

Old Dominion University

ODU Digital Commons

Communication Disorders & Special Education
Faculty Publications

Communication Disorders & Special Education

2019

Lexical-Level Predictors of Reading Comprehension in Third Grade: Is Spelling a Unique Contributor?

Kimberly A. Murphy

Old Dominion University, kamurphy@odu.edu

Laura M. Justice

Follow this and additional works at: https://digitalcommons.odu.edu/cdse_pubs



Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Elementary Education Commons](#), and the [Language and Literacy Education Commons](#)

Original Publication Citation

Murphy, K. A., & Justice, L. M. (2019). Lexical-level predictors of reading comprehension in third grade: Is spelling a unique contributor? *American Journal of Speech-Language Pathology*. doi:10.1044/2019_AJSLP-18-0299

This Article is brought to you for free and open access by the Communication Disorders & Special Education at ODU Digital Commons. It has been accepted for inclusion in Communication Disorders & Special Education Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Lexical-level Predictors of Reading Comprehension in Third Grade: Is Spelling a Unique Contributor?

Kimberly A. Murphy

Department of Communication Disorders and Special Education

Old Dominion University

Norfolk, VA

Laura M. Justice

Crane Center for Early Childhood Research and Policy

The Ohio State University

Columbus, OH

Conflict of Interest: The first author received doctoral funding from the Language and Reading Research Consortium. The second author was a principal investigator of the Language and Reading Research Consortium.

Funding statement: This work was supported by grant # R305F100002 of the Institute of Education Sciences' Reading for Understanding Initiative, awarded to the Language and Reading Research Consortium.

Correspondence concerning this article should be addressed to Kimberly A. Murphy, Department of Communication Disorders and Special Education, Old Dominion University, 4501 Hampton Blvd., Norfolk, VA 23529; kamurphy@odu.edu

Abstract

Purpose: Considerable research effort has focused on understanding reading comprehension and reading comprehension difficulties. The purpose of the present correlational study was to add to the small but growing body of literature on the role that spelling may play in reading comprehension, by investigating the full range of lexical-level literacy skills and whether spelling makes a unique contribution. The study also explored whether these relations vary with spelling scoring metric. **Method:** Data were collected from 63 children attending grade 3 in a Midwestern state. 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 for elements (SSE) and words (SSW). **Results:** All spelling metrics were significantly correlated with reading comprehension. 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 not only on orthographic accuracy, but also phonological and morphological accuracy (SSE, SSW). **Conclusions:** These results indicate that spelling is related to reading comprehension and have theoretical and clinical implications for the use of spelling assessment.

Introduction

To read for understanding is the ultimate goal of reading, yet, in the United States in 2015, 63% of fourth graders and 64% of eighth graders performed at basic or below basic levels in reading comprehension (National Center for Education Statistics, 2017). 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. This research has important implications for speech-language pathologists (SLPs) who work with children with language disorders and therefore may be at risk for reading comprehension difficulties.

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 vocabulary, word recognition, and spelling (Bowers, Kirby, & Deacon, 2010). The contribution of vocabulary and word recognition to reading comprehension is well-supported 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 when controlling for word recognition and vocabulary. In doing so, we begin to explore the possibility that spelling, like vocabulary and word recognition, could be used as a means of identifying children who may struggle with reading comprehension.

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). Researchers have argued for the importance of spelling to literacy development

(Foorman & Petscher, 2010; Pinto, Bigozzi, Tarchi, Accorti Gamannossi, & Canneti, 2015), but few studies have directly examined the relationship between spelling and reading comprehension. Like word recognition, spelling depends upon the integration of multiple sources of linguistic knowledge at the sublexical level, including phonological, orthographic, and morphological knowledge (Kim, Apel, & Al Otaiba, 2013), and therefore provides insight into developing competencies in children's linguistic systems. Spelling and word recognition are both skills that involve applying sublexical linguistic knowledge to words. Spelling can be thought of as the 'flip side of the coin' (Ehri, 2000) from word recognition. Whereas reading of words is a recognition task that involves the decoding or translation of print to speech, spelling is a production task that involves the encoding of speech to print. These two skills have much in common and are highly correlated (Abbott, Berninger, & Fayol, 2010), but are not perfectly overlapping.

Spelling and Reading Comprehension: Theoretical Rationale

An association between spelling and reading comprehension is supported both theoretically and empirically. The main hypothesis of the present study is that spelling will be significantly related to reading comprehension in our sample of third grade students, and may explain unique variance when controlling for other lexical-level skills. 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.

Lexical quality. The theoretical rationale for the present study is grounded in the lexical quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2002). This 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. The richer the quality of the representations in memory 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.

Although spelling and word recognition both use the same lexical representations, the quality of those representations may be better indexed by spelling than by word recognition (Perfetti, 1992). Accurate spelling of words requires precise, highly specified orthographic representations that are closely connected with phonological and semantic representations (Ehri, 2000; Perfetti, 1997). In contrast, accurate recognition of words can happen even in the absence of high quality phonologic and/or orthographic representations when context can be used to guess a word (Frith, 1980). Spelling therefore may be a better indicator of underlying lexical quality than word recognition and thus, potentially, a strong predictor of reading comprehension. Based on this rationale, we hypothesize that spelling will relate to reading comprehension and may contribute unique variance beyond the contribution of word recognition.

Orthographic learning. Another theoretical perspective, that of orthographic learning and the self-teaching hypothesis, further 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). Orthographic knowledge consists of two levels: lexical and sublexical. The lexical level consists of stored mental representations of known words and morphemes; the sublexical level consists of knowledge of rules and patterns that govern how and where letters are used in spellings (Apel, Henbest, & Masterson, 2019).

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 (Share, 1995). Newly learned 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. Research suggests that spelling may be superior to word recognition as a self-teaching mechanism for orthographic learning (Conrad, Kennedy, Saoud, Scallion, & Hanusiak, 2019; Ouellette, 2010; Shahr-Yames & Share, 2008). Shahr-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. These results were replicated in a similar study with English-speaking children in second grade (Ouellette, 2010) in which the author 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. Indeed, other researchers have interpreted findings of a relation between spelling and reading comprehension in the context of these hypotheses (Desimoni, Scalisi, & Orsolini, 2012; Retelsdorf & Köller, 2014).

Spelling and Reading Comprehension: Empirical Support

A substantial body of literature supports the role of the lexical-level skills of word recognition (e.g., García & Cain, 2014) and vocabulary (e.g., Quinn, Wagner, Petscher, & Lopez, 2015) for reading comprehension. Much less is known about spelling. Moderate to strong correlations with reading comprehension have been reported (Abbott et al., 2010; Kim, Petscher,

et al., 2013; Ritchey, Silverman, Schatschneider, & Speece, 2015), but researchers have only recently begun to directly examine the role that spelling may play in reading comprehension. In a large study of students in grades 3 to 12, Foorman and Petscher (2010) found that classrooms that had low levels of spelling also had low levels of reading comprehension. Retelsdorf and Köller (2014) found reciprocal effects between spelling and reading comprehension in German students from grade 5 to grade 7. Other studies have shown that poor reading comprehension can be predicted by earlier spelling ability (Desimoni et al., 2012; Ritchey et al., 2015) and that children classified as poor comprehenders performed below typical readers on measures of spelling (Etmanskie, Partanen, & Siegel, 2016).

To fully understand the relation between spelling and reading comprehension, we need to know the unique contribution of spelling while controlling for other skills known to relate to reading comprehension. Some research has taken this approach. For example, there is evidence that spelling predicts a small amount of unique variance in reading comprehension when controlling for earlier reading comprehension ability (Foorman, Petscher, and Bishop (2012), silent text reading fluency (Foorman & Petscher, 2010), and both of these variables together (Kim, Petscher, et al., 2013).

With regards to lexical-level literacy skills, there is some evidence for a unique contribution of spelling while controlling for vocabulary or for word recognition, but no studies have controlled for both. Reed, Petscher, and Foorman (2016) controlled for vocabulary and found a unique contribution of spelling to reading comprehension across grades 6 to 10, with the variance explained by spelling increasing from 2% in grade 6 to 9% in grade 10. Vocabulary made a greater unique contribution than spelling but the contribution decreased across grades, from 19% to 11%. A unique contribution of spelling when controlling for word reading accuracy

and speed was reported by Desimoni et al. (2012) in a study of Italian children. Spelling accounted for 3.5% unique variance in reading comprehension in grade 1 and 9.2% in grade 3. Thus, the importance of spelling may increase across grades when controlling for vocabulary or word recognition.

In sum, there is a growing body of literature on the relation of spelling to reading comprehension but few studies have investigated the unique role that spelling may play. Of those that have taken a multivariate approach, many have included only a single additional predictor (e.g., word recognition, vocabulary, or text fluency). Furthermore, some of the studies were conducted in languages other than English, so generalizations cannot be made due to variation in orthographic complexity across languages. Currently, there is no research on the unique contribution that spelling may make in reading comprehension while controlling for both of the other lexical-level literacy skills, namely word recognition and vocabulary. The present study expands the current literature and fills a gap by exploring the relations between all three lexical-level literacy skills and reading comprehension, with a focus on understanding the role of spelling in particular. In addition, the present study investigated the possible implications of using different scoring metrics for spelling.

Spelling Scoring Metrics

Spelling can be measured using a variety of scoring metrics which 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 or not 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 has not yet been investigated with respect to spelling and reading comprehension.

The most commonly used scoring metric is the number of words spelled correctly, a quick and simple method of scoring that provides an index of children's whole word 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), 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 (i.e., orthography) only, so linguistic plausibility of a spelling is not considered. The Spelling Sensitivity System is based on a multiple-linguistic approach to error analysis. Spelling errors are analyzed according to whether they are related to phonemic awareness, orthographic pattern knowledge, morphological awareness, or storage of word-specific mental representations. A computerized version of the Spelling Sensitivity System, available for free online, automates the scoring process and therefore reduces the time required for scoring.

There is some research systematically comparing spelling scoring metrics but no studies have done so in exploring the relation between spelling and reading comprehension. Treiman and colleagues examined the relations between spelling scoring metrics and future spelling ability (Treiman, Kessler, & Caravolas, 2019; Treiman, Kessler, Pollo, Byrne, & Olson, 2016). Even as early as the beginning of kindergarten, metrics based on accuracy of letters, such as whole-word accuracy and correct letter sequences, were better predictors of second grade spelling than phoneme-based metrics, such as phonological plausibility and phoneme distance.

Other research has examined the relations between spelling scoring metrics and word reading ability. Ritchey, Coker, and McCraw (2010) found equivalent correlations between four

spelling metrics (words correct, and three partial-credit metrics – correct letter sequences, sounds correct, and phonological coding) and early reading-related skills in kindergarten. Clemens et al. (2014) examined the concurrent (kindergarten) and predictive (kindergarten to grade 1) validity of spelling metrics to a variety of reading skills. Participants were children 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 System-Elements score. Results for the concurrent models indicated that the partial-credit metrics accounted for 6% to 14% more variance in reading than the words correct metric. Results were similar in the longitudinal model. These results suggest that spelling metrics are similar but those that give credit for partially correct spellings may reveal a stronger relation between spelling and reading. Furthermore, Treiman, Hulslander, et al. (2019) found that a partial-credit metric based on orthographic accuracy, measured at the end of kindergarten, was a significant, unique predictor of word reading in grades 1, 2, 4 and 9.

The present study is the first, of which we are aware, to examine whether scoring metrics matter in understanding the relation between spelling and reading comprehension. Based on the findings with word reading, however, we hypothesize that partial credit metrics may be better predictors of reading comprehension, particularly those that are based on orthographic accuracy.

The Present Study

The purpose of the present study was to investigate the contribution of spelling to reading comprehension while controlling for other lexical-level literacy skills, namely word recognition and vocabulary. Results using four different spelling scoring metrics were compared. The specific research questions are: (a) To what extent are the lexical-level literacy skills of spelling, word recognition, and vocabulary related to each other and to reading comprehension, and do the

relations vary with spelling scoring metric?, 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 spelling scoring metric? These research questions are theoretically supported and driven by the need to better understand the role that spelling may play in reading comprehension.

Method

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). 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 participants spoke English as their primary language. One child spoke Spanish as the primary language. Nonverbal intelligence, as measured with the Matrices subtest of the *Kaufman Brief Intelligence Test* (Kaufman & Kaufman, 1997), was average ($M = 103.90$, $SD = 19.23$; range 65-143). Fourteen percent of participants ($n = 9$) had an Individual Education Plan (29% unreported). Demographic information was collected from caregiver questionnaires. Caregivers filled out the questionnaires at home and did not always answer every question, resulting in unreported data of 8-30% across questions. 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. 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 (maximum group size was 6 children).

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 for 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), and are described in the supplemental materials (S1). Each of the selected measures is described in detail below. Sample-specific internal consistency reliabilities were estimated using Cronbach's alpha, and all were $> .80$ (Table 1). Since standard scores are not available for all measures, raw scores from each test were z -scored based on the study sample and the z -scores were used in our analyses.

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. Practice items were completed as a group and then students were instructed to read the test passages silently and answer the multiple-choice questions. Students were given a maximum of 35 minutes to complete the test. As reported by the test publisher, internal consistency reliability for this measure is .92 and criterion validity is .77 to .79.

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. Students pronounced as many pseudowords (Phonemic Decoding Efficiency subtest) and real words (Sight Word Efficiency subtest) as they could, with 45 seconds for each subtest. The test publishers report concurrent validity of .85 and .89 for the subtests. For our analyses, the z -scores from each subtest were averaged to form a composite word recognition score.

Vocabulary. Form A of the *Peabody Picture Vocabulary Test – 4th Edition* (PPVT; Dunn & Dunn, 2007) was administered to assess receptive vocabulary. Students selected a picture from a choice of four that represents the meaning of a word spoken by the examiner. 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*.

Spelling. Two measures of spelling ability were administered: *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 CBM is timed. Both WTW and CBM use a word dictation format and can be administered to groups. 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 two cases, both on the CBM). Students spelled the dictated words on test forms that were lined and numbered.

Words Their Way - Elementary (WTW). 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. Reliability and validity of the WTW inventories has been reported based on data collected in a school district in California (Sterbinsky, 2007). For internal consistency, Cronbach's alpha was reported as .915. For third grade, test-retest reliability was reported as .95 and concurrent validity ranged from .497 to .692.

Aimsweb Spelling Curriculum-Based Measure (CBM). The CBM was chosen as a measure of general spelling ability for grade-level words. It is a timed test, so it offers a measure of spelling fluency, and contains curriculum-independent graded word lists. For the present study, the winter benchmark spelling list for third grade was administered according to standardized instructions (Shinn & Shinn, 2002). Administration of the 17 test items was timed for two minutes, with new words dictated every seven seconds. Reliability and validity are not reported directly by the publishers; rather, research supporting the reliability and validity of spelling CBMs in general is cited (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. 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 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 later confirmed) and therefore composite scores were computed for use in analyses. The z -scores for each measure were averaged to create a composite. This procedure was completed for each of the 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 was 20 for WTW and 17 for 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 four 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 then averaged across words. We used the average and not total CLS so that the scale would be more similar to the other metrics. The maximum number of correct letter sequences possible on the WTW is 144 and on the CBM is 112; therefore, the maximum averaged CLS score is 7.20 for WTW (144 divided by 20 words) and 6.59 for CBM (112 divided by 17 words).

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. The *Computerized Spelling Sensitivity System* (CSSS; Masterson & Apel, 2010; Masterson & Hrbec, 2011) was used and procedures outlined in the CSSS Manual were followed. Using the CSSS, spellings are scored based on the linguistic accuracy of each element in the word; errors were analyzed and coded as being related to phonemic awareness, orthographic pattern knowledge, morphological awareness, or storage of word-specific mental representations. 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., doubling the <p> in *slipped*). 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 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 word-specific mental representations. The maximum possible score for both SSW and SSE is 3.

Coder training and intercoder reliability. For each of the partial-credit scoring metrics (CLS, SSE, SSW), scoring was conducted by the author (primary coder) and one of two graduate students. Coders were trained and procedures were followed to ensure that scoring was conducted in a reliable manner. Details are available in the supplemental materials (S2). Intraclass correlation coefficients (ICC) for intercoder reliability ranged from .99 to 1.00.

Results

All analyses were conducted using SPSS version 22. There were no missing data. All variables were standardized, as described above, to achieve uniformity in scaling. 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; all were met.

Question 1: Descriptives and Bivariate Relations

Descriptive statistics are presented in Table 1. Standard scores indicate that performance on reading comprehension, word recognition, and vocabulary was age appropriate. WTW is not normed, so standard scores are not available. 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. Comparing percentage of words correct across the two spelling measures, children performed slightly better on WTW than on CBM (e.g., 63.15% on WTW and 60.76% on CBM). The average SSW score of > 2 suggests that children's spellings were, generally, orthographically legal with errors occurring in word-specific mental 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 (r s ranging .802 to .841) and the lowest were between vocabulary and spelling (r s ranging .412 to .456). The spelling metrics were very highly correlated with each other (r s ranging .959 to .985), indicating almost complete overlap among them.

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%. WSC and word recognition each explained 41%. SSE and SSW explained 36.6% and 34.6%, respectively, 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 to reading comprehension while controlling for word recognition (Model A), vocabulary (Model B), and a full model with word recognition and vocabulary together (Model C). Results are presented in Table 3. Results for the SSW and SSE metrics were almost identical; therefore, for the sake of space, we report only the SSE results for each of the models.

Model A: Word recognition + Spelling. In Model A, we examined whether spelling contributed unique variance to reading comprehension beyond that explained by word recognition. The first set of models tested spelling using the WSC metric. The total variance in reading comprehension explained by word recognition and spelling was 42.8% (R^2_{adj}). The R^2 change for spelling was not significant (p value just above .05), meaning that spelling did not contribute unique variance after controlling for word recognition. Examining the individual predictors, neither word recognition nor spelling met traditional levels of significance in the model (p values just above .05). Using the CLS metric, the total variance explained was 45.7% (R^2_{adj}). The R^2 change for spelling was significant this time, meaning that it added significant unique variance to reading comprehension. Examining the individual predictors, spelling was significant but word recognition was not. Finally, using the SSE metric, the total variance in

reading comprehension explained was 40.6% (R^2_{adj}). The R^2 change was not significant. Word recognition was a significant predictor but spelling was not.

Model B: Vocabulary + Spelling. In Model B, we tested the contribution of spelling while controlling for vocabulary. For all metrics, the R^2 change for spelling was significant, meaning that spelling explained unique variance in reading comprehension, and vocabulary and spelling were both significant predictors. The variance explained in reading comprehension was 48.7 % (R^2_{adj}) using the WSC metric, 52.5 % (R^2_{adj}) using the CLS metric, and 45.3% (R^2_{adj}) using the SSE metric.

Model C: Word recognition + Vocabulary + Spelling. In this full model, we examined the unique contribution of spelling while controlling for both word recognition and vocabulary. At Step 1 using the WSC spelling metric, word recognition and vocabulary together explained 50.1% (R^2_{adj}) of the variance in reading comprehension. Both predictors were significant at this step. Adding spelling in Step 2 did not explain significant additional variance, and spelling was not a significant predictor in the model. Word recognition was also not significant, leaving vocabulary as the only predictor to make a unique contribution to reading comprehension when all three lexical-level skills were in the model together. Using the CLS metric, however, spelling explained additional unique variance in reading comprehension, for a total variance explained of 53.0%. Spelling and vocabulary were significant, but word recognition was not. Using the SSE metric, the total variance explained was 49.9%. Spelling did not explain unique variance, nor was it a significant predictor. Both word recognition and vocabulary were significant predictors.

Discussion

Gaining insight into the skills necessary for skilled reading comprehension is needed to inform practice related to the identification of children with or at risk for academic difficulties.

No research to date has examined the combined and unique contributions of all three lexical-level literacy skills to reading comprehension. Spelling, as a strong indicator of lexical quality, may also be related to children's reading comprehension (Perfetti & Stafura, 2014); therefore, it is imperative to gain a better understanding of the unique role that it may play. The present study is a first step in this direction, describing the lexical-level literacy skills of a sample of third-grade children and exploring the relation between spelling and reading comprehension. Additionally, this study investigated whether the relation between spelling and reading comprehension varied with spelling scoring metric.

The present work revealed several major findings. First, children's literacy skills were characterized by moderate to strong concurrent relations and a high degree of consistency across skills. Second, spelling was shown to be uniquely related to reading comprehension when using the CLS scoring metric and controlling for the other lexical-level skills. Third, the relation between spelling and reading comprehension varied slightly 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.

Correlations Amongst the Lexical-level Skills and With Reading Comprehension

Spelling was strongly related to word recognition in the present study, with correlations similar to or higher than those reported in other research between spelling and timed word recognition for children in grade 3 (e.g., Abbott et al., 2010; Ahmed, Wagner, & Lopez, 2014; Morris et al., 2012), with differences possibly due to the fact that one of our measures of spelling was also timed. The moderate correlation between spelling and vocabulary also aligns with other research (Chua, Rickard Liow, & Yeong, 2016; Cunningham & Stanovich, 1991; Kim, Apel, et al., 2013). There were slight differences across spelling scoring metric for the correlations with

word recognition (variance explained ranging from 64-70%) but not with vocabulary (variance explained ranging from 16.8-17.6%). Thus, in terms of understanding the relations between the lexical-level skills, there is not much variance with spelling scoring metric.

The correlations with reading comprehension were strong for spelling and word recognition, as expected, and moderate for vocabulary. We note that the correlation of vocabulary with reading comprehension (.56) is similar or slightly lower than that reported in other research for children in third grade, e.g., .54 and .63 in Tannenbaum, Torgesen, and Wagner (2006) and .62 and .65 in Quinn et al. (2015), with differences possibly due to differences in measures used. Turning to the spelling scoring metrics, we see a greater range in the variance explained for reading comprehension (35-46%) than we did with spelling and word recognition. Despite the fact that the spelling scoring metrics are very highly correlated with each other, there may be differences in how they relate to reading comprehension. We did not test whether the correlations were significantly different from each other, however. Also of note, for the WSC and CLS metrics, the correlation between spelling and reading comprehension was equal to or slightly higher than the correlation between word recognition and reading comprehension. This pattern has been reported in other research with children in third grade. For example, Ahmed et al. (2014) reported correlations with reading comprehension of .60-.66 for spelling and .51-.62 for word recognition, and Mehta, Foorman, Branum-Martin, and Taylor (2005) reported correlations of .73 for spelling and .66 for word recognition. Again, we did not test differences between the correlations, but these results highlight the strength of the relation between spelling and reading comprehension – which may be equal to the relation between word recognition and reading comprehension – and suggest that spelling scoring metric may matter in understanding the relationship.

The Unique Contribution of Spelling to Reading Comprehension

The results of our multiple hierarchical regressions reflect what we found in the bivariate correlations and support our hypothesis that spelling was significantly related to reading comprehension and that spelling explained unique variance when measured using the CLS metric. With CLS, spelling was significant while controlling for word recognition, vocabulary, and both combined, explaining an additional 6.5%, 22.2%, and 3.6% of variance, respectively. The unique contribution of spelling when controlling for word recognition was lower than that 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. However, the study was conducted in Italian – a language with an opaque orthography so results are not directly comparable – and the word recognition task involved text-level reading, not context-free word reading as in the present study. The unique contribution of spelling when controlling for vocabulary was much higher than that reported by Reed et al. (2016). They found that spelling accounted for 2% unique variance in reading comprehension, but this discrepancy may be explained by grade (they studied children in grade 6) and the fact that they measured vocabulary differently, using a written measure of depth of lexical knowledge, whereas a measure of oral vocabulary breadth was 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 while controlling for the other. The *p*-values for both were just above the traditional .05 level of significance. There are two points of note here. First, these results are likely a reflection of the high correlation between word recognition and spelling, as highly correlated predictors in a model can result in ‘competition’, 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), a post 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. Since spelling explained 3.7% unique variance, the shared variance was calculated to be 38.1% (total R^2 of 44.6% minus the unique contributions). Thus, word recognition and WSC share so much common variance in explaining reading comprehension that there is little room left for either predictor to contribute unique variance. Second, although word recognition and WSC were not significant together, 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 important practical effects. Furthermore, the sample size in the present study may have been insufficient to detect small but true effects.

Spelling Scoring Metrics

The third main finding of this study relates to how the association between spelling and reading comprehension varied with spelling scoring metric. Metrics based solely on orthographic accuracy – CLS and WSC – were more highly related to reading comprehension than were metrics that considered accuracy across multiple linguistic domains (i.e., phonological and morphological in addition to orthographic). CLS, a partial-credit metric, was the only spelling metric that was significant when controlling for word recognition. WSC, the whole-word accuracy metric, performed similarly to CLS when controlling for word recognition but did not meet traditional levels of significance. However, as noted in the preceding paragraph, WSC contributed more unique variance in reading comprehension than word recognition did. SSE and SSW, the multiple-linguistic partial-credit metrics, showed the weakest relation with reading comprehension. Considering the view that orthographic knowledge consists of lexical and

sublexical levels (Apel et al., 2019), our results indicate that it may be lexical spelling, which involves the stored mental representations of words (i.e., lexical orthographic accuracy), that is related to reading comprehension.

This pattern of metrics based solely on orthographic accuracy performing better than multiple-linguistic metrics aligns with recent research by Treiman, Kessler, et al. (2019). That study compared spelling metrics as predictors of later spelling performance, and our findings extend this result to the concurrent relation between spelling and reading comprehension. Our finding that CLS performed slightly better than WSC in our models aligns with (Clemens et al., 2014), who found that partial-credit spelling metrics explained more variance in word reading than the whole-word accuracy metric. It may be that scoring that allows for partial credit offers some advantage, at least with regards to metrics based solely on orthographic accuracy. In contrast to our study and the Treiman, Kessler, et al. study, however, Clemens et al. found that the partial-credit metric SSE also explained more variance than the whole-word metric. This discrepancy in findings may be due to the difference in dependent variable, which was a word reading factor in Clemens et al., reading comprehension in our study, and spelling in Treiman, Kessler, et al.

The difference across partial-credit metrics may be explained by lexical quality. In SSE/SSW scoring, some credit is given for words that are spelled with phonological accuracy but orthographic errors (e.g., *faver* for *favor*, and *flote* for *float*). This type of spelling may reflect low lexical quality of the child's representation for that word. High lexical quality requires phonological, orthographic, and morphological accuracy. In our sample of typically-developing children in third grade, it may be that measurement at the level of phonology is not sensitive or unique as a predictor of reading comprehension. Scoring for the CLS and WSC metrics is based

solely on orthographic accuracy, which could be considered to subsume accuracy in phonology as well, since spelling by sight requires proficiency in phonemic awareness (orthographic mapping; Ehri, 2014). In addition, we acknowledge that the reduced range in WSC scores (each word is scored 0 or 1) compared to those derived from CLS (range in score from 4-10) may impact the manner in which these variables relate to reading comprehension.

Implications

Several implications of the present research are evident. First, the results have theoretical implications regarding the simple view of reading and the lexical quality hypothesis. The simple view of reading states 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, Simos, Sideridis, and Mouzaki (2012) pointed out, however, it is unclear how these skills should be measured and indexed. Recently, researchers have used spelling as a ‘proxy’ for the decoding component (Kim, Wagner, & Lopez, 2012; Reed et al., 2016). We would argue, however, that spelling needn’t be a proxy for decoding; rather, the decoding component may be better conceptualized more broadly as a *print* component, to include both decoding (word recognition) and spelling. Decoding and spelling are both skills that require the application of orthographic knowledge, and therefore provide a window into children’s proficiency with print. The results of the present study provide preliminary support for this conceptualization in that spelling contributed equally or more highly to reading comprehension, at least when spelling was measured with the orthography-only metrics. The idea of reconceptualizing the decoding component of the simple view of reading as a print component has also been put forth by Protopapas et al. (2012).

Similarly, with regards to the lexical quality hypothesis, other researchers have operationalized lexical quality as vocabulary and decoding (Verhoeven & Van Leeuwe, 2008), but the results of the present study support that, at least for the prediction of reading comprehension, it may be better operationalized with spelling instead of word recognition. Using the WSC and CLS spelling metrics, we found that spelling and vocabulary together explained the same or slightly more variance in reading comprehension than word recognition and vocabulary. Our results offer support to Perfetti's claim that spelling may be a better indicator of underlying lexical quality than word recognition (Perfetti, 1992). More research is needed, but spelling should be a consideration in future investigations of the simple view of reading and lexical quality and how we conceptualize the print-related components of each.

The results of the present research also have important practical implications for SLPs and educators. Our finding that spelling explained equivalent or greater variance in reading comprehension than did word recognition supports the call of a growing number of researchers for spelling to be included in literacy assessment and screening batteries (Chua et al., 2016; Clemens et al., 2014; Morris et al., 2012; Ritchey et al., 2015; Robbins, Hosp, Hosp, & Flynn, 2010; Treiman, Hulslander, et al., 2019). In current practice, word recognition is frequently assessed as a predictor of future reading comprehension, but spelling is not. As Treiman, Hulslander, et al. (2019) state, spelling is not simply a proxy for phonological awareness and alphabet knowledge, nor is it redundant with tests of word recognition; rather, it provides valuable information about children's literacy skills. Gathering such data about a child's early skills is important for identifying children who may be a risk for reading comprehension difficulties.

An added practical benefit of spelling assessment is that it can be group administered, making it an efficient assessment tool in the classroom, unlike tests of word reading which have to be individually administered. SLPs should consider adding spelling assessment to their test batteries and advocating that schools do the same, including in the lower grades before reading comprehension has developed (and therefore cannot reliably be measured). Spelling assessment provides insight into children's sublexical and lexical skills and may help inform our decisions about which children may be at risk for reading comprehension difficulties. A spelling measure could be added to screening batteries in addition to word recognition, or, considering that our model with CLS and vocabulary performed as well as the full model, spelling may even be able to replace word recognition. Of course, longitudinal research and testing and development of a sensitive measure of spelling would be needed before considering such a recommendation.

Additionally, our findings provide preliminary support for a possible advantage of using the CLS scoring metric. Across all the models examined, the highest variance explained in reading comprehension was achieved using CLS, and CLS was the only spelling metric that was a unique predictor when combined with word reading. We suggest caution in interpreting this possible advantage, given the high correlation amongst all the scoring metrics and the fact that we did not test whether the correlations between the metrics and reading comprehension were significantly different from each other, but SLPs should be aware of the different methods of scoring spelling and the purposes and advantages of each. For example, the Spelling Sensitivity System was designed to measure developmental changes in spelling and its detailed error analysis provides information that can be useful for selecting intervention targets (Masterson & Apel, 2010), but it was the least related to reading comprehension amongst our metrics. CLS is likely the least well known amongst SLPs but our results suggest that it may be a useful metric in

capturing how spelling is related to reading comprehension, at least for children in third grade. SLPs should become familiar with the CLS and how to do the scoring, and this could be a collaborative opportunity with teachers. SLPs can also feel confident that even the most simple metric, the number of words spelled correctly, is strongly related to reading comprehension.

Overall, the present study adds to the growing body of literature on the importance of spelling as a literacy construct, in general, and its relation to reading comprehension in particular. Considering other research showing that spelling instruction transfers to word recognition (Graham & Santangelo, 2014), and that orthographic learning via spelling transfers to word recognition (Conrad et al., 2019), it may be that good spelling skills 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 & Stafura, 2014).

Limitations and Future Directions

The primary limitation of the present study is its small sample size and inherent methodological limitations. Future research using larger sample sizes could employ a latent variable approach and more advanced statistical modeling. Another limitation of our sample was that participants, for the most part, were from mid- to high-SES, English-speaking families in a single Midwest state. This is not a representative sample and results are therefore not generalizable to the population at large. Future research should explore the relations between spelling and reading comprehension in diverse populations such as children with language and/or reading difficulties and children who are English language learners.

The present study was also limited in its correlational design and its investigation of concurrent relations only. Spelling data were available at only a single time point, so

examination of longitudinal predictive relations was not possible. While evidence exists to suggest that spelling is predictive of reading comprehension over time (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 both word recognition and vocabulary. To gain a better understanding of the predictive ability of spelling over time 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 and age levels would also be beneficial. It may be the case, for example, that metrics capturing phonological accuracy may be important when children are first learning to spell.

Conclusion

This study adds to the small but growing body of literature on the importance of spelling for reading comprehension. In both research and practice, the focus has been on the other lexical-level skills of word recognition and vocabulary, but the results of this study suggest that more effort should be directed toward spelling. Although more research is needed, SLPs should consider adding spelling, and different scoring methods, to their assessment protocols. A sensitive measure of spelling may provide additional information beyond a measure of word recognition. Future research should explore whether spelling ability is predictive of reading comprehension over time and therefore could be used to identify children at risk for reading comprehension difficulties.

Acknowledgments

Many thanks are extended to the Language and Reading Research Consortium for providing access to data, support in collecting the spelling data, and doctoral funding; to dissertation committee members Laura Justice, Richard Lomax, and Monique Mills, and mentor Shayne Piasta, for their support and guidance; and to Patricia Biancone for her assistance with coding.

References

- Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281-298. doi:10.1037/a0019318
- Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology, 106*(2), 419-434. doi:10.1037/a0035692
- Apel, K., Henbest, V. S., & Masterson, J. (2019). Orthographic knowledge: Clarifications, challenges, and future directions. *Reading and Writing, 32*(4), 873-889. doi:10.1007/s11145-018-9895-9
- Bear, D. R., Invernizzi, M., Templeton, S., & Johnston, F. (2008). *Words their way: Word study for phonics, vocabulary, and spelling instruction*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*(2), 144-179. doi:10.3102/0034654309359353
- Chua, S. M., Rickard Liow, S. J., & Yeong, S. H. M. (2016). Using spelling to screen bilingual kindergarteners at risk for reading difficulties. *Journal of Learning Disabilities, 49*(3), 227-239. doi:10.1177/0022219414538519
- Clemens, N. H., Oslund, E. L., Simmons, L. E., & Simmons, D. (2014). Assessing spelling in kindergarten: Further comparison of scoring metrics and their relation to reading skills. *Journal of School Psychology, 52*(1), 49-61. doi:dx.doi.org/10.1016/j.jsp.2013.12.005

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, New Jersey: Erlbaum.
- Conrad, N. J., Kennedy, K., Saoud, W., Scallion, L., & Hanusiak, L. (2019). Establishing word representations through reading and spelling: Comparing degree of orthographic learning. *Journal of Research in Reading, 42*(1), 162-177. doi:10.1111/1467-9817.12256
- Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*(2), 264-274. doi:10.1037/0022-0663.83.2.264
- Desimoni, M., Scalisi, T. G., & Orsolini, M. (2012). Predictive and concurrent relations between literacy skills in grades 1 and 3: A longitudinal study of Italian children. *Learning and Instruction, 22*(5), 340-353. doi:dx.doi.org/10.1016/j.learninstruc.2012.02.002
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Minneapolis, MN: Pearson.
- Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*(3), 19-36. doi:Doi 10.1097/00011363-200020030-00005
- Etmanskie, J. M., Partanen, M., & Siegel, L. S. (2016). A longitudinal examination of the persistence of late emerging reading disabilities. *Journal of Learning Disabilities, 49*(1), 21-35. doi:10.1177/0022219414522706
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. E., Michaels, R., & Shaywitz, S. E. (2015). Achievement gap in reading is present as early as first grade and persists through adolescence. *The Journal of Pediatrics, 167*(5), 1121-1125.e1122. doi:10.1016/j.jpeds.2015.07.045

- Foorman, B. R., & Petscher, Y. (2010). Development of spelling and differential relations to text reading in grades 3-12. *Assessment for Effective Intervention, 36*(1), 7-20.
doi:10.1177/1534508410379844
- Foorman, B. R., Petscher, Y., & Bishop, M. D. (2012). The incremental variance of morphological knowledge to reading comprehension in grades 3–10 beyond prior reading comprehension, spelling, and text reading efficiency. *Learning and Individual Differences, 22*(6), 792-798. doi:dx.doi.org/10.1016/j.lindif.2012.07.009
- Frith, U. (1980). Unexpected spelling problems. In *Cognitive processes in spelling* (pp. 495-515). London: Academic Press.
- García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in english. *Review of Educational Research, 84*(1), 74-111.
doi:10.3102/0034654313499616
- Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703-1743.
doi:10.1007/s11145-014-9517-0
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*(2), 127-160. doi:10.1007/BF00401799
- Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10-17.
doi:<https://doi.org/10.1080/1045988X.2003.10871074>
- Kaufman, A. S., & Kaufman, N. L. (1997). *Kaufman brief intelligence test* (2nd ed.). Minneapolis, MN: Pearson.

- Kim, Y.-S., Apel, K., & Al Otaiba, S. (2013). The relation of linguistic awareness and vocabulary to word reading and spelling for first-grade students participating in response to intervention. *Language, Speech, and Hearing Services in Schools, 44*(4), 337-347. doi:10.1044/0161-1461(2013/12-0013)
- Kim, Y.-S., Petscher, Y., & Foorman, B. (2013). The unique relation of silent reading fluency to end-of-year reading comprehension: Understanding individual differences at the student, classroom, school, and district levels. *Reading and Writing, 28*(1), 131-150. doi:10.1007/s11145-013-9455-2
- Kim, Y.-S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from grade 1 to grade 2. *Journal of Experimental Child Psychology, 113*(1), 93-111. doi:<http://dx.doi.org/10.1016/j.jecp.2012.03.002>
- Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*(2), 151-169. doi:10.1002/rrq.99
- Language and Reading Research Consortium, Farquharson, K., & Murphy, K. A. (2016). Ten steps to conducting a large, multi-site, longitudinal investigation of language and reading in young children. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.00419
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie reading tests (4th ed.)*. In. Itasca, IL: Riverside Publishing.
- Masterson, J., & Apel, K. (2010). The spelling sensitivity score: Noting developmental changes in spelling knowledge. *Assessment for Effective Intervention, 36*(1), 35-45. doi:10.1177/1534508410380039

- Masterson, J., & Hrbec, B. (2011). Computerized spelling sensitivity system [computer software]. Springfield, MO: Missouri State University Language-Literacy Lab. Retrieved from <https://www.missouristate.edu/csd/LLL/CSSS.htm>
- Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading, 9*(2), 85-116. doi:10.1207/s1532799xssr0902_1
- Morris, D., Trathen, W., Lomax, R. G., Perney, J., Kucan, L., Frye, E. M., . . . Schlagal, R. (2012). Modeling aspects of print-processing skill: Implications for reading assessment. *Reading and Writing, 25*(1), 189-215. doi:10.1007/s11145-010-9253-z
- National Center for Education Statistics. (2017). *The nation's report card: 2015 mathematics and reading assessments*. Washington, DC: U.S. Department of Education.
- NCS Pearson Education. (2012). *AimswEB technical manual*. In. Retrieved from <http://www.aimswEB.com/wp-content/uploads/aimswEB-Technical-Manual.pdf>.
- Ouellette, G. P. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology, 107*(1), 50-58. doi:dx.doi.org/10.1016/j.jecp.2010.04.009
- Ouellette, G. P., & Sénéchal, M. (2008). A window into early literacy: Exploring the cognitive and linguistic underpinnings of invented spelling. *Scientific Studies of Reading, 12*(2), 195-219. doi:10.1080/10888430801917324
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

- Perfetti, C. A. (1997). The psycholinguistics of spelling and reading. In C. A. Perfetti, L. Rieben, & M. Fayol (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 21-38). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi:10.1080/10888430701530730
- Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, the Netherlands: John Benjamins.
- Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(1), 22-37. doi:10.1080/10888438.2013.827687
- Pinto, G., Bigozzi, L., Tarchi, C., Accorti Gamannossi, B., & Canneti, L. (2015). Cross-lag analysis of longitudinal associations between primary school students' writing and reading skills. *Reading and Writing, 28*(8), 1233-1255. doi:10.1007/s11145-015-9569-9
- Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology, 33*(3), 217-240. doi:10.1080/02702711.2010.507626
- Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*(1), 159-175. doi:10.1111/cdev.12292
- Reed, D. K., Petscher, Y., & Foorman, B. R. (2016). The contribution of vocabulary knowledge and spelling to the reading comprehension of adolescents who are and are not english language learners. *Reading and Writing, 1-25*. doi:10.1007/s11145-015-9619-3

- 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
- Ritchey, K. D., Coker, D. L., & McCraw, S. B. (2010). A comparison of metrics for scoring beginning spelling. *Assessment for Effective Intervention, 35*(2), 78-88.
doi:10.1177/1534508409336087
- Ritchey, K. D., Silverman, R. D., Schatschneider, C., & Speece, D. L. (2015). Prediction and stability of reading problems in middle childhood. *Journal of Learning Disabilities, 48*(3), 298-309. doi:10.1177/0022219413498116
- Robbins, K. P., Hosp, J. L., Hosp, M. K., & Flynn, L. J. (2010). Assessing specific grapho-phonemic skills in elementary students. *Assessment for Effective Intervention, 36*(1), 21-34. doi:10.1177/1534508410379845
- Shahar-Yames, D., & Share, D. L. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading, 31*(1), 22-39. doi:10.1111/j.1467-9817.2007.00359.x
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151-218. doi:dx.doi.org/10.1016/0010-0277(94)00645-2
- Shinn, M. R., & Shinn, M. M. (2002). Aimsweb training workbook: Administration and scoring of spelling curriculum-based measurement (s-cbm) for use in general outcome measurement. *Obtained through <http://www.aimsweb.com>.*
- Sterbinsky, A. (2007). Words their way spelling inventories: Reliability and validity analyses. *Center for Research in Educational Policy, The University of Memphis.*

- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398. doi:10.1207/s1532799xssr1004_3
- The Annie E. Casey Foundation. (2012). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation. Retrieved from <http://www.aecf.org/resources/double-jeopardy/>.
- Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency* (2nd ed.). Austin, TX: Pro-Ed.
- Treiman, R., Hulslander, J., Olson, R. K., Willcutt, E. G., Byrne, B., & Kessler, B. (2019). The unique role of early spelling in the prediction of later literacy performance. *Scientific Studies of Reading, 1*-8. doi:10.1080/10888438.2019.1573242
- Treiman, R., Kessler, B., & Caravolas, M. (2019). What methods of scoring young children's spelling best predict later spelling performance? *Journal of Research in Reading, 42*(1), 80-96. doi:10.1111/1467-9817.12241
- Treiman, R., Kessler, B., Pollo, T. C., Byrne, B., & Olson, R. K. (2016). Measures of kindergarten spelling and their relations to later spelling performance. *Scientific Studies of Reading, 20*(5), 349-362. doi:10.1080/10888438.2016.1186168
- Verhoeven, L., & van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*(3), 407-423. <https://doi.org/10.1002/acp.1414>

Supplemental Materials

S1: Choice of assessments. This is an explanation of how measures from the larger study were selected for inclusion in the present study.

S2: CSSS details. This describes procedures related to use of the CSSS software; specifically, how new words were entered into the dictionary, as well as coder training and intercoder reliability.