemic words. The list of academic vocabulary words, for example, consists largely of morphologically complex words (Coxhead, 2000; Nagy, Townsend, Lesaux, & Schmitt, 2012).

**Spelling and Reading: Theoretical Rationale**

An association between spelling and reading is supported both theoretically and empirically. Much is known about the relation between spelling and word recognition but little attention has been paid to the role of spelling in reading comprehension. The main hypothesis of the present study is that spelling makes an important contribution to reading comprehension, perhaps more so than word recognition. As argued by Kim, Petscher, and Foorman (2013), spelling is informative as a lexical-level predictor of reading comprehension because it is indicative of underlying linguistic skills and requires greater precision and memory than word recognition. In this section I will first present the theoretical rationale for a relation between spelling and reading comprehension, followed by a discussion of the research.

**Lexical quality.** The primary theoretical rationale for the present study is the lexical quality hypothesis, which emphasizes the role of word knowledge in reading (Perfetti & Stafura, 2014). The constituents of word knowledge include representations of phonology, orthography, and morphosyntax. The degree to which these representations
are established in memory and integrated with one another determines how quickly and accurately words will be read (Perfetti, 2007; Perfetti & Hart, 2002). The richer the quality of the representations and the more closely the representations are connected, the more efficient word retrieval will be. Once words are identified they can be understood, thereby initiating the process of reading comprehension. Thus, high lexical quality drives efficient word identification which facilitates reading comprehension. Importantly, lexical quality may be better indexed by spelling than by word recognition (Perfetti, 1992) since accurate spelling of words requires precise, highly specified orthographic representations that are closely connected with phonological and semantic representations. Accurate recognition of words, however, can happen even in the absence of high quality phonologic and/or orthographic representations when context can be used to “guess” a word. Spelling therefore may be a stronger indicator of underlying lexical quality than word recognition and thus potentially a better predictor of reading comprehension.

**Orthographic learning.** Another theoretical perspective, that of orthographic learning and the self-teaching hypothesis, also supports an association between spelling and reading comprehension. Orthographic learning refers to the acquisition of orthographic knowledge and involves the integration of phonological, orthographic, and semantic representations (Ouellette 2010). Self-teaching is a process by which orthographic learning occurs - orthographic knowledge is gained as a result of feedback in attempts to read or spell unfamiliar words (Shahar-Yames & Share, 2008; Share, 1995). For example, in attempting to decode an unfamiliar word, known letter-sound
patterns within the word can be used to attempt a pronunciation, and correct attempts provide an opportunity for the reader to learn any previously unfamiliar letter-sound patterns in the word. New patterns are stored in memory and available for use in subsequent attempts to read or spell unfamiliar words. Thus, self-teaching is a mechanism by which knowledge of orthographic representations is acquired. This knowledge is important for the development of word recognition and therefore also for reading comprehension.

Research suggests that spelling may be superior to reading as a self-teaching mechanism for orthographic learning (Ouellette, 2010; Shahar-Yames & Share, 2008). Shahar-Yames and Share (2008) conducted a study of Hebrew-speaking children in third grade to compare orthographic learning via reading as compared to spelling. Results showed that orthographic learning occurred in both conditions but was stronger in the spelling condition than in the reading condition. The authors concluded that it is the process of print-to-sound translation that is important for orthographic learning to occur and, although that translation happens during both reading and spelling, spelling appears to be superior. These results were replicated in a similar study with English-speaking children in second grade (Ouellette, 2010). Ouellette concluded that spelling offers a ‘superior milieu’ for orthographic learning to occur.

Together, these theoretical frameworks lend support to the hypothesized relation between spelling, as an indicator of lexical quality and orthographic learning, and reading comprehension. In the next section, I turn to the research base for what is currently known about spelling and its relation with word recognition and reading comprehension.
Spelling and Word Recognition

Compared to the expansive research base for reading, spelling has received far less attention in the literature. In terms of the work that has been done, the association between spelling and word recognition is the most studied. Spelling is closely related to word recognition. In some ways, these skills are ‘two sides of a coin’, sharing similarities in acquisition and the knowledge sources that they depend on (Ehri, 2000). High correlations, typically above .70, have been reported in the literature (e.g., Abbott et al., 2010; Ehri, 2000; Juel, Griffith, & Gough, 1986; Morris et al., 2012).

A reciprocal relation between spelling and word recognition is supported in studies comparing unidirectional models, i.e., spelling influencing reading or reading influencing spelling, to an interactive model in which both skills influence each other (Shanahan & Lomax, 1986, 1988). The interactive model provided a better fit than the unidirectional models. Reciprocity between spelling and word recognition is further supported by research showing that instruction in spelling improves word recognition, and vice versa (e.g., Conrad, 2008). Based on their meta-analyses, Graham and colleagues reported effect sizes of .62 and .40 for spelling instruction on word recognition (Graham & Hebert, 2011; Graham & Santangelo, 2014).

Despite the many similarities between word recognition and spelling, however, they are not simply mirror images of one another. Spelling is a more difficult skill than word recognition, due to the nature of the task itself and the depth of English orthography (as described in Chapter 1). The memory demands of spelling are greater than those of reading (Ehri, 2000) Spelling is a production task and is therefore more difficult than
reading which is a recognition task. Correct spelling of a word requires the complete, precise representation of the phonology, orthography, and morphology of the word (Ehri, 2000; Kim, Petscher, et al., 2013; Perfetti, 1997). In contrast, reading does not necessarily require such precision. When reading unfamiliar words in text, children can use a partial decoding strategy in which context facilitates retrieval of the word’s pronunciation even if the decoding attempt is not completely accurate. In a study of twelve-year olds, Frith (1980) concluded that children who relied on a partial decoding strategy were good readers but poor spellers. In other words, incomplete orthographic representations were sufficient for adequate reading but not for adequate spelling. Individuals with strong skills in both reading and spelling (‘lexical experts’), in comparison to those with strong reading only, have more accurate context-independent word recognition skill and therefore have more cognitive resources freed up for comprehension (Hersch & Andrews, 2012). Other studies have shown that good spellers are usually good readers but the reverse is less likely to be true (Fayol, Zorman, & Lété, 2009). In sum, research suggests that spelling requires more precision than word recognition and therefore may be a stronger index of underlying linguistic knowledge (Ehri, 2000; Perfetti, 1997).

**Spelling and Reading Comprehension**

Given that spelling and word recognition are closely related and given that, theoretically, spelling may be a stronger indicator of lexical quality and orthographic learning, both of which facilitate reading comprehension, it follows that spelling may be an important predictor of reading comprehension. The print-dependent component of the simple view of reading can be thought of as comprising multiple, interconnected sources
of lexical knowledge that permit learning of the English orthographic system and thus the
development of word recognition and spelling skills and, subsequently, reading
comprehension. Within the framework of the lexical quality hypothesis, word knowledge
is a key factor in reading comprehension. Word knowledge comprises not only print-
based representations of words, i.e., phonology, orthography, and morphology, but also
semantic representations, i.e., vocabulary. As described in a previous section, there is a
substantial body of literature supporting the role of both word recognition and vocabulary
in reading comprehension. In contrast, however, much less is known about the potential
relation between spelling and reading comprehension, particularly within the context of
other lexical-level literacy skills, i.e., word recognition and vocabulary. I will review the
existing literature on this topic next.

Correlations between spelling and reading comprehension are typically reported
in the moderate to strong range. For example, Abbott et al. (2010) reported concurrent
correlations of .76 to .43, decreasing across grades 1 to 7. Kim, Petscher, et al. (2013)
found relatively stable correlations across grades 3 to 10, ranging from .45 to .58
concurrently and longitudinally (within grade; fall spelling to spring reading
comprehension). The longitudinal correlations were slightly higher than the concurrent
reported a correlation of .45 between spelling fluency (a timed measure of spelling) in
grade 4 and reading comprehension in grade 6. In comparison, the correlation between
word recognition fluency and reading comprehension was .33 to .41.
Much of the existing research on the relation between spelling and reading comprehension has not included word recognition or vocabulary measures. For instance, Foorman, Petscher, and Bishop (2012) examined the contribution of spelling to reading comprehension in grades 3 to 10 while controlling for the autoregressor (winter reading comprehension). Winter spelling scores accounted for 0-5% of variance at the student level in spring reading comprehension. In grade 3, the pseudo-$R^2$ for the autoregressor alone was .36 and with spelling added it was .41. Berninger, Abbott, Abbott, Graham, and Richards (2002) used structural equation modeling to examine the contribution of spelling to reading comprehension in the presence of handwriting automaticity in grades 1 to 6. The path between spelling and reading comprehension was significant in all grades except grade 5.

Foorman and colleagues (Foorman & Petscher, 2010) conducted another study in which they examined the contribution of spelling to reading comprehension in the presence of silent text reading fluency (a maze task – a timed test in which children read text and select the correct word from maze boxes placed at designated intervals throughout the text). In grades 3 to 12, classrooms in which children had low spelling scores also demonstrated low maze and reading comprehension scores. In other words, high and low spelling ability classrooms were differentiated on both text reading fluency and reading comprehension, with large effect sizes for reading comprehension.

Similarly, Kim, Petscher, et al. (2013) sought to determine the unique contribution of spelling to reading comprehension in grades 3 to 10 while controlling for silent text reading fluency (maze). At the student level, fall spelling scores accounted for
1-3% of unique variance in spring reading comprehension after accounting for fall reading comprehension and silent reading fluency, as compared to 2-7% unique variance accounted for by silent reading fluency. The total variance accounted for by all three predictors ranged from 28-45% at the student level.

In a study that controlled for word recognition, Desimoni, Scalisi, and Orsolini (2012) examined concurrent and predictive relations between spelling and reading comprehension in a sample of Italian children in grades 1 and 3. Both word recognition accuracy and speed were measured, but within the context of reading a text as opposed to reading words in isolation. In the concurrent analysis, results of hierarchical regressions showed that spelling accounted for 21.3% of variance in reading comprehension in grade 1 and 19.4% in grade 3 after controlling for age (which was significant in grade 1 but not in grade 3). Controlling for age as well as word reading accuracy and speed, spelling accounted for an additional 3.5% of variance in reading comprehension in grade 1 and 9.2% in grade 3. The unique variance accounted for by spelling was significant in all instances. In comparison, the unique contribution of word recognition was significant in grade 1 but not in grade 3. The authors suggested that these results may be explained by the self-teaching hypothesis; that spelling supports the acquisition of orthographic knowledge more strongly than does word recognition, which in turn facilitates automatic, efficient word recognition and reading comprehension (Shahar-Yames & Share, 2008). This study did not account for vocabulary and was also limited in that it measured reading comprehension with only a single narrative text. A single text is not likely to give a reliable measure of reading comprehension, especially given the importance of
background knowledge for reading comprehension (Langer, 1984; Pearson, Hansen, & Gordon, 1979).

Looking at longitudinal contributions of spelling and word recognition with reading comprehension, Desimoni et al. (2012) conducted multivariate regression and found that spelling in grade 1 was not a significant unique predictor of reading comprehension in grade 3 when controlling for grade 1 reading comprehension, reading errors, and reading speed. Mixed findings of a longitudinal relationship between spelling and reading comprehension were reported by Abbott et al. (2010). Results of longitudinal path analyses across adjacent grades from grade 1 through 7 indicated that the path from spelling to reading comprehension was significant only from grade 2 to 3 and grade 4 to 5, with the latter having a small but negative coefficient.

The unique role of spelling in the presence of vocabulary was studied by Reed et al. (2016). The results of hierarchical regression analyses showed that spelling made a unique and increasing contribution to reading comprehension across grades 6 to 10, from 2% in grade 6 to 9% in grade 10. Vocabulary, measured with a written task that required knowledge of derivational forms of words, made a greater unique contribution than spelling but the contribution decreased across grades, from 19% in grade 6 to 11% in grade 10. Another finding of this study was that spelling partially mediated the relation between vocabulary and reading comprehension. The authors interpreted this finding as support for the alternative view of the simple view of reading presented by Tunmer and Chapman (2012), in which the language comprehension component has both direct and indirect effects on reading comprehension.
Some researchers have measured spelling using a recognition task, in contrast to the production tasks more commonly used. For instance, Katzir et al. (2006) used an orthographic choice task in which children were required to select the correct spelling from a set of four homophones. In this study of dyslexic and younger reading-matched typical readers, results of hierarchical regression analyses showed that spelling recognition was a unique and significant predictor of reading comprehension after controlling for reading disability status and phonemic awareness (in one model) and rapid letter naming and word reading fluency (in another model). Another type of spelling recognition task involves identifying spelling errors within a text. Retelsdorf and Köller (2014) conducted a longitudinal study in Germany to test the reciprocal effects between spelling and reading comprehension from grade 5 to grade 7. Controlling for cognitive ability, school track, parents’ occupational status, gender, and ethnicity, results of cross-lagged panel models indicated significant effects in both directions, but the effect from reading comprehension to spelling was greater than the effect from spelling to reading comprehension. The authors suggested that the effect of spelling on reading comprehension may be explained by the lexical quality hypothesis; strong spelling builds strong orthographic representations which support reading.

Recently, classification studies have examined whether spelling scores can accurately predict reading comprehension status (good vs. poor comprehenders). Using logistic regression, Ritchey et al. (2015) found that spelling fluency and reading fluency measured in grade 4 significantly predicted poor reading status in comprehension, decoding, or fluency in grade 6. A receiver operating curve analysis yielded an area under
the curve of .91, indicating highly reliable prediction. Using discriminant analysis, Desimoni et al. (2012) found spelling in grade 1 to have significant discriminatory power in predicting high and low literacy group membership in grade 3 (Partial Lambda = 0.92, \( p < .05 \)). Other research has shown that children classified as poor comprehenders in grade 4 performed below typical readers on measures of spelling between grade 2 and grade 7, and that kindergarten spelling ability was an early indicator of poor comprehension (Etmanskie, Partanen, & Siegel, 2016). Together, these results suggest that spelling may serve as a valuable screening tool for children at risk for reading comprehension difficulties.

In summary, there is a growing body of literature on the role of spelling in reading comprehension but most studies have either taken a univariate approach or have included only a single additional predictor (e.g., word recognition, vocabulary, or text fluency). Another limitation in the existing literature is that some of these studies were conducted in languages other than English. Orthographic complexity varies greatly across languages and therefore cross-linguistic comparisons should not be made without empirical evidence. There is no existing research of the role of spelling in the presence of multiple lexical-level literacy skills, that is, word recognition and vocabulary. It is important to understand the unique contributions of each of these skills as they have been shown to influence each other and reading comprehension. Within the framework of the lexical quality hypothesis, these skills comprise word knowledge which provides the foundation for reading comprehension (Perfetti, 2007; Perfetti & Stafura, 2014). The present study
therefore aims to explore the relations between lexical-level literacy skills and reading comprehension, with a focus on understanding the role of spelling in particular.

**Spelling Scoring Metrics**

Spelling can be measured using a variety of scoring metrics but whether the metric matters – in understanding the relation between spelling and reading comprehension, for example – is a question that remains to be answered. The most commonly used metric is the number of words spelled correctly, a quick and simple method of scoring that provides an index of children’s overall spelling accuracy. Other scoring metrics give credit for spellings that are partially correct; although a drawback of such metrics is that they require more training and more time to score. Two such metrics are correct letter sequences, commonly used in curriculum-based measures (Hosp & Hosp, 2003), and the Spelling Sensitivity System (Masterson & Apel, 2010, 2013), commonly used in measuring developmental changes in children’s spelling. Correct letter sequences are pairs of letters that are placed in the proper sequence within a word. Scoring is based on accuracy of letters only, so phonological plausibility of a spelling is not taken into account. The Spelling Sensitivity System is a fine-grained scoring metric that uses a multiple-linguistic approach to error analysis. Errors are analyzed according to whether they are related to phonemic awareness, orthographic pattern knowledge, morphological awareness, or storage of mental graphemic representations (MGRs). Two scores are obtained, one for the whole word and one for the individual elements in each word. These will be described in Chapter 3.
Only two studies have systematically compared spelling scoring metrics but neither has done so in exploring the relation between spelling and reading comprehension. In the first study, a variety of metrics (including words correct, correct letter sequences, sounds correct, and phonological coding) were found to have equivalent correlations with word recognition, phonological awareness, letter name fluency, nonsense word fluency, and writing skills in kindergarten (Ritchey, Coker, & McCraw, 2010). These correlations ranged from .73 to .78. The metrics were also highly correlated with one another (range .84 to .98). Of note is that the mean number of words correct for children in this study was 1.88 (range 0 to 8; maximum possible = 10). With such a low mean, it is possible that many children scored at floor and there may be insufficient variability on this metric for it to be statistically reliable.

In the second study, Clemens et al. (2014) used structural equation modeling to examine the concurrent (kindergarten) and predictive (kindergarten to grade 1) validity of spelling metrics to a variety of reading skills. Participants were children who had been identified as at risk for reading difficulties. The spelling metrics were the same as those used by Ritchey et al., with the addition of the spelling sensitivity word score. Two concurrent models were run, one in which spelling metrics predicted phonological awareness and pseudoword decoding factors, and the other in which spelling metrics predicted a reading factor consisting of word recognition accuracy and fluency and oral text reading fluency. Results indicated that the metrics that provided credit for partially correct spellings accounted for 6% to 14% more variance than the words correct metric. For the reading factor, words correct accounted for 70% of the variance, whereas the
metrics that gave credit for partially correct spellings accounted for 74-76% of the variance.

Results were similar in the longitudinal model, where the reading factor was comprised of word recognition accuracy, pseudoword decoding, oral text reading fluency, and passage comprehension. Words correct explained 39% of the variance whereas the metrics that gave credit for partially correct spellings accounted for 44-46% of the variance. These results therefore suggest that spelling metrics are similar but those that give credit for partially correct spellings may reveal a stronger relation between spelling and reading. Of note, however, is that when kindergarten word recognition was added to the model, the metrics all performed similarly, explaining 49-51% of the variance.

The Present Study

The purpose of the present study is to investigate the relation between spelling and reading comprehension in the presence of other lexical-level literacy skills, namely word recognition and vocabulary. The first aim of the study is to explore the bivariate relations among the three lexical-level literacy skills and reading comprehension, to gain an understanding of children’s abilities across these skills. The second aim is to explore the unique contribution of spelling to reading comprehension when accounting for word recognition and vocabulary. Results using four different scoring metrics will be compared.

The specific research questions of the study are as follows: (a) What are the bivariate, concurrent relations between reading comprehension, spelling, word
recognition, and vocabulary in third grade?, and (b) To what extent is spelling uniquely related to reading comprehension when accounting for word recognition and vocabulary, and does this relation vary depending on the spelling scoring metric used?

With regard to the first research question, I hypothesize that children will demonstrate consistency across skills but there may be a few children who show dissociations. For example, some children may be unexpected poor spellers (i.e., have good word recognition but poor spelling). If these cases exist, I hypothesize that the reading comprehension of these children will be lower than expected based solely on their word recognition skills. With regard to the second research question, I hypothesize that spelling will contribute unique variance to the prediction of reading comprehension over and above the variance predicted by word recognition or vocabulary. In addition, I hypothesize that spelling will be more highly correlated with and explain more unique variance in reading comprehension when using metrics that give credit for partially correct spellings, especially those that utilize in-depth linguistic error analysis based on phonology, orthography, and morphology. In other words, the finer the unit of analysis the stronger the relationship is expected to be.
Chapter 3: Methods

Participants

Participants were 63 children in third grade who represented a subset of children from a larger study by the Language and Reading Research Consortium (LARRC). The larger study was a five-year longitudinal cohort study, conducted 2010-2015, investigating the language basis of reading comprehension in children from preschool to third grade in four states. The subsample of children in the present study consisted of the participants from 35 classrooms in one Midwestern state who were in third grade in the final year of the study (2015).

Recruitment for the LARRC study began with districts and/or schools. Once permission was obtained at this level, all teachers in the relevant grades were invited to participate. Recruitment packets were then sent home with all children in the participating teachers’ classrooms. Packets contained information about LARRC, a caregiver consent form, and a family questionnaire. Family and teacher questionnaires were used to ensure that all consented children could adequately participate in the study, meaning that they spoke English fluently and did not have severe or profound disabilities. Additionally, preschool children were selected only if they were expected to matriculate to kindergarten in the following year. See Language and Reading Research Consortium, Farquharson, and Murphy (2016) for additional study details and recruitment procedures.
Participants in the present study had a mean age of 9.08 years (SD = .36 years; range 8.42 years to 10.00 years) at the time of testing and were primarily male (64%; n = 40). The majority of children spoke English as their primary language. One child spoke Spanish. Nonverbal intelligence, as measured with the Matrices subtest of the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1997), for this sample was average (M = 103.90, SD = 19.23; range 65-143). Fourteen percent of children (n = 9) had an Individual Education Plan (29% unreported). Demographic information was collected from caregiver questionnaires. With regard to highest level of maternal education, one mother had no high school diploma, 22% of the mothers (n = 14) had a high school diploma but no college, 24% (n = 15) had some college but no degree, 30% (n = 19) had a 2- or 4-year degree, and 14% (n = 9) had a graduate degree (8% unreported).

Information about family income was reported as follows: < $25 000 (11%, n = 7), $25 001-75 000 (21%, n = 13), > $75 001 (41%, n = 26), and unreported (27%). Nineteen percent (n = 12) qualified for free/reduced price lunch (30% unreported). Caregivers’ race and ethnicity was primarily Caucasian (84%, n = 53) and non-Hispanic (90%, n = 57; 8%); additionally, 3% (n = 2) were Black/African American, 2% (n = 1) American Indian or Alaska Native, 2% (n = 1) Asian, and 8% unreported.

**Study Procedures**

As part of the larger study, children were administered a battery of language and reading measures between January and May of third grade. All measures except the reading comprehension test were individually administered. The majority of children were assessed in a quiet room in their schools, but 10 children were assessed in alternate
locations (in a library or at home) when it was not possible to conduct the assessment during the school day. Testing took place across 10 sessions that lasted up to one hour each.

In addition to the data collected in the larger study, data on children’s spelling ability were collected specifically for the present study. The spelling assessment was administered in the same session as the reading comprehension assessment because both tests could be group-administered. Group size ranged from 1 to 6 children.

**Assessor training.** Research staff underwent comprehensive training in the administration and scoring of all measures. Training procedures included on-line training modules, quizzes, and mock administrations with supervising assessors. All assessors had to meet criteria of 100% on the quiz and at least 90% fidelity to administration procedures before they were permitted to begin assessments. As an example, see Appendix A for the spelling measures’ fidelity assessment used to ensure that assessors had met training criteria. Scoring for measures used in the present study was conducted in the field, and additional checks for accuracy were conducted in the lab before data were entered into the database.

**Measures**

The measures used in the present study were reading comprehension, vocabulary, and word recognition fluency assessments administered as part of the larger study, in addition to two spelling assessments administered as part of the present study. In the larger study, the constructs of reading comprehension, vocabulary, and word recognition were assessed using multiple measures. The sample size in the present study was not
sufficient for a latent variable approach, however, so a single measure of each construct had to be selected. The choices were based on the results of structural equation models published by the larger study (Language and Reading Research Consortium, 2015). The measure with the highest factor loading on each construct in the third grade sample was selected for use in the present study. Clarification on the selection process for word recognition requires note. The LARRC study tested the influence of accuracy and fluency separately and found that, in third grade, the fluency construct had a stronger relationship to reading comprehension than the accuracy construct. A contextual (i.e., passage-level) fluency measure was included in the fluency factor, and it loaded slightly higher than the word-level fluency measure (.89 versus .85). Given the focus of the present study on word-level skills, the word-level fluency measure was chosen. Furthermore, the LARRC study used two subtests of the word-level fluency measure (word and pseudoword reading), which both loaded at .85. In the present study, a composite score of these subtests was used. Each of the selected measures is described in detail below. Sample-specific internal consistency reliabilities were estimated using Cronbach’s alpha (Table 1).

**Reading comprehension.** The Comprehension subtest of the *Gates-MacGinitie Reading Test – 4th Edition* (GMRT; MacGinitie, MacGinitie, Maria, & Dreyer, 2000) was administered to assess reading comprehension. This measure consists of 11 fiction and nonfiction passages with three to five multiple-choice questions per passage. The questions require understanding information that is explicitly or implicitly provided in the passage. Practice items were completed as a group and then students were instructed to
read the test passages silently and answer the questions by bubbling in the correct responses on the test form. Students were given a maximum of 35 minutes to complete the test. As reported in the test manual, internal consistency reliability for this measure is .92 and criterion validity is .77 to .79. Raw scores – the total number of questions answered correctly – were z-scored for use in analyses.

**Word recognition.** Form A of the *Test of Word Reading Efficiency – 2nd Edition* (TOWRE; Torgesen, Wagner, & Rashotte, 2012) was administered to assess the accuracy and speed (i.e., fluency) of pseudoword and real word reading. The TOWRE consists of two subtests. For the Phonemic Decoding Efficiency subtest, students pronounced as many pseudowords as they could in 45 seconds. For the Sight Word Efficiency subtest, students read aloud as many real words as they could in 45 seconds. Average alternate forms reliability for both subtests is reported to exceed .90, and concurrent validity is .85 for Phonemic Decoding Efficiency and .89 for Sight Word Efficiency. The raw scores from each subtest were z-scored based on the study sample and then the two z-scores were averaged to form a composite. The composite score was used in analyses.

**Vocabulary.** Form A of the *Peabody Picture Vocabulary Test – 4th Edition* (PPVT; Dunn & Dunn, 2007) was administered to assess receptive vocabulary breadth. This test required students to select a picture from a choice of four that represented the meaning of a word spoken by the examiner. Testing continued until a ceiling of 8 or more errors in a set was reached. Split-half reliability is reported as .93 for the spring of Grade 3; alpha reliability is .96. Criterion validity is reported as .67 with the Core Language score of the *Clinical Evaluation of Language Fundamentals – 4th Edition*, and .80 with
the Expressive Vocabulary Test – 2nd Edition. Raw scores were z-scored for use in analyses.

**Spelling.** Several factors were taken into consideration when choosing the spelling measures for the present study. First, research has demonstrated variability in word characteristics (e.g., morphological and orthographic qualities) across different measures of spelling (Calhoon, Greenberg, & Hunter, 2010; Calhoon & Masterson, 2011). Calhoon and colleagues concluded that a single test cannot adequately measure spelling (Calhoon et al., 2010). Thus, it was decided to use two measures of spelling for the present study.

The choice of which two spelling measures to use was based on three criteria. The first requirement was for established measures that have acceptable reliability and validity. The second requirement was for measures that do not employ basal or ceiling rules (e.g., testing is discontinued after a particular number of consecutive incorrect responses – a common procedure in standardized assessments). This would not only make group administration difficult but would result in children spelling different words. The characteristics of the words they spelled could therefore not be considered equivalent. For this reason, researchers have recommended against using measures that employ basal and ceiling rules (McCarthy, Hogan, & Catts, 2012).

The third requirement was for a reasonable number of test items. Many studies of spelling use a small number of test items but comprehensive assessment is recommended in order to cover a wide range of orthographic knowledge in the test items; e.g., Masterson and Apel (2000) recommend 50-100. However, given the length of the
existing assessment protocol for the larger study, it was not feasible within the present study to add such a lengthy, comprehensive spelling assessment to the battery. Instead, two smaller measures of spelling were chosen, with a total of 37 items that would be administered to all children regardless of performance: a shortened version of *Words Their Way-Elementary* spelling inventory (WTW; Bear, Invernizzi, Templeton, & Johnston, 2008) and *Aimsweb Spelling Curriculum-Based Measurement* (CBM; Shinn & Shinn, 2002).

WTW and CBM have similar administration procedures, except that the CBM is a timed measure. These common procedures will be reported first, followed by a detailed description of each measure. Both WTW and CBM use a word dictation format and can be administered to groups or individuals. In the present study, children were provided with pencils and paper and were seated such that they could not view each other’s responses. All words were administered to all children. The assessor dictated each word twice, a few seconds apart. Sentences were provided only for words that have homonyms (of which there were 2 cases, both on the CBM – *pear* and *weigh*). In these cases, the word was spoken, then the sentence, and then the word again. Students spelled the dictated words on test forms that were lined and numbered for each word. Once each test was complete, assessors were instructed to ask students to name aloud any letters that were printed illegibly. See Appendix B for complete administration instructions.

WTW is a spelling inventory consisting of 25 words intended to represent a variety of spelling features that range in difficulty, including words with inflected and derivational morphemes. The elementary version is intended for first through sixth grade.
WTW was modified for the present study by shortening the word list from 25 to 20 for the sake of reducing administration time. Item analysis by Sterbinsky (2007) was used in guiding which words to remove. For the full 25-word test, the reported item difficulty ranged from 15.0 (most difficult) to 98.9 (least difficult), and the index of discrimination ranged from 2.2 (low discrimination) to 65.3 (high discrimination). The words omitted in the present study all had an index of discrimination of less than 32 and included the least difficult word and the most difficult word.

Reliability and validity of the WTW inventories has been reported based on data collected in a school district in California (Sterbinsky, 2007). Internal consistency was examined using data from 862 students in grades 2 to 5 who completed the elementary spelling inventory; Cronbach’s alpha was reported as .915. Test-retest reliability was examined using data from 901 students, including ELL, gifted, and special education students. Estimates based on 2 administrations of the test given a week apart in the spring ranged from .742 to .974. For third grade specifically, the test-retest reliability was reported as .95. Concurrent and predictive validity were examined based on data from 685 students, using the spring administration of the California State Standards – English Language Arts (CST-ELA) tests as the criterion. For third grade, coefficients for concurrent validity with the CST-ELA subtests ranged from .497 to .692, and from .135 to .553 for predictive validity.

The second spelling assessment, the CBM, was chosen as a measure of general spelling ability for grade-level words. The CBM contains graded word lists that are tied to general curriculum. Since the CBM is timed, it offers a measure of spelling fluency.
The test words are randomly sampled from a number of widely used spelling series and therefore are considered “curriculum independent” (Shinn & Shinn, 2002). The use of curriculum-based measures for assessment of spelling ability is common in practice and supported in the literature (e.g., Hosp & Fuchs, 2005; Hosp & Hosp, 2003). For the present study, the winter benchmark spelling list for third grade was administered.

The CBM was administered according to the standardized instructions in the test manual (Shinn & Shinn, 2002). Administration was timed for 2 minutes exactly, with new words dictated every 7 seconds for a total of 17 words. To facilitate the rapid and precise timing of administration, the administration test sheet (Appendix B) provided the exact second at which each new word was to be dictated. Assessors were then able to dictate at the correct pace using a timer. All items were administered to all students.

Reliability and validity of CBM are not reported directly by the publishers. Instead, the publishers cite research supporting the reliability and validity of spelling CBMs in general (NCS Pearson Education, 2012). Test-retest reliability for words spelled correctly and number of correct letter sequences (to be described below) was reported as .91 and .86 respectively. Reliability for 2 parallel forms administered at the same time was .96 and .97. Median criterion validity ranged from .88 to .95 for words spelled correctly and .81 to .98 for correct letter sequences.

**Spelling Scoring Metrics**

Four scoring metrics were utilized to score both the WTW and the CBM. The first is a binary accuracy score (correct/incorrect) and the remaining three allow credit for partially correct spellings. High correlations between WTW and CBM were anticipated
and therefore composite scores were computed for use in analyses. The scores for each measure were converted into z-scores based on the study sample and then the z-scores were averaged to create a composite. This procedure was completed for each of the scoring metrics.

**Words Spelled Correctly (WSC).** The first metric is a simple accuracy score. Each word was scored as correct or incorrect and the total number of words spelled correctly was summed. The maximum possible score is 20 for WTW and 17 for S-CBM. The first author scored all data and a second graduate student double-checked a random 15% of the sample \( (n = 10) \) for accuracy. There were no discrepancies in scoring.

**Correct Letter Sequences (CLS).** The second metric is the number of correct letter sequences, a commonly used metric in curriculum-based measures such as the CBM. A correct letter sequence is a pair of letters correctly spelled and placed together in a word, and includes the space before and after the word. The maximum possible CLS score for a word, therefore, is the number of letters in the word plus one. For example, if the word *cat* is correctly spelled, the correct sequences are \((space)-c, c-a, a-t, \) and \(t-(space)\), for a total of 4 correct letter sequences. Scoring procedures were conducted according to the Aimsweb CBM Training Manual (Shinn & Shinn, 2002). The number of correct letter sequences in each word were counted and averaged. The maximum CLS is 120 for WTW and 112 for CBM and therefore the maximum averaged score is 6 for WTW and 6.59 for CBM. Reliability procedures for scoring of this and subsequent metrics will be discussed below.
Spelling Sensitivity Score for Elements (SSE) and Words (SSW). The final two metrics are based on analysis of the underlying linguistic knowledge demonstrated in children’s spellings. An established system, the Computerized Spelling Sensitivity System (CSSS; Masterson & Apel, 2010, 2013; Masterson & Hrbec, 2011), was used and procedures outlined in the CSSS Manual were followed. Children’s spellings were scored based on the linguistic accuracy of each element in the word. Elements are the graphemes that represent phonemes in a base word as well as any affixes (e.g., the suffix -ful) and juncture changes. Juncture changes are spelling modifications required at the juncture of a base and an affix (e.g., dropping the <e> in hoping, doubling the <p> in slipped, or changing <y> to <i> in carries). Based on the CSSS procedures, each element is scored on a scale of 0 to 3. The maximum score of 3 points is awarded for elements spelled correctly. If an element is incorrect but spelled with a ‘legal’ alternative, it receives a score of 2. ‘Legal’ is defined as orthographically or morphologically plausible. For example, the spelling noat for note is considered legal because the grapheme <oa> is a legal spelling for the long <o> sound. When an illegal spelling is used, the element is scored as 1 (e.g., noot for note). Elements that are omitted receive a score of 0.

The CSSS includes a large dictionary of words that have already been parsed into elements and have legal spellings assigned for each element. The complete list of words from the WTW assessment is already included in the CSSS dictionary, but the S-CBM words are not. The author and another graduate student, also trained in the CSSS, together entered these words into the dictionary, coming to agreement for each word on the parsing into elements and the list of legal spellings for each element. In cases of
uncertainty, the new entries were compared to similar existing entries in the dictionary and agreement was reached between the two coders.

The CSSS program calculates two scores. The Element Score (SSE) is the average score for all elements in the sample, calculated by dividing the number of element points awarded by the total number of elements. The Word Score (SSW) is an overall classification of a word that reflects the lowest single element score in that word; i.e., 3 points if all elements are spelled correctly, 2 points if all misspelled elements are legal, 1 point if any misspelled elements are illegal, and 0 points if any elements are omitted. The SSW score is calculated by dividing the number of word points awarded by the total number of words in the sample. An SSW score < 1.00 suggests that a child is likely making many phonological errors; 1.00-2.00 suggests that a child’s spellings are phonologically accurate but do not follow orthographic conventions; > 2.00 suggests that the child’s spellings are orthographically legal but errors are occurring in the mental graphemic representations.

**Coder training and intercoder reliability.** For each of the partial-credit scoring metrics (CLS, SS-E, SS-W), scoring was conducted by the author (primary coder) and one of two other graduate students. Procedures were followed to ensure that scoring was conducted in a reliable manner. First, each coder underwent training. For CLS, this included reading the Aimsweb S-CBM training manual (Shinn & Shinn, 2002) and completing the five practice tests in the manual. For SSE and SSW, this included reading the CSSS manual (Masterson & Apel, 2015) and watching the online demonstration videos (http://www.missouristate.edu/csd/LLL/174784.htm).
Second, intercoder reliability was measured for each of the partial-credit scoring metrics. For CLS, an estimate of intercoder reliability was obtained using an intraclass correlation coefficient (ICC) for total scores. Acceptable criterion based on the literature is an ICC > .80 (Nunnally & Bernstein, 1994). Data for 10 participants (15%) were double-coded and an ICC of .998 was obtained. For SSW and SSE, multiple steps were taken to ensure reliability. First, children’s spellings were examined for analyzability. Analyzable spellings are those that contain two elements that were either spelled correctly or represented by a legal grapheme. The author and a graduate student rated the spellings of a random 10 participants as analyzable or not. Initially, there was disagreement on two items. The raters discussed the items and came to an agreement. Finally, children’s spellings were imported into the CSSS software and compared to the dictionary. Each spelling was manually checked by the author to ensure that the parsings aligned as expected within the software prior to exporting the scores. Data for a random 10 participants was re-entered by a graduate student to check for reliability of the parsing alignment. An ICC of 1.00 was obtained for SSW and .996 for SSE.

**Analytic Strategy**

The present study aims to answer two major questions related to children’s spelling skills. Preliminary analysis will test the data for the assumptions of multiple regression (i.e., independence of the residuals, homogeneity of variance, linearity, and normality of the residuals), multicollinearity, and outliers and influential cases. Next, statistical analyses will be conducted using SPSS to address the two research questions. These analyses are described next.
Research question 1. To address the first research aim, to describe the bivariate relations between reading comprehension, spelling, word recognition, and vocabulary, descriptive statistics (e.g., means, standard deviations, and ranges) will be provided for each variable of interest. Each spelling scoring metric will be examined separately. Next, correlations and scatterplots will be provided. The correlation between each predictor and reading comprehension, when squared, will indicate the percent of variance in reading comprehension accounted for by that variable alone. Of particular interest to the present study is whether the strength of the correlation between reading comprehension and spelling will vary with spelling scoring metric. Scatterplots will depict the bivariate distributions of the correlations, enabling a visual examination of any possible dissociations, or unexpected high/low scores (e.g., children who score high on reading but low on spelling, or low on reading but high on spelling). To facilitate comparisons, z-scores will be plotted.

Research question 2. To address the second research question, to determine the extent to which spelling is uniquely related to reading comprehension and whether the relation varies by spelling scoring metric, a series of hierarchical regression analyses will be run. These analyses will determine the relative contribution (amount of variance explained) of spelling to reading comprehension in the presence of word recognition fluency and vocabulary. Three separate regressions will be run for each spelling metric. In Model 1, word recognition fluency will be entered at Step 1 and spelling will be added at Step 2. This analysis will provide an estimate of the model improvement as a result of adding spelling as a predictor. Specifically, the $R^2$ change (the squared semipartial
correlation coefficient) will indicate the unique contribution of spelling when accounting for word recognition fluency. In Model 2, vocabulary will be entered at Step 1 and spelling at Step 2. This will estimate the unique contribution of spelling after accounting for vocabulary. In Model 3, both word recognition fluency and vocabulary will be entered at Step 1. Spelling will again be entered at Step 2. This analysis will provide the $R^2$ change to indicate the unique contribution of spelling after accounting for both word recognition fluency and vocabulary. The significance tests for the $R^2$ change and for the predictors themselves will be examined and compared in all models.

All models will be run for each of the four spelling metrics. The $R^2$ change for spelling will be compared to determine whether the strength of the relationship between spelling and reading comprehension varies with the scoring metric used.

There are limitations to using the indices produced by multiple regression for measuring the importance of predictors, particularly with correlated predictors (e.g., Darlington, 1968). Therefore, supplementary analysis will be conducted to more fully explore the importance of spelling relative to word recognition fluency and vocabulary in the prediction of reading comprehension. The Pratt Index will be used for this purpose. An $R^2$-based statistic, the Pratt Index provides a means of partitioning the explained variance in the model to each predictor. It is calculated simply as the product of the standardized regression coefficient and the Pearson correlation between the predictor and the outcome, divided by the $R^2$ of the model (Thomas, Zhu, & Decady, 2007). The resulting Pratt Index for all predictors will be compared to determine the relative importance of each in predicting reading comprehension.
Chapter 4: Results

All analyses were conducted using SPSS version 22. There were no missing data. All variables were standardized, based on the sample statistics, to achieve uniformity in scaling. A composite was created for word recognition by averaging the \( z \)-scores from the sight word and phonemic decoding subtests of the TOWRE. Similarly, composites for each spelling metric were created by averaging the WTW and CBM \( z \)-scores. All analyses were run using these \( z \)-scores and composites.

Preliminary inspection of the data involved screening for violations of the assumptions of multiple linear regression. This screening was conducted for variables in each of four models (for the four spelling metrics; to be described below). Scatterplots of studentized residuals were plotted against unstandardized predicted values and against each independent variable. Based on visual inspection of the plots, the assumptions of independence, homogeneity of variance, and linearity were met. Review of the Shapiro-Wilk tests (\( ps > .05 \)) and skewness and kurtosis values (all < \( |2| \)) suggested normality of the residuals. Q-Q plots and histograms provided further evidence of normality. Casewise diagnostics showed no evidence of influential cases. Data were also examined for multicollinearity. Based on the recommendations of Lomax and Hahs-Vaughn (2012), no evidence of multicollinearity was found. Tolerance values were > .10 in all models and the variance inflation factors (VIF) were < 10. Furthermore, eigenvalues were always >
.10 and condition indices were < 15. Thus, all assumptions of multiple linear regression – independence, homogeneity of variance, linearity, normality, and noncollinearity – have been met.

**Question 1: Descriptives and Bivariate Relations**

Descriptive statistics are presented in Table 1. Standardized scores indicate that children’s performance on reading comprehension, word recognition, and vocabulary was age appropriate. Standard scores for the WTW spelling assessment are not available since WTW is not a normed measure. For CBM, Aimsweb provides national norms in the form of percentile ranks for each grade. Based on the percentiles for winter of grade 3, 51% (n = 32) of the children in the present study scored at or above the 50th percentile. Thus, performance on the CBM was grade appropriate. Comparing raw scores on the two spelling measures revealed that children performed slightly better on WTW than on CBM (e.g., percent words spelled correctly was 63.15% on WTW and 60.76% on CBM). This was true for each metric except SSE for which the two measures were equivalent. The average SSW score of > 2 suggests that children’s spellings were, generally, orthographically legal with errors occurring in mental graphemic representations. Errors in morphology were also common.

Concurrent, bivariate relationships between variables were examined using Pearson correlations (Table 2). All correlations were moderate to strong and significant at p < .001. Reading comprehension was strongly and most highly correlated with CLS (r = .676) and least correlated with vocabulary (r = .564). Overall, the highest correlations were between word recognition and spelling (rs ranging .802 to .841) and the lowest were
between vocabulary and spelling (rs ranging .412 to .456). The spelling metrics were very highly correlated with each other (rs ranging .959 to .985), indicating almost complete overlap between them. The correlations between spelling metrics and reading comprehension ranged from .588 to .676; between spelling metrics and word recognition the range was .802 to .841; and between spelling and vocabulary the range was .412 to .426.

The relation between reading comprehension and the predictors was further explored using scatterplots. Scatterplots reflect the strength and direction of the relationship as indicated by the correlation coefficient but allow for visual inspection of each case. As seen in Figures 1 to 6, there was a high degree of consistency across measures, such that children who scored high or low on any predictor also scored high or low on reading comprehension. Specifically, no children scored high (≥ 1 SD) on word recognition or spelling but low (≤ -1 SD) on reading comprehension. This was true across all spelling metrics. Only one child scored high on vocabulary but low on reading comprehension.

Scatterplots of word recognition and spelling metrics were also examined. As seen in Figures 7-10, no evidence of dissociation was found. Children who were good or poor at word recognition were also good or poor at spelling and vice versa. Specifically, no children scored high (≥ 1 SD) on word recognition but low (≤ -1 SD) on spelling, or high on spelling but low on word recognition. Again, this held true across all spelling metrics.
In an effort to further understand the relation between the three lexical-level skills, a principal components analysis (PCA) was conducted ad hoc. PCA is a variable reduction method that can be used to identify the internal structure of the data. Of interest in the present study is whether spelling may be a stronger indicator of lexical-level literacy than word recognition and vocabulary, given that it is hypothesized to be a stronger indicator of underlying lexical quality. Thus, PCA was conducted with three variables in order to compare their loadings: word recognition, vocabulary, and spelling. A separate analysis was run for each spelling metric. Results are presented in Table 3. As expected, the solution for each analysis was a single component, assumed here to represent lexical-level literacy. All loadings were high. Spelling scored using the WSC, CLS, and SSW metrics had higher loadings than word recognition and vocabulary. The loading for the SSE metric was equivalent to the loading for word recognition. These results indicate that spelling may be the strongest indicator of lexical skill. The results must be interpreted with caution, however, given the small sample size in this study. Definitive guidelines for minimal sample size in PCA do not exist but results drawn from small sample sizes (e.g., $N < 100$) are considered prone to error and lacking in generalizability (Osborne & Costello, 2004). Thus, the present results provide initial insights only and are not intended as conclusive or generalizable.

**Question 2: The Contribution of Spelling to Reading Comprehension**

To first examine the percent of variance in reading comprehension explained by each predictor on its own, the correlation coefficients (Table 2) were squared. CLS
predicted the most variance at 45.7%. Word recognition and WSC each explained 41%. SSE explained 36.6%, SSW explained 34.6%, and vocabulary explained 31.8%.

Next, to examine the contribution of spelling to reading comprehension while controlling for word recognition and vocabulary, a series of hierarchical regressions were run. Four sets of models were tested, one for each spelling scoring metric. Each set had three models, designed to evaluate the unique contribution of spelling in the presence of word recognition (Model A), vocabulary (Model B), and a full model with word recognition and vocabulary together (Model C).

**Words spelled correctly (WSC).** The first set of models tested spelling using the WSC metric. Results are presented in Table 4. In the word recognition model, word recognition was entered at Step 1 and spelling was added at Step 2. This allowed for examination of whether spelling, measured using WSC, contributed unique variance to reading comprehension beyond that explained by word recognition. The total variance in reading comprehension explained by word recognition and spelling together was 42.8% ($R^2_{adj}$). The $R^2$ change for spelling was not significant, $\Delta F (df 1, 60) = 3.993, p = .05$, meaning that spelling did not contribute unique variance after accounting for word recognition. Examining the individual predictors, neither word recognition nor spelling were significant in the model (word recognition $\beta = .348, t = 1.994, p = .05$; spelling $\beta = .349, t = 1.998, p = .05$). Both were closely predictive but the alphas did not meet traditional levels of significance.

Next, the contribution of spelling was tested while controlling for vocabulary. Vocabulary was entered at Step 1 and spelling was added at Step 2. The total variance in
reading comprehension explained by vocabulary and spelling together was 48.7% ($R^2_{adj}$). The $R^2$ change for spelling was significant, $\Delta F (df 1, 60) = 22.305, p < .001$, meaning that spelling explained variance in reading comprehension beyond that explained by vocabulary. Both vocabulary ($\beta = .344, t = 3.364, p = .001$) and spelling ($\beta = .483, t = 4.723, p < .001$) were significant predictors in the model.

The full model for the WSC metric controlled for both word recognition and vocabulary. At Step 1, word recognition and vocabulary together explained 50.1% ($R^2_{adj}$) of the variance in reading comprehension. Both word recognition ($\beta = .490, t = 4.963, p < .001$) and vocabulary ($\beta = .361, t = 3.654, p = .001$) were significant predictors at this step. Adding spelling in Step 2 did not explain additional variance, $\Delta F (df 1, 59) = 1.926, p = .170$, and spelling was not a significant predictor in the model ($\beta = .230, t = 1.388, p = .170$). Word recognition was also not significant ($\beta = .310, t = 1.911, p = .061$), leaving vocabulary as the only predictor to make a significant unique contribution to reading comprehension ($\beta = .330, t = 3.292, p = .002$) when all three lexical-level skills were in the model together.

To conclude, when spelling was scored using the WSC metric it added significant unique variance in predicting reading comprehension in the presence of vocabulary only. Word recognition and WSC were not significant in the presence of each other. Vocabulary was significant in the presence of word recognition and WSC individually and together.

**Correct letter sequences (CLS).** In the second set of models, spelling was measured using the CLS metric. Results are presented in Table 5. In the model with word
recognition at Step 1 and spelling at Step 2, the total variance in reading comprehension explained by both variables together was 45.7% \((R^2_{adj})\). The \(R^2\) change for spelling was significant, \(\Delta F (df 1, 60) = 7.459, p = .008\). Examining the individual predictors, spelling was significant \((\beta = .472, t = 2.731, p = .008)\) but word recognition was not \((\beta = .243, t = 1.404, p = .165)\).

In the vocabulary model, the total variance in reading comprehension explained by vocabulary and spelling together was 52.5 % \((R^2_{adj})\). The \(R^2\) change for spelling was significant, \(\Delta F (df 1, 60) = 28.934, p < .001\). Both vocabulary \((\beta = .323, t = 3.288, p = .002)\) and spelling \((\beta = .529, t = 5.379, p < .001)\) were significant predictors.

In the full model with word recognition and vocabulary at Step 1 and spelling at Step 2, spelling explained additional unique variance in reading comprehension, \(\Delta F (df 1, 59) = 4.737, p = .034\), for a total variance explained of 53.0%. Spelling was a significant predictor in the model \((\beta = .358, t = 2.176, p = .034)\). Vocabulary was also significant \((\beta = .315, t = 3.210, p = .002)\) but word recognition was not \((\beta = .207, t = 1.286, p = .203)\).

In contrast to the WSC metric, then, spelling measured using the CLS metric added significant unique variance in predicting reading comprehension in the presence of both word recognition and vocabulary. Word recognition was not significant in the presence of CLS. Vocabulary was significant in the presence of both word recognition and CLS.

**Spelling sensitivity-words (SSW).** Results for the models using the SSW metric are presented in Table 6. In the word recognition model, the total variance in reading comprehension explained by word recognition and spelling together was 40.9% \((R^2_{adj})\).
The $R^2$ change for spelling was not significant, $\Delta F (df \ 1, 60) = 1.935, p = .169$.

Examining the individual predictors, word recognition was significant ($\beta = .441, t = 2.550, p = .013$) but spelling was not ($\beta = .241, t = 1.391, p = .169$).

In the vocabulary model, the total variance in reading comprehension explained by vocabulary and spelling together was 45.9% ($R^2_{adj}$). The $R^2$ change for spelling was significant, $\Delta F (df \ 1, 60) = 18.144, p < .001$. Both vocabulary ($\beta = .371, t = 3.567, p = .001$) and spelling ($\beta = .443, t = 4.260, p < .001$) were significant predictors in the model.

In the full model with word recognition and vocabulary, spelling did not explain additional unique variance in reading comprehension, $\Delta F (df \ 1, 59) = 0.728, p = .397$, nor was it a significant predictor ($\beta = .138, t = 0.853, p = .397$). Word recognition was significant in the model ($\beta = .382, t = 2.385, p = .020$), as was vocabulary ($\beta = .345, t = 3.426, p = .001$). The total variance explained by the three predictors was 53.0%.

To conclude, when spelling was scored using the SSW metric it added significant unique variance in predicting reading comprehension in the presence of vocabulary but not word recognition. Both word recognition and vocabulary were significant in the presence of SSW and each other.

**Spelling sensitivity-elements (SSE).** Results for the SSE metric (Table 7) are very similar to the SSW results. In the word recognition model, the total variance in reading comprehension explained by word recognition and spelling together was 40.6% ($R^2_{adj}$). The $R^2$ change for spelling was not significant, $\Delta F (df \ 1, 60) = 1.660, p = .203$, and spelling was not a significant predictor in the model ($\beta = .211, t = 1.288, p = .203$). Word recognition was significant ($\beta = .471, t = 2.875, p = .006$).
In the vocabulary model, the total variance in reading comprehension explained by vocabulary and spelling together was 45.3% ($R^2_{adj}$). The $R^2$ change for spelling was significant, $\Delta F (df 1, 60) = 17.304, p < .001$. Both vocabulary ($\beta = .388, t = 3.765, p < .001$) and spelling ($\beta = .429, t = 4.160, p < .001$) were significant predictors in the model.

In the full model with word recognition and vocabulary, spelling did not explain additional unique variance in reading comprehension, $\Delta F (df 1, 59) = 0.788, p = .378$, nor was it a significant predictor ($\beta = .135, t = 0.888, p = .378$). Both word recognition ($\beta = .387, t = 2.541, p = .014$) and vocabulary ($\beta = .348, t = 3.482, p = .001$) were significant predictors. The total variance explained by the three predictors was 49.9%.

To conclude, the results of the SSE metric mirrored those of the SSW metric. Spelling as measured by SSE added significant unique variance in predicting reading comprehension in the presence of vocabulary but not word recognition. Both word recognition and vocabulary were significant in the presence of SSE and each other.

**Relative importance.** To further explore the relative importance of the three lexical-level literacy skills to reading comprehension, the Pratt Index was calculated for each predictor in the full models. The Pratt Index partitions the explained variance in a model to each predictor and is designed to overcome some of the limitations of multiple regression indices for determining relative importance of predictors. As shown in Table 8, results were consistent with those of the regression analyses. Spelling had the highest Pratt Index of all predictors in the model using the CLS metric. For the other metrics, word recognition had the highest Pratt Index and spelling the lowest. Thus, spelling was the most important predictor of reading comprehension in the presence of word
recognition and vocabulary when CLS was the scoring metric. It was the least important predictor when the scoring metric was WSC, SSW, and SSE.
Chapter 5: Discussion

Gaining insight into the skills necessary for skilled reading comprehension is an endeavor that consumes many research agendas and informs practice related to the identification and instruction of children with or at risk for academic difficulties. The importance of word recognition and vocabulary for skilled reading comprehension has been well established in the literature. The role of spelling, however, has been largely ignored, despite the fact that our understanding of spelling has evolved from being considered purely a visual skill to being recognized as a linguistic skill that derives from the same language underpinnings as word recognition. Spelling and word recognition are highly correlated and, along with vocabulary, comprise the set of lexical-level literacy skills. No research to date has examined the combined and unique contributions of these skills to reading comprehension. Given that spelling may be a strong indicator of lexical quality, a construct hypothesized to be a pressure point in reading comprehension, it is imperative to gain a better understanding of the unique role that spelling may play in reading comprehension. The present study is a first step in this direction, describing the literacy skills of a sample of third-grade children and exploring the relation between spelling, amongst other lexical-level literacy skills, and reading comprehension. Descriptive statistics and hierarchical regressions were used to answer the research questions.
Additionally, this study investigated whether the relationship between spelling and reading comprehension varied with spelling scoring metric. To date, little research has been conducted on the differences between scoring metrics, with none exploring the relation between spelling and reading comprehension. If spelling were to provide insight into children’s reading comprehension skills, it would be useful to know which method of scoring is most strongly related. From a practical perspective, some metrics are more time consuming than others to score and may therefore only be desirable if they were able to explain more variance in reading comprehension. This study, therefore, sought to answer the question of whether scoring metric matters in the relation between spelling and reading comprehension.

The present work revealed several major findings. First, children’s literacy skills were characterized by strong concurrent relations and a high degree of consistency across skills. Second, spelling was shown to be a significant, unique predictor of reading comprehension when using the CLS scoring metric. Third, differential relations were revealed between spelling and reading comprehension depending on the spelling scoring metric used. In the following section, these results will be discussed in detail and situated within the context of the extant literature. In addition, future research directions, possible implications, and limitations of this work are discussed.

**Characteristics and Bivariate Relations**

The third-grade children in this study demonstrated age-appropriate abilities on measures of reading comprehension, word recognition, vocabulary, and spelling. Children’s spellings, on average, were characterized by accurate representation of
phonemes and use of legal orthographic patterns, with errors primarily in mental
graphemic representations and morphology. In other words, the children knew many
orthographic conventions but lacked in word-specific orthographic knowledge as well as
morphological knowledge, including the correct spelling of affixes and spelling changes
at the juncture of the base and suffix. That children showed difficulty with mental
graphemic representations and affixing is not unexpected, given other literature
describing children’s pattern of spelling errors. For example, Arndt and Foorman (2010)
examined the spelling errors of a sample of children at the beginning of second grade and
found that the most common errors were morphological, followed by errors in
orthographic patterns and mental graphemic representations. As in the present study,
phonological errors were rare. Bahr, Silliman, Berninger, and Dow (2012) showed that
morphological errors increased across grades 1 to 9 while phonological and orthographic
errors decreased. These results align with what we know about growth in children’s
linguistic skills. For example, Berninger, Abbott, Nagy, and Carlisle (2010) conducted a
longitudinal study children in grades 1 to 6 and found that growth in phonological
awareness occurred mostly in grades 1 to 3, whereas growth in orthographic awareness
and morphological awareness continued through grade 6.

**Relations amongst the lexical-level skills.** Spelling was strongly correlated with
word recognition in the present study (.80 to .84 across spelling metrics) and moderately
correlated with vocabulary (.41 to .43 across spelling metrics). These correlations are
similar to or higher than those reported in other research between spelling and timed
word recognition for children in grade 3. For instance, Abbott et al. (2010) reported the
correlation to be .79, Morris et al. (2012) reported .77, and Ahmed, Wagner, and Lopez (2014) reported .67 and .69. Correlations between vocabulary and spelling have been reported at .42 in kindergarten (Chua, Rickard Liow, & Yeong, 2016), .41 in grade 1 (Kim, Apel, et al., 2013), and .32 in grades 4 to 6 (Cunningham & Stanovich, 1991).

There was strong consistency across skills and little evidence of dissociation between spelling and word recognition skills in the present study. Although other researchers have described children who are good readers but poor spellers, or vice versa (Fayol et al., 2009; Frith, 1980), we did not find examples of this in the present study. We found no children who scored above average in one skill but below average in the other, and only a few children who scored average in one skill but below average in the other.

Two factors may explain why dissociations were not found in the present study. First, the participants in both Fayol et al. (2009) and Frith (1980) were older, age 10 and 12 respectively, than those in the present study. Dissociations may be more likely to occur in older children as gaps may widen over time. Novice readers who use a guessing strategy may not develop precise orthographic representations of words (Frith, 1980; Snowling, Gallagher, & Frith, 2003; Wolter & Apel, 2010). This strategy may be less detrimental to reading because of the contextual support available, but imprecise orthographic representations are necessary for accurate spelling (Perfetti, 1992). Over time, for these children the gap between what they can read and what they can spell would widen. Second, the type of word recognition assessed by Fayol et al. (2009) differed from the present study. Fayol et al. used text-level reading (i.e., reading in context) to classify children as good or poor readers whereas the present study assessed
context-free word recognition. Reading words in context employs semantic processing more so than orthographic processing (Martin-Chang & Levesque, 2015), so dissociation between reading and spelling may be more evident when assessing contextual reading.

Spelling requires greater precision than word recognition and has therefore been posited as a stronger indicator of the quality of underlying lexical representations (Ehri, 2000; Perfetti, 1992). Results of an exploratory PCA provided preliminary support for this notion. Word recognition, vocabulary, and spelling formed a single component, with spelling loading more highly than word recognition for all scoring metrics except SSE and higher than vocabulary regardless of scoring metric. To my knowledge, the structure of lexical-level literacy has not previously been empirically tested, at least not with all three lexical-level skills together; however, researchers have proposed the conceptualization of a lexical skill domain (Protopapas et al., 2013), and there is evidence for word recognition and spelling representing a unitary literacy construct (Mehta, Foorman, Branum-Martin, & Taylor, 2005).

**Relations with reading comprehension.** Associations were also strong between the lexical-level literacy skills and reading comprehension. Specifically, word recognition ($r = .64$) and spelling ($r = .59$ to .68 across metrics) were highly correlated with reading comprehension. For two spelling metrics, WSC and CLS, the correlation of spelling and reading comprehension was greater than or equal to the correlation of word recognition and reading comprehension. This pattern of results has been reported in other research with children in third grade. Ahmed et al. (2014) reported correlations of .60 and .66 for spelling with two measures of reading comprehension and .51 and .62 for word
recognition and the same two measures of reading comprehension. Abbott et al. (2010) reported a correlation of .53 for spelling and .67 for word recognition, and Mehta et al. (2005) reported .73 for spelling and .66 for word recognition.

Vocabulary was moderately correlated with reading comprehension ($r = .56$). This correlation is in the lower range of what has been reported in other research for children in third grade. For example, Tannenbaum et al. (2006) reported correlations of .54 and .63 between vocabulary and two measures of reading comprehension with PPVT, and Quinn et al. (2015) reported correlations of .62 and .65.

**The Contribution of Spelling to Reading Comprehension**

The primary purpose of the present study was to investigate whether spelling was uniquely associated with reading comprehension when controlling for other lexical-level literacy skills. It was hypothesized that spelling, as a strong index of lexical quality, would contribute unique variance to reading comprehension above and beyond word recognition or vocabulary and that the association would vary with scoring metric.

First, it is of note that the lexical-level literacy skills together accounted for a sizeable amount of the variance (50-53% across spelling scoring metrics) in reading comprehension. Regarding the unique contribution of spelling, results supported the hypothesis in that spelling made a unique contribution to reading comprehension when measured using the CLS metric and was closely predictive ($p = .05$) with the WSC metric. With CLS, spelling was significant in the presence of word recognition, vocabulary, and both combined. It explained an additional 6.5% of variance when controlling for word recognition, 22.2% when controlling for vocabulary, and 3.6% when
controlling for both together. With WSC, these values were 3.7%, 18.5%, and 1.5%, respectively. A slightly higher unique contribution for spelling when controlling for word recognition was reported by Desimoni et al. (2012). In that study, also conducted with children in grade 3, spelling contributed an additional 9.2% of the variance in reading comprehension. The study was conducted in Italian, however, a language with an opaque orthography, so results are not directly comparable. Additionally, the word recognition task involved text-level reading, not context-free word reading as in the present study.

Reed et al. (2016) found that spelling accounted for 2% unique variance in reading comprehension in grade 6 when controlling for vocabulary. This number is much lower than the present findings. The difference in grade may explain this discrepancy. Given what we know about the shift in the simple view of reading developmentally, from being more heavily dependent on the print component in early grades to the language component in later grades (Catts et al., 2005; García & Cain, 2014; Language and Reading Research Consortium, 2015), it is not surprising that vocabulary is more important than spelling for reading comprehension in grade 6. Another possible reason for the difference between results is that the vocabulary task used by Reed et al. was a written measure of depth of lexical knowledge. This type of vocabulary knowledge may have a different relation to reading comprehension and spelling than the measure of oral vocabulary breadth used in the present study.

Despite the fact that models using the WSC metric were significant and explained a large percent of variance in reading comprehension, neither word recognition nor spelling were significant in the presence of each other. In other words, when WRF and
WSC were together, neither one was a unique predictor. The \( p \)-values for both were just above the traditional .05 level of significance. Two points are of note here. First, although word recognition and spelling were not significant, that is not to say that the unique variance explained is unimportant. Indeed, according to Cohen’s criteria (Cohen, 1988), \( R^2 \) values of 2-12% are small but can represent practically important effects. Furthermore, the sample size in the present study may have been insufficient to detect small but true effects.

Second, these results are likely a reflection of the high correlation between word recognition and spelling and indicate that it is shared variance, not unique contributions, that is significantly related to reading comprehension. Having highly correlated predictors in a model can result in ‘competition’ between them, resulting in one or both not being significant (Morris et al., 2012). To confirm the amount of shared versus unique variance (similar to Reed et al., 2016), an ad hoc regression was run to test the unique contribution of word recognition when controlling for spelling. Using the WSC metric, word recognition explained an additional 2.8% of variance in reading comprehension. The unique contribution of spelling was 3.7%; therefore the shared variance was 38.1% (total \( R^2 \) of 44.6% minus the unique contributions). Thus, word recognition and spelling share so much common variance in explaining reading comprehension that there is little room left for either predictor to contribute unique variance. In essence, they are competing with each other and neither reaches significance for unique contribution.
Spelling Scoring Metrics

The third main finding of this study relates to how the association between spelling and reading comprehension varied with the spelling scoring metric. In line with results from other research (Clemens et al., 2014; Ritchey et al., 2010), the four spelling metrics used in the present study were very highly correlated with each other. The range in the present study was .96 to 99; in Clemens et al. it was, .94 to .99; and in Ritchey et al. it was .84 to .98. WSC and CLS were similarly correlated with reading comprehension, as were SSW and SSE. The highest correlation with reading comprehension was with the CLS metric and the lowest was with the SSE metric. The strength of the correlations aligned with the results of the hierarchical regressions. CLS was the only spelling metric that uniquely contributed to reading comprehension in the presence of word recognition and vocabulary, and WSC was closely predictive, just above the cutoff for significance ($p = .05$). That CLS was a better predictor than WSC is similar to findings of Deno et al. (1982) in which CLS was better than WSC at discriminating performance on a written expression task for children in grades 3 to 6.

Regarding the relation between spelling and reading comprehension, it was hypothesized that the unique contribution of spelling would be greater when using metrics that give credit for partially correct spellings, particularly for those that assess multiple linguistic domains. This hypothesis was made based on prior research showing that partial-credit metrics explained more variance in word reading for kindergarteners (Clemens et al., 2014), and that multiple-linguistic metrics provide a more specific assessment of children’s underlying skills (Masterson & Apel, 2013; Ritchey et al.,
The results of the present study partially supported the hypothesis that partial-credit metrics would be more highly related to reading comprehension, but did not support a stronger association for the multiple-linguistic metrics. The CLS metric, a partial-credit metric that scores based on orthographic accuracy only, showed the strongest relation with reading comprehension and was the only metric that was significant in the presence of word recognition. SSW and SSE, the multiple-linguistic partial-credit metrics, showed the weakest relation. WSC, the whole-word accuracy metric, performed similarly to CLS in the presence of word recognition but did not meet traditional levels of significance.

Ritchey et al. (2010) argued that partial-credit metrics offer more precise analysis since they go beyond whole-word accuracy to examine smaller units within spellings. CLS measures orthographic accuracy of letter pairs and SSW and SSE, even more nuanced, measure phonological, orthographic, and morphological accuracy of words and individual elements. From this perspective of precision in skills measured, the metrics used in the present study can be ordered from least to most precise as WSC, CLS, SSW, SSE. This rationale was used in forming the hypotheses of the present study.

Alternatively, precision can be considered from the perspective of the overall lexical precision required by the child and reflected in the spelling. Using this conceptualization, the order of the metrics is reversed. The metrics requiring the least amount of lexical precision from the child are SSW and SSE since they give partial credit for even very imprecise spellings. Lower scores represent less linguistic knowledge. The SSW metric reflects accuracy and legality at word level for each linguistic domain and
represents the lowest element-level score in a given word. SSW, by definition, is always less than or equal to the SSE score. The SSE metric scores the linguistic accuracy and legality of every element in the word. A very imprecise spelling, i.e., one that is phonologically correct but orthographically and/or morphologically incorrect, would still be given partial credit. For example, consider the spelling \textit{wegl} for \textit{weigh}. The child represented the vowel phoneme but did not use an orthographically plausible spelling for it. Therefore, the SSE score would be 2 (the average of a score of 3 for the first element, the \textit{<w>}, and a score of 1 for the vowel), and the SSW score would be 1. In contrast, CLS captures a higher degree of precision since it is a measure purely of orthographic accuracy, which, by nature, encompasses phonological and morphological accuracy. WSC, as a metric of whole-word orthographic accuracy, captures the highest degree of precision. To get the whole word correct requires that the child have a complete, accurate orthographic representation that is fully integrated with phonological and morphological representations.

Coming back to the results of the present study, WSC and CLS performed similarly in the hierarchical regressions, with CLS being a significant unique predictor in the presence of both word recognition and vocabulary and WSC being closely predictive in the presence of word recognition. SSW and SSE performed similarly to each other in the regressions and had much larger \(p\)-values than WSC and CLS. These results make sense in the framework of precision just presented. WSC and CLS both measure only orthographic accuracy, the highest level of precision. SSW and SSE, in contrast, measure accuracy across multiple linguistic domains and give credit even for very imprecise
spellings. Thus, it appears that orthography-only metrics are more closely associated with reading comprehension than are multiple-linguistic metrics, at least in third grade, and that this may be because the orthography-only metrics assess the underlying precision or lexical quality evidenced in children’s spellings.

In contrast to the present results, Clemens et al. (2014) found that, in their concurrent model, all partial-credit metrics accounted for more variance in reading than WSC. There are likely several reasons for this discrepancy between results. Importantly, the participants in the Clemens et al. study were in kindergarten. The relation between spelling and reading may differ for novice readers compared to more experienced and skilled readers. And certainly, as Clemens et al. noted, partial-credit metrics can be more sensitive to phonological awareness, a skill known to be especially important in early reading skills (National Early Literacy Panel, 2008). Furthermore, the dependent variable in Clemens et al. was a word reading factor, comprised of word recognition accuracy and fluency, and oral text reading fluency. In the model predicting to grade 1, reading comprehension was added to the factor, but this still contrasts with the present study where reading comprehension was the only outcome measure. When kindergarten word recognition was controlled, however, all metrics performed similarly.

Another possible explanation for CLS being the strongest of the spelling predictors is statistical. The possible range for CLS scores on each item was greater than that for the other metrics. The WSC score was dichotomous, either 0 or 1, and the SSW and SSE scores ranged from 0 to 3. In contrast, the possible range for raw CLS scores for each item in this study was 4 to 10. A larger range generates more variability and may
have contributed to CLS being more strongly associated with reading comprehension than the other metrics.

**Future Research**

The results of the present study shed light on the relation between spelling and reading comprehension in third grade but also point to important directions for future research. Many facets of the relation remain to be investigated. First and foremost is whether spelling is uniquely predictive of reading comprehension over time. While evidence exists to suggest that spelling is predictive (Chua et al., 2016; Etmanskie et al., 2016; Kim, Petscher, et al., 2013; Ritchey et al., 2015), other research has shown mixed results (Abbott et al., 2010; Desimoni et al., 2012). Furthermore, no studies have examined the longitudinal contribution of spelling while controlling for other lexical-level literacy skills. To gain a better understanding of the predictive ability of spelling could inform screening and identification procedures for children at risk for reading comprehension difficulties. Such longitudinal work should include investigation of spelling metrics to determine whether there are differential effects based on how spelling is scored. Expanding the work of the present research to include other grade levels would also be beneficial.

Methodological improvements on the present study offer another avenue for future research. In particular, with a larger sample size, methods such as a latent variable approach and structural equation modeling would provide important advantages. A latent variable approach, using multiple measures of each construct, would decrease measurement error and therefore improve the reliability of results. All individual
measures are inherently flawed to some degree and cannot fully assess a construct. Individual measures of the same construct often produce different results. For example, measures of reading comprehension and word recognition have been shown to vary in important characteristics and this may have an impact on test outcomes (Cutting & Scarborough, 2006; García & Cain, 2014; Keenan, Betjemann, & Olson, 2008). Use of structural equation modeling allows specification and testing of directionality as well as direct and indirect effects. For example, using structural equation modeling, Morris et al. (2012) found that both word recognition and spelling were directly related to oral reading rate which was directly related to silent reading rate. Reed et al. (2016) found that spelling partially moderated the relationship between vocabulary and reading comprehension. Continued investigations such as these would be valuable in expanding our understanding of the relations between lexical-level literacy skills and reading comprehension.

Given the results of the exploratory PCA in the present study showing that the lexical-level literacy skills all loaded onto a single component, it would be of interest to further examine the structure of lexical quality and the contribution of word recognition, vocabulary, and spelling to the construct. Again, a latent variable approach would be needed. If such investigations were to support the findings of the present study, that spelling loaded most highly onto the lexical construct, this, coupled with research showing that spelling instruction transfers to reading (Conrad, 2008; Graham & Hebert, 2011), would have implications for assessment and instruction.
The results of the present study suggest that orthographic-only metrics are more highly related to reading comprehension than multiple-linguistic metrics, but it may be productive to evaluate a scoring metric for morphology. While the CSSS scoring system, from which SSE and SSW are derived, scores elements according to morphological accuracy, this information is collapsed into the SSE and SSW scores. As can be seen in the authors’ description of how the scores should be interpreted, <1 indicating phonological errors, 1-2 indicating orthographic convention (word-general) errors, and >2 indicating orthographic image (word-specific) errors, the morphological information is ‘lost’ in this scheme. Given that morphological knowledge continues to develop throughout the school years (Berninger et al., 2010), and given research showing that morphological knowledge is important for both spelling and reading comprehension (Apel et al., 2012; John R Kirby et al., 2012), it would be worth exploring whether a scoring metric for morphology would show a different association with reading comprehension.

Other avenues for future research include investigating scoring metrics for word recognition and whether these have a differential effect on outcomes, and the development of a spelling screener, to include detailed psychometric work on item characteristics such as word frequency and orthotactic probability. Additionally, it would be beneficial to explore the relations between spelling and reading comprehension in children with language and/or reading difficulties.
Implications

Several implications of the present research are evident. First, the results have theoretical implications regarding the simple view of reading. The simple view of reading posits that the components of reading comprehension are decoding (referred to as word recognition in the present study) and linguistic comprehension (Hoover & Gough, 1990). As Protopapas et al. (2012) pointed out, however, it is unclear how these skills should be measured and indexed. Recently, researchers have begun to use spelling as a ‘proxy’ for decoding (Kim, Wagner, & Lopez, 2012; Reed et al., 2016). I would argue, however, that spelling needn’t be a proxy for decoding; rather, the component is better conceptualized more broadly as a print component, to include both word recognition and spelling. The results of the present study support this conceptualization in that spelling contributed equally or more highly to reading comprehension, at least when spelling was measured with the orthographic-only metrics. The idea of a print-dependent component was also put forth by Protopapas et al. (2012), who investigated the structure of this component from word and pseudoword accuracy and fluency. Spelling was not measured in their study, however.

An alternative conceptualization, consistent with the lexical quality hypothesis, would be to think of decoding as a lexical domain and linguistic comprehension as a supralexical (i.e., discourse-level) domain. The idea of a lexical domain is supported by the exploratory PCA conducted in the present research in which word recognition, vocabulary, and spelling all loaded onto a single component. Whether vocabulary, the lowest-loading variable in the PCA, would load differently when including other
linguistic comprehension measures is not known. Protopapas et al. (2012) tested both print-dependent and print-independent models. Vocabulary fit well in the print-independent model, but an alternative with vocabulary loading on the print-dependent factor was not tested. More work needs to be done of course, but the present research is a step in this direction of reconceptualizing how the components of the simple view of reading are defined and indexed.

The results of the present research also have important practical implications. First, results support the call of a growing number of researchers for spelling to be included in reading assessment and screening batteries (Chua et al., 2016; Clemens et al., 2014; Morris et al., 2012; Robbins et al., 2010). Measured with metrics that credit orthographic precision, spelling explained equivalent or greater variance in reading comprehension than did word recognition, suggesting that spelling can provide unique information on reading comprehension beyond that obtained by traditional reading assessment.

An added benefit of spelling assessment is that it can be group administered, making it a more efficient assessment tool in the classroom. Assessments of word recognition are typically administered individually and therefore require more time. Yet another advantage of assessing spelling is that analysis of errors provides a clear window into children’s underlying linguistic abilities and therefore can inform instruction (Apel & Masterson, 2001).

Second, the present results provide preliminary support for a relative advantage of using orthography-only scoring metrics, especially CLS. Across all the models examined,
the best prediction of reading comprehension was achieved using the CLS metric, with the full model (word recognition, vocabulary, and spelling) and the vocabulary model (vocabulary and spelling) explaining equivalent amounts of variance (53% and 52.5%, respectively). Spelling was a unique predictor in each of these models. Using WSC, the other orthography-only metric, slightly less variance was explained, and spelling contributed more unique variance than word recognition. Interestingly, the vocabulary model with CLS explained more variance than the vocabulary model with word recognition (50.1%) and more than the full model with WSC (50.8%). These differences are slight but suggest that the CLS metric may offer advantages in explaining reading comprehension. Although further investigation is required, this has potential implications for reading comprehension assessment and screening. Spelling appears to provide unique information in the assessment of reading comprehension, and, combined with vocabulary, may provide more accurate prediction of reading comprehension than word recognition and vocabulary.

Third, the present study adds to the growing body of literature on the importance of spelling as a literacy construct and, as such, a target of instruction. Spelling is known to be closely associated with word recognition and its unique association with reading comprehension, supported by the results of the present study, is beginning to be understood. These results combined with research showing that spelling instruction transfers to word recognition (Graham & Santangelo, 2014), support the notion that spelling instruction may help children become ‘lexical experts’, proficient in both word
recognition and spelling. Such expertise and efficiency at the lexical level is likely to support reading comprehension (Perfetti, 2007; Perfetti & Stafura, 2014).

Limitations

The primary limitation of the present study is its small sample size. The sample size not only limits the generalizability and reliability of results but also determined the correlational design of the study. With a small sample size, a latent variable approach could not be employed, nor sophisticated statistical methods such as structural equation modeling. Latent variable approaches are preferred because they allow for more accurate measurement. Use of these methods requires sample sizes of at least 100, however (Osborne & Costello, 2004). The design of the present study was therefore limited to the use of observed variables and linear regression. Regression is strictly correlational in nature so conclusions about causality or directionality cannot be inferred. Another limitation of a small sample size is limited power to detect effects. Some of the nonsignificant results in this study could actually be important, though small, effects that were undetectable with a sample size of 63.

The present study was also limited in its investigation of concurrent relations only. Spelling data were available at only a single time point, thus examination of predictive relations over time was not possible. Additionally, the participants in this study lived in a single Midwest state and were from mid- to high-SES, English-speaking families. This is not a representative sample and results are therefore not generalizable to the population at large.
Conclusion

The purpose of the present study was to investigate the association between spelling and reading comprehension. There is a paucity of research on how spelling relates to reading comprehension and this study was the first to examine the contribution of spelling in the presence of other lexical-level literacy skills, as well as the possible effect of spelling scoring metric. Findings suggest that spelling explained equivalent or greater variance in reading comprehension than did word recognition when it was measured using metrics that credit orthographic precision (WSC and CLS). Furthermore, spelling was a unique predictor of reading comprehension, controlling for both word recognition and vocabulary, when using the CLS metric. It was uniquely predictive of reading comprehension when controlling for vocabulary only, regardless of scoring metric.

Spelling scoring metrics were differentially related to reading comprehension in the sense that the orthography-only metrics performed similarly to each other, as did the multiple-linguistic metrics. The orthography-only metrics were more strongly associated with, and explained more unique variance in, reading comprehension. Thus, it may not be about precision in skill across phonology, orthography, and morphology that matters most, but precision in orthography, specifically, that is important in the relation between spelling and reading comprehension. CLS may be the metric most strongly associated with reading comprehension because it captures orthographic accuracy (i.e., lexical precision) while also giving credit for partially correct spellings. Multiple-linguistic metrics give credit for less precise spellings and resulted in a weaker relationship.
between spelling and reading comprehension. While this may be the case in the present sample of third grade children, it may not hold true across all grade levels, in children with lesser and greater literacy experience and skill. Future research is needed to further explore the relation between spelling and reading comprehension and to explore the possibility of assessing spelling as a means of identifying children who may have or be at risk for reading comprehension difficulties.


doi:dx.doi.org/10.1016/j.jsp.2013.12.005


doi:10.1037/a0012544


spelling. *Journal of Educational Psychology, 83*(2), 264-274. doi:10.1037/0022-0663.83.2.264


Learning and Instruction, 22(5), 340-353.

doi:dx.doi.org/10.1016/j.learninstruc.2012.02.002


García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of


Retelsdorf, J., & Köller, O. (2014). Reciprocal effects between reading comprehension and spelling. *Learning and Individual Differences, 30*(0), 77-83. doi:dx.doi.org/10.1016/j.lindif.2013.11.007


Appendix A: Tables

Table 1

Descriptives

<table>
<thead>
<tr>
<th>Measures</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Max</th>
<th>Max Possible</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standardized scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NVIQ</td>
<td>103.90</td>
<td>19.23</td>
<td>65</td>
<td>143</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>5.27</td>
<td>2.05</td>
<td>2</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>99.98</td>
<td>14.07</td>
<td>66</td>
<td>127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>108.81</td>
<td>12.83</td>
<td>82</td>
<td>139</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Raw and average scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>31.95</td>
<td>9.85</td>
<td>14</td>
<td>46</td>
<td>48</td>
<td>.920</td>
</tr>
<tr>
<td>SWE</td>
<td>64.29</td>
<td>9.68</td>
<td>37</td>
<td>84</td>
<td>108</td>
<td>*</td>
</tr>
<tr>
<td>PDE</td>
<td>32.16</td>
<td>10.68</td>
<td>10</td>
<td>56</td>
<td>66</td>
<td>*</td>
</tr>
<tr>
<td>VOC</td>
<td>150</td>
<td>14.50</td>
<td>114</td>
<td>185</td>
<td>228</td>
<td>.935</td>
</tr>
<tr>
<td>WSC-WTW</td>
<td>12.63</td>
<td>4.05</td>
<td>2</td>
<td>19</td>
<td>20</td>
<td>.866</td>
</tr>
<tr>
<td>WSC-CBM</td>
<td>10.33</td>
<td>3.81</td>
<td>1</td>
<td>17</td>
<td>17</td>
<td>.834</td>
</tr>
<tr>
<td>CLS-WTW (average)</td>
<td>5.85</td>
<td>0.86</td>
<td>3.75</td>
<td>7.10</td>
<td>7.20</td>
<td>.895</td>
</tr>
<tr>
<td>CLS-CBM (average)</td>
<td>5.41</td>
<td>0.76</td>
<td>3.71</td>
<td>6.59</td>
<td>6.59</td>
<td>.818</td>
</tr>
<tr>
<td>SSW-WTW (average)</td>
<td>2.25</td>
<td>0.44</td>
<td>1.15</td>
<td>2.95</td>
<td>3</td>
<td>.871</td>
</tr>
<tr>
<td>SSW-CBM (average)</td>
<td>2.19</td>
<td>0.48</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>.807</td>
</tr>
<tr>
<td>SSE-WTW (average)</td>
<td>2.76</td>
<td>0.17</td>
<td>2.30</td>
<td>2.99</td>
<td>3</td>
<td>.884</td>
</tr>
<tr>
<td>SSE-CBM (average)</td>
<td>2.75</td>
<td>0.17</td>
<td>2.33</td>
<td>3</td>
<td>3</td>
<td>.802</td>
</tr>
</tbody>
</table>

Note. NVIQ = Kaufman Brief Intelligence Test, matrices subtest; RC = Gates-MacGinitie Reading Test, comprehension subtest; WRF = the Test of Word Reading Efficiency, total; VOC = Peabody Picture Vocabulary Test; SWE = TOWRE sight word efficiency subtest; PDE = TOWRE phonemic decoding subtest; WSC = words spelled correctly; CLS = correct letter sequences; SSW = spelling sensitivity-words; SSE = spelling sensitivity-elements; WTW = Words Their Way; CBM = Spelling Curriculum-Based Measure.

* Reliability not calculated because this is a timed measure.
Table 2

**Correlations between the Predictors and Reading Comprehension**

<table>
<thead>
<tr>
<th>Measures</th>
<th>RC</th>
<th>WRF</th>
<th>VOC</th>
<th>WSC</th>
<th>CLS</th>
<th>SSW</th>
<th>SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading comprehension (RC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word recognition (WRF)</td>
<td>0.640</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary (VOC)</td>
<td>0.564</td>
<td>0.416</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Words spelled correctly (WSC)</td>
<td>0.640</td>
<td>0.835</td>
<td>0.456</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct letter sequences (CLS)</td>
<td>0.676</td>
<td>0.841</td>
<td>0.456</td>
<td>0.976</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spelling sensitivity - words (SSW)</td>
<td>0.605</td>
<td>0.825</td>
<td>0.438</td>
<td>0.985</td>
<td>0.961</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spelling sensitivity - elements (SSE)</td>
<td>0.588</td>
<td>0.802</td>
<td>0.412</td>
<td>0.959</td>
<td>0.963</td>
<td>0.980</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* All correlations are significant at the .01 level (2-tailed).
Table 3

*Loadings for Principal Components Analysis of Lexical Literacy Skill*

<table>
<thead>
<tr>
<th>Indicator</th>
<th>WSC Metric</th>
<th>CLS Metric</th>
<th>SSW Metric</th>
<th>SSE Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td>0.912</td>
<td>0.913</td>
<td>0.912</td>
<td>0.909</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.690</td>
<td>0.688</td>
<td>0.684</td>
<td>0.678</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.925</td>
<td>0.927</td>
<td>0.920</td>
<td>0.908</td>
</tr>
</tbody>
</table>

*Note.* WSC = words spelled correctly; CLS = correct letter sequences; SSW = spelling sensitivity-words; SSE = spelling sensitivity-elements.
Table 4

Hierarchical Regression Results for the Words Spelled Correctly Metric

<table>
<thead>
<tr>
<th>Models/Steps</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>$R^2_{adj}$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.400**</td>
<td>0.409**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.640</td>
<td>0.104</td>
<td>6.503</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.428</td>
<td>0.446</td>
<td>0.037</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.348</td>
<td>0.184</td>
<td>1.994</td>
<td>0.051</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSC</td>
<td>0.349</td>
<td>0.179</td>
<td>1.998</td>
<td>0.050</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.307**</td>
<td>0.319**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.564</td>
<td>0.106</td>
<td>5.340</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.487**</td>
<td>0.503**</td>
<td>0.185**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.344</td>
<td>0.102</td>
<td>3.364</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSC</td>
<td>0.483</td>
<td>0.105</td>
<td>4.723</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.501**</td>
<td>0.517**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.490</td>
<td>0.104</td>
<td>4.963</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.361</td>
<td>0.099</td>
<td>3.654</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.508</td>
<td>0.532</td>
<td>0.015</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.310</td>
<td>0.171</td>
<td>1.911</td>
<td>0.061</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.330</td>
<td>0.100</td>
<td>3.292</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WSC</td>
<td>0.230</td>
<td>0.170</td>
<td>1.388</td>
<td>0.170</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. WRF = word recognition; WSC = words spelled correctly; VOC = vocabulary. ** $p < .01$. 

112
Table 5

*Hierarchical Regression Results for the Correct Letter Sequences Metric*

<table>
<thead>
<tr>
<th>Models/Steps</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>$R^2_{adj}$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word Recognition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.400**</td>
<td>0.409**</td>
<td></td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.640</td>
<td>0.104</td>
<td>6.503</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.457**</td>
<td>0.475**</td>
<td>.065**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.243</td>
<td>0.182</td>
<td>1.404</td>
<td>0.165</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLS</td>
<td>0.472</td>
<td>0.178</td>
<td>2.731</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.307**</td>
<td>0.319**</td>
<td></td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.564</td>
<td>0.106</td>
<td>5.340</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.525**</td>
<td>0.540**</td>
<td>0.222**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.323</td>
<td>0.098</td>
<td>3.288</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLS</td>
<td>0.529</td>
<td>0.101</td>
<td>5.379</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>0.501**</td>
<td>0.517**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.490</td>
<td>0.104</td>
<td>4.963</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.361</td>
<td>0.099</td>
<td>3.654</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>0.530*</td>
<td>0.553*</td>
<td>0.036*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.207</td>
<td>0.170</td>
<td>1.286</td>
<td>0.203</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.315</td>
<td>0.098</td>
<td>3.210</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLS</td>
<td>0.358</td>
<td>0.169</td>
<td>2.176</td>
<td>0.034</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* WRF = word recognition; CLS = correct letter sequences; VOC = vocabulary. * p < .05, ** p < .01.
Table 6

*Hierarchical Regression Results for the Spelling Sensitivity-Words Metric*

<table>
<thead>
<tr>
<th>Models/Steps</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>$R^2_{adj}$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word Recognition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.400**</td>
<td>0.409**</td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.640</td>
<td>0.104</td>
<td>6.503</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.409</td>
<td>0.428</td>
<td>0.018</td>
</tr>
<tr>
<td>WRF</td>
<td>0.441</td>
<td>0.183</td>
<td>2.550</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSW</td>
<td>0.241</td>
<td>0.178</td>
<td>1.391</td>
<td>0.169</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.307**</td>
<td>0.319**</td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.564</td>
<td>0.106</td>
<td>5.340</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.459**</td>
<td>0.477**</td>
<td>0.158**</td>
</tr>
<tr>
<td>VOC</td>
<td>0.371</td>
<td>0.104</td>
<td>3.567</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSW</td>
<td>0.443</td>
<td>0.107</td>
<td>4.260</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.501**</td>
<td>0.517**</td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.490</td>
<td>0.104</td>
<td>4.963</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.361</td>
<td>0.099</td>
<td>3.654</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.499</td>
<td>0.523</td>
<td>0.006</td>
</tr>
<tr>
<td>WRF</td>
<td>0.382</td>
<td>0.169</td>
<td>2.385</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.345</td>
<td>0.101</td>
<td>3.426</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSW</td>
<td>0.138</td>
<td>0.167</td>
<td>0.853</td>
<td>0.397</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* WRF = word recognition; SSW = spelling sensitivity-words; VOC = vocabulary.  
**p < .01.
Table 7

**Hierarchical Regression Results for the Spelling Sensitivity-Elements Metric**

<table>
<thead>
<tr>
<th>Models/Steps</th>
<th>Beta</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>$R^2_{adj}$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word Recognition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.640</td>
<td>0.104</td>
<td>6.503</td>
<td>&lt; .001</td>
<td>0.400**</td>
<td>0.409**</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.471</td>
<td>0.173</td>
<td>2.875</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSE</td>
<td>0.211</td>
<td>0.170</td>
<td>1.288</td>
<td>0.203</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.564</td>
<td>0.106</td>
<td>5.340</td>
<td>&lt; .001</td>
<td>0.307**</td>
<td>0.319**</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.388</td>
<td>0.103</td>
<td>3.765</td>
<td>&lt; .001</td>
<td>0.453**</td>
<td>0.471**</td>
<td>0.153**</td>
</tr>
<tr>
<td>SSE</td>
<td>0.429</td>
<td>0.107</td>
<td>4.160</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Word Recognition and Vocabulary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.490</td>
<td>0.104</td>
<td>4.963</td>
<td>&lt; .001</td>
<td>0.501**</td>
<td>0.517**</td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>0.361</td>
<td>0.099</td>
<td>3.654</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WRF</td>
<td>0.387</td>
<td>0.161</td>
<td>2.541</td>
<td>0.014</td>
<td>0.499</td>
<td>0.523</td>
<td>0.006</td>
</tr>
<tr>
<td>VOC</td>
<td>0.348</td>
<td>0.100</td>
<td>3.482</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSE</td>
<td>0.135</td>
<td>0.158</td>
<td>0.888</td>
<td>0.378</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* WRF = word recognition; SSE = spelling sensitivity-elements; VOC = vocabulary.

** p < .01.
### Table 8

**Pratt Index for Models with each Spelling Metric**

<table>
<thead>
<tr>
<th>Predictor</th>
<th>WSC Metric</th>
<th>CLS Metric</th>
<th>SSW Metric</th>
<th>SSE Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word recognition</td>
<td>0.373</td>
<td>0.240</td>
<td>0.468</td>
<td>0.474</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.350</td>
<td>0.321</td>
<td>0.372</td>
<td>0.375</td>
</tr>
<tr>
<td>Spelling</td>
<td>0.277</td>
<td>0.438</td>
<td>0.160</td>
<td>0.151</td>
</tr>
</tbody>
</table>

*Note.* WSC = words spelled correctly; CLS = correct letter sequences; SSW = spelling sensitivity-words; SSE = spelling sensitivity-elements.
Figure 1. Scatterplot of word recognition with reading comprehension.
Figure 2. Scatterplot of vocabulary with reading comprehension.
Figure 3. Scatterplot of words spelled correctly with reading comprehension.
Figure 4. Scatterplot of correct letter sequences with reading comprehension.
Figure 5. Scatterplot of spelling sensitivity—words with reading comprehension.
Figure 6. Scatterplot of spelling sensitivity-elements with reading comprehension.
Figure 7. Scatterplot of words spelled correctly with word recognition.
Figure 8. Scatterplot of correct letter sequences with word recognition.
Figure 9. Scatterplot of spelling sensitivity-words with word recognition.
Figure 10. Scatterplot of spelling sensitivity-elements with word recognition.
Appendix C: Assessor Training Materials - Spelling Assessment Fidelity Checklist

Fidelity checklist for Words Their Way

☐ Has list of spelling words and instruction sheet

☐ Reads directions accurately

☐ Dictates words at the students’ pace, but spending no more than approximately 60 seconds on any word.

☐ Dictates words at appropriate volume

☐ Speaks naturally, without emphasizing any sounds or syllables in the words

☐ Says all words twice (and only twice), with a pause of 2-3 seconds in between

☐ Occasionally uses “place” cues (e.g., “You should be on the fifth word, which is...”)

☐ Does not provide any assistance with the spelling of any word

☐ Monitors students for getting lost, and points to correct place if child has lost their place

☐ If group administered, ensures that students do not copy from each other

☐ Keeps talking to a minimum and does not answer questions about the test words

☐ Ensures that students stop writing, as directed at the end of the test

☐ Asks for clarification for any illegible spellings, but does not repeat the words
Fidelity checklist for Curriculum-Based Measure

☐ Has list of spelling words and instruction sheet

☐ Reads directions accurately

☐ Starts stopwatch at first word

☐ Dictates words at correct pace (every 7 seconds)

☐ Dictates words at appropriate volume

☐ Speaks naturally, without emphasizing any sounds or syllables in the words

☐ Says all words twice (and only twice), with a pause of 2-3 seconds in between

☐ Uses homonyms in a sentence (as written on admin sheet); says word-sentence-word

☐ Occasionally uses “place” cues (e.g., “You should be on the fifth word, which is...”)

☐ Does not provide any assistance with the spelling of any word

☐ Monitors students for getting lost, and points to correct place if child has lost their place

☐ If group administered, ensures that students do not copy from each other

☐ Keeps talking to a minimum and does not answer questions about the test words

☐ Accurate 2-minute timing

☐ Ensures that students stop writing immediately, as directed at the end of the test

☐ Asks for clarification for any illegible spellings, but does not repeat the words
Appendix D: Spelling Administration and Instructions

Words Their Way

Say to the students:
“*We are going to do 2 spelling activities. The second one will be quicker than the first. You may spell the words by printing them or by using cursive/handwriting – whatever you are best at. I need to be able to read all the letters so make sure you write them carefully. If you make any mistakes, I don’t want you to erase the letters. Instead, just cross them out with one line and then write it over again*”.

Hand out the test sheets for Words Their Way, lined and numbered 1 to 20.

Say to the students:
“*I am going to say some words that I want you to spell on the sheet of paper in front of you. Write the first word on the first line, the second word on the second line, and so on. Some of the words may be easy to spell; some may be difficult. If you are not sure how to spell a word, spell it the best you can. I will count every letter you get correct, even if you don’t know the whole word. I will say every word 2 times. Please keep your eyes on your own paper - do not look at your neighbor’s paper. Are there any questions? (pause) Let’s begin.*”

Say each word aloud and repeat it (each word is spoken two times, with a pause of 2-3 seconds between).

Do not spend more than approximately 60 seconds on a word. If children are still writing at about the 45-second mark on any word, say “*Let’s finish up this word now, I’m ready to move onto the next one*”. Then at about the 60-second mark, move on to the next word even if they are not finished. Say, “*OK, here’s the next word. Move to the next line and spell…*(say the next word)”.

Use place cues as indicated (every 5th word).

At the end of the test, say, “*Stop. Please put your pencils down*”.

Collect the test sheets and check to see if the spellings are legible. For any word that cannot be deciphered, say to the student: “*I can’t read what letter this is/what these letters are? Can you tell me?*” Then write the letter(s) yourself on the line next to the
child’s spelling. If the child asks you what the word was, say, “I can’t say the word again. Just try your best to tell me what the letters are”.

Curriculum Based Measure

In order to accurately time the CBM test, you should use a stopwatch and follow the timing as written on the admin sheet, saying a new word every 7 seconds.

Hand out the lined test paper, numbered 1 to 17.

Say to the students:

“You’re going to do another spelling activity. This one will be a 2-minute spelling activity. Again, I will say some words that I want you to spell on the sheet of paper in front of you. Write the first word on the first line, the second word on the second line, just like you did in the last activity. I’ll give you only 7 seconds to spell each word, which is not a lot so you will have to work quickly.

When I say the next word, write it down, even if you haven’t finished the last one. I will count every correct letter that you write, so try all the words even if you aren’t sure how they are spelled. Remember, don’t waste time erasing words or letters, just cross them out with one line. Please keep your eyes on your own paper - do not look at your neighbor’s paper. You will need to work as quickly as you can on this activity. Are there any questions? (Pause) Let’s begin.”

Say the first word and start your stopwatch. The time at which you should say each new word is printed on the administration sheet.

Say each word twice, pausing 2-3 seconds between.
Even the words that have sentences still only have 7 seconds per word.
You say the word, the sentence, and then the word again. Do not make up sentences for the words that don’t have them.

Say a new word every 7 seconds. This is quite fast paced. Use place cues as indicated (every 5th word).

If students get very frustrated or upset about the pace (to the point that they will not continue the test), you may pause the timer and get them back on track. Tell them it’s OK if they can’t finish a word, they should just try their best, and that you will count all letters that they get correct. You may stop the timer to give this kind of reassurance only once during the test. If you have to alter the timing of the test in this way, be sure to make a note on the test sheet.
After 2 minutes (17 words) say, “Stop. Please put your pencils down.”

Collect the test sheets from students. Make sure the spelling is legible. Ask the student to read the letters in any word that cannot be deciphered. Then write the letters yourself on the line next to the child’s spelling. If a student asks you what the word was, say, “I can’t say the word again. Just try your best to tell me what the letters are”.

If something happens to disrupt the timing of the test, you may stop timing until the problem is solved. Then start wherever you left off, and resume timing. Make a note about what happened on the test sheet.