EXAMINING THE QUICK SPELL TEST: HOW DOES IT RELATE TO PHONOLOGICAL PROCESSES, NAMING SPEED, ORTHOGRAPHIC PROCESSING, AND READING?

by

Sandy S. Y. Lai

Queen's University
Kingston, Ontario, Canada
August 2007

ABSTRACT

I examined the relationship between performance on the Quick Spell Test (Bowers, 1996) and reading ability in an archival sample of Grade 3 (*N* = 192) and Grade 4 (*N* = 149) children from Kirby, Parrila, Deacon, and Wade-Woolley's (2004-2007) longitudinal study. The data included a battery of phonological awareness, naming speed, phonological decoding, orthographic processing, and reading ability measures administered in both grades. I found that the Quick Spell Test was a good concurrent predictor of reading ability in Grade 3, but not a good longitudinal predictor. Orthographic processing and phonological decoding were consistently good predictors of the relationships between the Quick Spell Test and reading ability and of the difference in variance between the Quick Spell Test subscales.

ACKNOWLEDGEMENTS

First I would like to thank Pat Bowers for making the Quick Spell Test materials available and for providing me with detailed background information.

Second, I would like to thank my supervisor, John, for his wonderful mentorship and valuable resources. I've had an enjoyable time in his lab, and this experience has not only initiated my contact with the educational world but has also opened the door to my future endeavours.

Third, I would like to thank all the faculty and peers with whom I've been able to take courses during this past year. Their collective wisdom and experiences, their enthusiasm for education and research, and their diverse personalities have definitely created an unforgettable academic year.

This thesis concludes my time here at Queen's University. The past five years have been truly amazing and I am grateful for everyone who has made my experience so memorable.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: EXAMINING THE QUICK SPELL TEST

Because the modern, knowledge-based society relies heavily upon written communication, early education has placed great emphasis on reading instruction. Proficient reading involves the mastery of skills that enable the reader to interpret text by extracting ideas from letters, graphemes, words, and syntax (Heath, 1991), and is dependent on numerous cognitive processes in which individuals differ. To the extent that these cognitive processes are responsive to change, interventions targeting these factors may improve reading ability. Moreover, many of the cognitive skills required for reading may emerge and be assessable prior to formal reading instruction. Therefore, to the extent that these early reading skills are measurable and reading proficiency is predictable from pre-reading assessment, educators can evaluate children's individual reading potential and provide any necessary support as soon as possible. In light of these possibilities, many researchers and educators have been attempting to identify these key contributors to reading, as a deeper understanding of the underlying mechanisms is necessary for devising effective education and intervention practices.

For many years, research mainly focused on the predictive power of phonological skills, such as phonological awareness (PA) and phonological decoding (PD), which enable the reader to convert written text into oral language by parsing individual graphemes into their corresponding phonemes. However, researchers (e.g., Bus & van IJzendoorn, 1999) have since argued that phonological processes are insufficient in explaining all of the variance in reading ability and have more recently turned their attention to additional predictors. In particular, naming speed (NS) and orthographic processing (OP) have been widely researched, and many researchers have proposed theories explaining their relationship to reading. One account in particular posits NS as a precursor of OP, arguing that NS affects the rate at which readers can acquire good OP skills (e.g., Wolf & Bowers, 1999); however, the literature evaluating this hypothesis has yet to draw confident conclusions.

Bowers (1996) developed the Quick Spell Test (QST) with the goal of examining the proposed relationship between NS and OP. In the QST, readers recall four-letter strings immediately after a brief exposure. The stimuli are categorized as real words, pseudowords (i.e., words conforming to English regularities but not in the English lexicon), and consonant clusters. Because the real words and pseudowords have greater orthographic structure than the consonant clusters, readers should report these stimuli more accurately. Moreover, readers with reading difficulties should have more trouble performing the QST than normal readers. If NS is related to OP, readers with NS deficits should show a distinct pattern of performance on the QST that is distinguishable from the performance of normal readers and poor readers with phonological deficits. However, the literature on the QST has been scarce and inconclusive. The studies have also not examined the relationship between QST and reading, and whether the QST measures the processes that Bowers (1996) designed it to measure.

In the current study, I use regression analyses to determine the predictive ability of QST in both concurrent and longitudinal reading achievement, and the relationship of the four major reading predictors (i.e., PA, NS, PD, and OP) and the QST. The data analyzed in this study are extracted from Kirby, Parrila, Deacon, and Wade-Woolley's (2004-2007) longitudinal study, which followed a cohort of children from Grades 3 to 5. The findings of this study will add not only to the QST literature, but also the general literature on reading ability in normal samples. Moreover, the evaluation of the QST's validity will be useful in determining whether practitioners and educators can use this measure as a viable means of testing reading skills in young or pre-readers.

The following sections give an overview of the literature regarding these four critical predictors of reading: PA, PD, NS, and OP. Table 1 shows a list of the abbreviations in this study. First, I describe the characteristics of these four reading predictors, and discuss the research assessing their independent predictive abilities in reading and some of the hypothesized relationships between the constructs. Next I explain the QST measure and review the findings of the QST literature. Finally I describe the current study and its potential contribution to this growing literature.

Table 1

List of Abbreviations in Alphabetical Order

Term	**Abbreviation**
High Frequency	HF
Low Frequency	LF
Naming Speed	NS
Orthographic Processing	OP
Phonological Awareness	PA
Phonological Decoding	PD
Quick Spell Test	QST
Test of Word Reading Efficiency	TOWRE

Literature Review

Phonological Awareness

Phonological awareness (PA) refers to the oral language ability to perceive and manipulate the sound components of words, such as the basic speech units of phonemes, rimes, and syllables (Adams, 1990). This ability to identify the individual sounds in speech strengthens children's ability to remember or see individual letters and spelling patterns, and the development of PA, particularly at the phonemic level, may require explicit instruction. Some researchers also argue that poor PA hinders the acquisition of the "alphabetic principle," so that understanding the relationship between letters and sounds is difficult (Stanovich, 1986). Common tasks that measure PA are Phoneme Elision, in which children must repeat a given word but with a specified phoneme deleted (e.g., "cat" without the "/k/"), and Word Blending, in which children must combine a series of separated phonemes into a word (e.g., "k-ă-t" into "cat").

In the reading literature, PA has surfaced as one of the most prominent and powerful factors in reading development (see Adams, 1990, and Goswami & Bryant, 1990, for reviews). For example, Hogan, Catts, and Little (2005) found through a path analysis that a kindergarten measure of PA accounted for unique variance in Grade 2 word reading after controlling for kindergarten letter identification ability. In a longitudinal study of PA, Sprugevica and Høien (2003) followed 53 children from the end of kindergarten to the middle of Grade 2. A kindergarten PA factor, formed from phonemic segmentation and phoneme deletion tasks, predicted 27.4% of the variance in word reading and 15.9% of the variance in sentence reading in the middle of Grade 1, 9.3% of the variance in word reading at the end of Grade 1, and 13.3% of the variance in sentence reading in the middle of Grade 2. It also significantly predicted 4.9% of the variance in word reading in the middle of Grade 1 after controlling for word reading scores at the end of kindergarten, and 15.2% of the variance in the slope of the word reading growth curve. In another longitudinal study, Kirby, Parrila, and Pfeiffer (2003) followed 79 children from kindergarten to Grade 5. They measured PA in kindergarten using four tasks: phoneme elision, phoneme blending, syllable blending, and sound isolation (i.e., identifying a sound from the beginning, middle, or end of a word). After controlling for general mental ability and letter recognition, kindergarten PA was a strong individual predictor of reading ability from kindergarten to Grade 2. Thus the literature has consistently shown PA to be a strong predictor of concurrent and later reading ability, particularly in the early grades.

Phonological Decoding

Closely related to PA, phonological decoding (PD) refers to the ability to convert written letters into sounds by applying grapheme-phoneme correspondence rules. In other words, PD is the phonic analysis of print. Accurate decoding requires not only familiarity with phonemes and sounds but also insight into the “alphabetic principle” and understanding that the writing system represents sound (Bowey, 1996). Because decoding involves more than merely manipulating sounds, PA is therefore a necessary but insufficient prerequisite for PD. For example, being able to segment the ‘/k/’ sound in the oral word “cat” is not equivalent to being able to translate the ‘c’ grapheme into the ‘/k/’ phoneme. Both PA and PD are especially important to beginning reading, as early readers must often rely on decoding to manage the high frequency of unfamiliar words when processing print. Share (1995) also argued that PD was the self-teaching mechanism through which beginning readers develop both word-specific and general orthographic knowledge, subsequently leading to efficient holistic word recognition. Specifically, decoding allows beginning readers to pay attention to the individual letters and orthographic units in words, thereby providing a solid foundation for the acquisition or refinement of an orthographic representation. Moreover, even the most proficient readers must rely on PD when encountering unfamiliar words.

A typical measure of PD is pseudoword reading, which involves reading non-words that conform to the phonological and orthographic structures of English. Pseudowords are only pronounceable through decoding, as they are not part of the reader's lexicon and therefore cannot be recognized holistically. Correlation coefficients between pseudoword reading and word recognition in the early elementary grades generally exceed .70, indicating that PD accounts for a majority of the variance in word recognition (Share, 1995). In a longitudinal study, Jorm, Share, Maclean, and Matthews (1984) divided kindergarten children into a group with no PD skills and a group with some PD skills, based on their scores on a pseudoword reading task. The two groups did not differ on any other measures in kindergarten. The group with some PD skills performed significantly better than the group with no PD skills in subsequent reading measures in both Grades 1 and 2. More recently, Kyte and Johnson (2004) administered a lexical decision task to 32 children, half of whom decided between real words and pseudowords by reading them aloud and half of whom performed the task while concurrently articulating a nonsense syllable. One day later, the researchers administered measures in which children had to identify the pseudowords they had encountered the previous day. These measures assessed whether the children remembered both the correct pronunciation and spelling of the target pseudowords. Children in the read aloud condition, which promoted PD, performed better than children in the concurrent articulation condition, which inhibited PD, providing support for the role of PD in orthographic learning and for Share's (1995) self-teaching hypothesis.

Although phonological skills have consistently predicted a large amount of variance in reading achievement, they do not account for all of the variance in reading (e.g., Wolf & Bowers, 1999). For instance, Share's (1995) claim that the correlation between PD and word recognition was .70 indicates that approximately 50% of the variance in reading is yet unexplained. Skilled readers also often recognize words immediately and holistically without requiring any grapheme-phoneme conversion. In addition, the literature on reading disabilities describes some individuals, known as surface dyslexics, who have adequate phonological skills but nevertheless are impaired in reading (e.g., Castles & Coltheart, 1993; McDougall, Borowsky, MacKinnon, & Hymel, 2005). Therefore, researchers have started seeking other predictors of reading achievement that would explain additional variance. The next sections discuss two of the leading candidates in the literature: naming speed (NS) and orthographic processing (OP).

Naming Speed

Naming speed (NS), also known as rapid automatized naming or speed of lexical access, refers to the speed at which children can name a set of stimuli in any of four symbolic categories: letters, numbers, colours, and objects (Wolf, O'Rourke, Gidney, Lovett, Cirino, & Morris, 2002). Wolf (1997) argued that NS represents a microcosm of reading, in that both processes require a rapid and precise integration of different cognitive systems. In particular, rapid naming involves attention to the stimuli, integration of visual information with stored orthographic (or logographic in the case of objects) and phonological representations, retrieval of phonological labels, and activation of articulation (Wolf & Bowers, 1999). The literature has consistently shown NS to independently predict reading both concurrently (e.g., Bowers, Steffy, & Swanson, 1986; Plaza & Cohen, 2003) and longitudinally (e.g., Kirby et al., 2003; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004). The best known measures of NS are the Rapid Automatized Naming Tests (RAN; Denckla & Rudel, 1976), which involve presenting a visual array of 50 stimuli consisting of five high-frequency digits, letters, colours, or objects that are presented ten times in random order. The NS is the speed at which a child serially or continuously names the array. Although research has found that discrete naming tasks also predict reading (e.g., Bowers & Swanson, 1991), continuous naming measures are considered to be more reflective of continuous reading because of their heightened timing demands and integration of other visual and inhibitory subprocesses (e.g., Bowers, 1995).

Moreover, all four types of NS stimuli predict later reading abilities when measured in kindergarten or prior to reading; however, in later grades, alphanumeric stimuli (i.e., letters and digits) predict reading ability better than non-alphanumeric stimuli (i.e., colours and objects; Wolf, 1991). This differentiation between NS stimuli appears because children automatize alphanumeric symbols during the first years of formal school instruction and therefore use different subprocesses for naming alphanumeric stimuli than for naming non-alphanumeric stimuli. Specifically, learning to unitize orthographic codes (i.e., alphanumeric stimuli) allows readers to have more working memory capacity for higher-order processes such as word reading and reading comprehension (LaBerge & Samuels, 1974). For example, Wolf, Bally, and Morris (1986) found that all kindergarten measures of NS predicted subsequent Grade 2 reading measures. However, when measured in Grades 1 or 2, non-alphanumeric symbols only weakly predicted Grade 2 single-word reading and reading comprehension whereas alphanumeric stimuli predicted all Grade 2 reading measures. Similarly, van den Bos, Zijlstra, and lutje Spelberg (2002) examined various age groups and found that the relationship between NS and word reading speed increased with age only for alphanumeric stimuli and not for non-alphanumeric naming. Therefore, researchers should determine the suitability of their NS measures for the age group of their respective samples.

The literature has also documented considerable debate on whether NS merely reflects phonological skills or pure processing speed, and whether NS uniquely predicts reading growth. First, some researchers have claimed that NS is not phonological in nature because it has consistently accounted for variance in reading after controlling for the effects of PA (see Wolf, Bowers, & Biddle, 2000 for a review). For example, Manis, Doi, and Bhadha (2000) examined 85 Grade 2 children and found NS to significantly predict 4.3 to 27.7% of the variance in various reading tasks, after controlling for vocabulary and PA. In addition, Swanson, Trainin, Necoechea, and Hammill (2003) synthesized the correlational research on PA, NS, and reading and found a mean correlation between PA and NS of .38. They also found that measures of NS and PA did not load onto the same factor in a factor analysis. Therefore NS and PA appear to be distinct, though correlated, constructs. On the other hand, other researchers have defined phonological processing to include not only PA, but also NS and PD (e.g., Wagner &

Torgesen, 1987). In this definition of phonological processing, NS measures the speed of access to phonological codes and is necessary for successful recoding of written symbols into their phonological representations. Thus, in this view, NS must be considered phonological in nature because it is included in the definition of phonological processing.

Second, Kail, Hall, and Caskey (1999) argued that the NS predicted reading because of age-related changes in general processing speed, which refers to the speed with which children can complete a series of simple cognitive tasks. However, Bowers, Storey, and Ferguson (2004) have shown NS to uniquely predict 11% of variance in word reading after controlling for the effects of both age and general processing speed. Thus NS appears to reflect a skill more specific to reading than pure processing speed. Third, Torgesen, Wagner, Rashotte, Burgess, and Hecht (1997) found that Grade 2 alphanumeric measures of NS did not predict Grade 4 reading measures beyond the variance accounted for by Grade 2 reading ability. Similarly, Grade 3 alphanumeric NS did not predict Grade 5 reading after accounting for Grade 3 reading ability. The researchers therefore concluded that NS was not a unique predictor of reading development. Conversely, Grade 1 non-alphanumeric measures of NS predicted Grade 3 reading measures after controlling for Grade 1 reading ability in Parrila, Kirby, and McQuarrie's (2004) longitudinal study. The discrepancy between these two studies may be due to differences in measures of NS (i.e., alphanumeric vs. non-alphanumeric stimuli) and age groups examined. Hence, further research should examine whether NS predicts reading development using more consistent methodology.

Another explanation that researchers have posited to explain the relationship between NS and reading ability is the orthographic processing speed account, in which

NS may likely be a precursor of orthographic processing (OP) abilities in children (Bowers & Wolf, 1993; Wolf & Bowers, 1999). Specifically, NS may affect the rate at which children can make connections between letter patterns and their respective sounds (i.e., grapheme-phoneme conversion), the quality of the orthographic codes in memory, and the amount of practice needed to ensure adequate orthographic coding of letter patterns. Thus readers with poor NS do not register letter sequences quickly enough to "chunk" them into orthographic units and instead process the letters independently for an extended period of time (Bowers & Newby-Clark, 2002), thereby hindering the acquisition of good OP skills. Although the research on this hypothesis has been sparse, some studies have provided evidence for the association between NS and OP. For example, Bowey et al. (2004) examined a sample of 125 children in Grade 4 to 6 and found that alphanumeric measures of NS predicted an additional 12.6% of the variance in word reading skill after controlling for the effects of age and non-alphanumeric measures of NS. In a sample of Grade 4 readers, Bowey, McGuigan, and Ruschena (2005) subsequently showed alphanumeric NS to predict 12.5% of the variance in word reading after controlling for global processing speed, non-verbal IQ and non-alphanumeric NS and 5.9% of the variance in word reading after also controlling for articulation rate. Moreover, alphanumeric measures of NS accounted for more variance in orthographic tasks than in PA measures in Manis et al.'s (2000) study, providing further evidence for the association between NS and OP. However, further research is needed to sufficiently evaluate the orthographic processing speed account.

Orthographic Processing

Because writing systems have specific conventions and rules concerning the visual and orthographic aspects of print, reading requires knowledge about orthographic constraints and regularities (Levy, Gong, Hessels, Evans, & Jared, 2006). Moreover, although all regular words can be read through decoding, PD is a slow and labourious process that requires constant effort and is also futile when reading irregular words. Consequently, good readers also use orthographic processing (OP), which refers to a more visual and holistic reading skill and relies on memory for letters, letter patterns, and words (Barker, Torgesen, & Wagner, 1992). OP is partially dependent on phonological skills, in that readers may acquire orthographic knowledge through repeated successful experiences of decoding words (Share, 1995). Thus accurate decoding creates and reinforces an association between a word's visual and phonological representations. Similarly, OP is also related to print exposure, as repeated exposure to print leads to a stable visual representation of an orthographic unit, which may be a word or a subunit of a word (Stanovich & West, 1989). Nevertheless, despite OP's relationships with other reading predictors, phonological skills and print exposure do not fully explain OP's ability to predict reading. For example, Barker et al. (1992) examined a sample of Grade 3 children and found that two OP tasks significantly accounted for 8 to 14% of the variance in various reading measures after controlling for the effects of age, IQ, phonological abilities, and print exposure. Therefore, OP contributes to reading independently of phonological abilities and print exposure.

Tasks measuring OP usually attempt to reduce the reader's reliance on phonological processing. Common OP tasks involve stimuli for which phonological processing would not produce an accurate response, such as irregular words (i.e., words that do not conform to standard grapheme-phoneme rules and therefore cannot be accurately decoded), homophones, and pseudohomophones (i.e., nonwords which sound the same as real words). For example, Cunningham, Perry, and Stanovich (2001) used six measures of OP in their study: three Wordlikeness tasks, which asked the children to identify the non-word letter combination which most closely resembled a real word from a pair of two non-word letter combinations (e.g., beff-ffeb); two Orthographic Choice tasks, which asked the children to identify the orthographically-correct word from a pair consisting of a real word and its pseudohomophone (e.g., rain-rane); and a Homophone Choice task, which presented the children with a question (e.g., "Which is a part of the body?") and asked them to identify the appropriate response from a pair consisting of a context-appropriate word and its context-inappropriate homophone (e.g., feat-feet). These six OP measures moderately converged onto one factor in a principal components analysis, indicating that they measured the same construct. A Grade 2 composite measure of OP formed from these six tasks also accounted for 16.3% of the variance in Grade 3 word recognition after controlling for Grade 1 phonological abilities and explained 8.9% of the variance in Grade 3 word recognition after controlling for Grade 2 PA and Grade 3 PD.

Although the literature generally poses OP skills to be measurable only after reading acquisition and print exposure, Levy et al. (2006) were able to measure young children's understanding of the orthographic aspects of print prior to reading acquisition. In their study, 4- to 6-year-old children were to choose the correct option when presented with two flash cards containing either two words or two sentences: one flash card was a correct representation of the word/sentence and the other flash card depicted a violation of an English print convention. Violations of word shape included using scribbles or pictures instead of letters or words, spacing out the letters in a word or words in a sentence inappropriately, writing the word/sentence onto more than one line, and using a single letter to represent a whole word (e.g., “W” instead of “weave”). Violations of word element included interspersing numbers among letters, repeating a single letter to represent all of a word/sentence (e.g., “Sssss” instead of “swamp”), and writing the word/sentence backwards or upside-down. Violations of spelling included using only vowels or consonants in a word, and misspelling words with a pseudohomophone (e.g., “perce” instead of “purse”). Principal components analyses showed that the three categories of violations developed separately for 4-year-olds, but that the word shape violations had reached ceiling performance in the 5- and 6-year-olds and hence only the word element and spelling components were developing. Using component scores for the three violation categories to predict reading ability and controlling for age and phonological sensitivity, the word shape component predicted 2% of unique variance for the 4-year-olds, the word element and spelling components uniquely predicted 3 and 6% of the variance, respectively, for the 5-year-olds, and the word elements and spelling components added 5 and 19% of the variance, respectively, for the 6-year-olds. Thus a primary orthographic knowledge exists prior to formal reading instruction that can be measured by these preliminary concepts of print and significantly predict reading ability.

Furthermore, the predictive power of OP seems to increase with age. For instance, Badian (2001) measured OP in preschoolers by having them indicate which one of four stimuli matched a target stimulus. After controlling for prior reading ability, verbal IQ, verbal memory, and age, this OP measure developed from being a nonsignificant predictor of reading comprehension in Grade 1, to significantly predicting 4% of the variance in Grade 3, and finally accounting significantly for 6% of the variance in Grade 7. Conversely, after controlling for the same variables, kindergarten syllable segmentation (a PA task) significantly predicted 4% of the variance in reading comprehension in Grade 1 but became a nonsignificant predictor of reading comprehension thereafter. Evidently, after the acquisition of reading skills, the predictive power of PA decreased with age while the predictive power of OP increased with age. The findings of this study are consistent with Share's (1995) self-teaching hypothesis, in which beginner readers rely mainly on phonological skills to decode words but become increasingly dependent on orthographic knowledge as they become skilled readers.

Overall, the literature supports the validity of OP as an independent construct and its independent predictive power in reading. In addition to phonological skills and print exposure, and as previously discussed, OP may also be related to NS. Wolf and Bowers (1999) suggested that NS potentially influences the rate at which readers can acquire orthographic representations and the quality of the acquired representations. However, research has yet to examine the proposed relationships between OP and NS and the predictive ability of OP after controlling for the effects of NS.

Quick Spell Test

In 1996, Bowers devised the Quick Spell Test (QST) to evaluate whether NS and OP are related to each other as proposed by Wolf and Bowers (1999). In the QST, readers see brief exposures (approximately 250 ms) of four-letter strings. These strings may be real words (e.g., "like"), pseudowords (e.g., "hool"), or consonant clusters (e.g., dlhw), and are randomly presented to the reader. The task measures the reader's ability to correctly report each string immediately after its presentation. Rueffer (2000) later added a second set of consonant clusters to create two separate categories: high frequency (HF) bigrams and low frequency (LF) bigrams. In the HF bigram category, the first two letters (i.e., the first bigram) occur frequently together in the English language and the last two letters (i.e., the second bigram) also occur frequently together. The first and second bigrams may or may not occur together often. For example, the HF bigram stimulus "clnd" consists of "cl" (often seen in words such as "cloud") and "nd" (often seen in words such as "land"). On the other hand, in the LF bigram category, the pairs of letters in each bigram do not occur frequently together (e.g., "mv" and "hw" in the stimulus "mvhw"). The LF bigram category contains the same stimuli as the original consonant cluster category (see Appendix A for a list of stimuli).

The four categories of strings are also meant to reflect different levels of orthographic structure, depending on the extent to which the stimuli in the category cohere to the English written language. Thus real words have the most orthographic structure, as they are part of the written lexicon. Pseudowords have the second highest orthographic structure, as the four letters can be read as a whole unit but are not part of the lexicon. The bigram categories have less orthographic structure than pseudowords, as the four letters are separated into two units (i.e., the two bigrams). The HF bigrams also have higher orthographic structure than the LF bigrams, as the pairs of letters in each HF bigram are more likely to be found in the written language than the pairs of letters in each LF bigram.

According to Bowers (2001), readers with good OP abilities should be more sensitive to the orthographic structure differences between the categories. More specifically, they should be better at identifying real words and pseudowords than the consonant clusters. On the other hand, readers with poor OP should perform poorly on all types of letter strings, and should not benefit from the higher orthographic structure of real words and pseudowords. In other words, the difference between types of letter strings should be greater between readers with good OP than readers with poor OP. Moreover, if NS affects reading through OP, then readers with low NS should show also reduced effects of orthographic structure (i.e., reduced difference between the strings). Specifically, Bowers (1996) hypothesized that NS would affect the ability of readers to process stimuli fast enough to benefit from the orthographic structure of pseudowords and real words.

Thus far, the studies using the QST have not found the purported effects of orthographic structure on readers with NS deficits. For instance, Bowers, Sunseth, and Golden (1999) administered the original version of the QST (i.e., without HF bigrams) to a sample of children in Grades 2 ($N = 31$) and 3 ($N = 32$). On the basis of performance on PA and NS measures, the researchers categorized the children into four groups: no deficit (normal performance on both PA and NS measures), single NS deficit (normal PA performance but poor NS performance), single PA deficit (poor PA performance but normal NS performance), and double deficit (poor PA and NS performance). For the Grade 2 children, the groups with PA and NS deficits (i.e., all groups except the no deficit group) performed worse on all QST string types. The double deficit group also performed worse than the other three groups on pseudowords. In Grade 3, those with poor PA performed worse on real words than those with normal PA, the groups with poor NS performed worse on pseudowords and LF bigrams than those with normal NS, and the double deficit group performed worse on pseudowords than the single NS deficit group. Replicating the study with 15 single PA deficit and 17 single NS deficit children, selected from a sample of 122 Grade 3 children, the researchers found that the two groups did not differ on real word or pseudoword performance, but that the NS deficit group performed worse on the LF bigrams. Therefore, the only consistent finding in these studies was that children with NS deficits performed poorly on the LF bigram strings.

One limitation with the Bowers et al. (1999) study, however, lies in the criteria used to form the four groups. First, the no deficit groups in both Grades 2 ($N = 15$) and 3 ($N = 15$) were larger than the deficit groups (*N*s = 4 to 7 in Grade 2; *N*s = 5 to 6 in Grade 3). Thus the disproportionate cell sizes may have affected the subsequent group comparisons. Second, their double deficit groups had lower and more variable NS scores than the single NS deficit groups. According to Schatschneider, Carlson, Francis, Foorman, and Fletcher (2002), using univariate criteria on PA and NS tasks (i.e., percentile cutoffs) to create the four groups is problematic because of the positive correlation between PA and NS. Schatschneider et al. demonstrated that, if PA and NS were uncorrelated, the single NS deficit and double deficit groups would both have equally low levels of NS and the single PA deficit and no deficit groups would both have equally adequate levels of NS. In this case, differences between the groups in each pair could be attributed to the difference in, and hence effect of, PA. Similarly, the single PA deficit and double deficit groups would both have equally low levels of PA and the single NS deficit and no deficit groups would both have equally adequate levels of PA; and the differences in each pair could be attributed to the effect of NS. However, when PA and NS are positively correlated, the double deficit group may have lower PA scores than those in the single PA deficit group and lower NS scores in the single NS deficit group.

This inequality between the group PA and NS means would obscure the interpretation of group differences. For example, the finding in Bowers et al.'s (1999) study that the double deficit group performed worse on pseudowords than the single NS deficit group is difficult to interpret because the double deficit group had lower PA *and* NS scores than the single NS deficit group.

More recently, Sunseth and Bowers (2002) administered the QST to a sample of Grade 3 children. From a larger sample of 201 Grade 3 children, the researchers selected 68 children for the four categories: 17 in the no deficit group, 17 in the single PA deficit group, 18 in the single NS deficit group, and 16 in the double deficit group. Addressing the previously noted concern with using percentile cutoffs to create the four groups, Sunseth and Bowers ensured that the groups had similar sizes and that the double deficit groups were comparable to the single PA deficit group in PA scores and comparable to the single NS deficit group in NS scores. Administering the same QST as in Bowers et al.'s (1999) study, their findings indicated that the PA deficit groups performed worse on real words than the groups with normal PA, the NS deficit groups performed worse on pseudowords and LF bigrams than groups with normal NS, and the double deficit group performed worse on LF bigrams than the other groups. However, the researchers observed that the group means for the single PA deficit and single NS deficit groups did not differ greatly from each other for all three string types. Although they did not conduct significance testing on the difference between the two groups, this observation was inconsistent with Bowers' (1996) hypothesis that children with NS deficits would perform poorly on the QST due to the link between NS and OP.

In conclusion, these QST studies did not confirm Bowers' (1996) original hypothesis that having a NS deficit would lead to poor performance on the QST due to an inability to benefit from orthographic structure. In all the studies, having any deficit (i.e., whether a NS or PA deficit, or both) led to poorer performance on the QST, and the double deficit groups performed worse than the single deficit groups only on the LF bigram strings, although the interpretation of this finding may be obscured by the method used to create the four groups in Bowers et al.'s (1999) study. Moreover, all groups were able to benefit from orthographic structure, thereby performing better on the real words and pseudowords than on the LF bigram strings. In light of these findings, Bowers (2001) refined the previous hypothesis concerning the link between NS and OP. The refined hypothesis suggests that NS affects the processing of stimuli with low orthographic structure (i.e., LF bigram strings), which represent a baseline upon which further orthographic knowledge can be built. Those with NS deficits have a lower baseline and accordingly require more orthographic knowledge to attain the same level of OP skills as those with adequate NS. Thus, the original hypothesis suggests NS affects the process through which readers acquire OP skills, whereas the refined hypothesis posits NS as affecting the foundation on which readers build OP skills.

Examining the refined hypothesis, Baker (2002; Bowers & Baker, 2002) administered the revised QST (i.e., with both HF and LF bigrams) to a Grade 2 sample, in which 20 children had no deficits, 19 had a single NS deficit, 21 had a single PA deficit, and 21 had both NS and PA deficits. The single NS deficit group performed worse than the no deficit group on pseudoword and both types of bigram strings, and performed worse than the single PA deficit group on both bigram strings. The double deficit group also did not differ from the single NS deficit group on the two bigram strings, suggesting that the NS component of the double deficit was responsible for their performance on the bigram strings. These findings supported the refined hypothesis that NS mainly affected performance on the strings with low orthographic structure. In another study, Ishaik (2003) administered the HF and LF bigrams to a sample of 90 Grade 2 children, of which 30 were good readers, 30 were poor readers with adequate NS, and 30 were poor readers with slow NS. Also consistent with the refined hypothesis, the slow NS group were less accurate on both bigram strings than the fast NS group, and both poor reader groups were less accurate than the good reader group. All groups were better at reporting the HF bigrams than the LF bigrams, indicating that they benefited equally from the higher orthographic structure. However, the poor readers did not benefit from increased exposure in a Letter Probe task, in which the children were shown four-letter consonant strings and asked to decide whether a subsequent probe was in the previous string. This finding was inconsistent with Bowers' (1996) hypothesis that readers with slow NS do not have sufficient time to process QST stimuli. Therefore, although the QST studies have concurred that NS mainly affects performance on the LF bigrams, the literature has yet to provide a sustainable explanation for the details of this influence.

Current Study

Since the development of the QST, the literature has focused on the relationship between QST and poor readers and has not examined the relationship between QST and normal reading ability. As well, the QST studies have mainly conducted group comparisons to show the respective influences of PA and NS. This methodology may be problematic for several reasons. First, as mentioned before, using percentile cutoffs on PA and NS tasks is problematic for group comparisons because the double deficit group may have lower PA or NS scores than the respective single deficit group. Second, none of the QST studies controlled for other variables that may affect reading ability, such as age or IQ. Therefore, even though the later QST studies attempted to resolve the problem of the double deficit group having lower PA or NS scores, the group differences in these studies nevertheless could have been due to extraneous factors. Third, studies were inconsistent with the criteria used to specify PA and NS deficits, thereby hindering comparisons across studies. Some studies with liberal criteria may have examined groups in which the lower scores of the “normal” group were similar to the higher scores of the “deficit” group. These analyses cannot be easily compared with other studies which used stringent criteria to create groups with greater skill differences. Fourth, comparing group means can reduce the predictive power of continuous measures; thus the lack of predictive ability in these studies may be due to inadequate power.

In addition to these methodological issues, few, if any, studies have examined the influence of predictors other than PA and NS on QST performance. As previously discussed, many researchers have shown processes such as PD and OP to also account for unique variance in reading. Moreover, although the QST is an orthographic task, few studies have examined whether OP accounts for more variance in the QST than other reading processes. This analysis is important for confirming the validity of the QST as an OP skill so that future studies can use the QST as an additional OP task. Furthermore, researchers have also rarely investigated the relationship between QST, other reading measures, and the four reading predictors. For example, the relationship between NS and QST found in the studies could have been merely due to a common speed requirement in both tasks; thus, it is essential to ensure that the relationship between the two measures is related to reading ability and not other cognitive abilities.

Therefore, the current archival study examined the relationship between QST and reading in a normal sample of Grade 3 and 4 children. I used regression analyses instead of group comparisons to preserve maximum power in the continuous measures and to avoid the aforementioned problems with grouping. Moreover, I controlled for the effects of age and IQ in all analyses, in order to eliminate variance due to these background factors. My first research question evaluated whether readers with good OP skills would show a greater difference between the types of QST stimuli than readers with poor OP skills. Bowers (1996; 2001) proposed that readers with good OP skills should benefit more from higher orthographic structure and therefore show a greater discrepancy between the different types of strings than readers with poor OP skills. If NS is related to OP, then readers with good NS should also show a greater difference between the different string types than readers with poor NS. The results of these analyses, together with the findings from the analyses addressing the second research question, would provide more information as to whether the QST predominantly measures OP or not. As well, if the analyses with NS produce findings similar to the OP analyses, then these results would provide support to the hypothesized link between NS and OP.

The second research question pertained to the relationship between QST and reading ability. To answer this question, I examined the amount of variance in various established reading measures attributable to performance on the QST, both concurrently and longitudinally. This is an essential first step in determining the predictive value of the QST: if the QST does not relate well to the reading measures, then studies examining the relationship between the QST and other reading predictors would not be greatly applicable to the study of reading.

My third research question asked whether the QST measures the same processes as it purports. To address this issue, I determined which of the four reading predictors (i.e., PA, NS, PD, and OP) best accounted for the shared variance between QST and various established reading measures. Although previous studies used the QST as a sole outcome variable, I used the shared variance between QST and various reading measures as the outcome variables. This procedure ensured that the examined variance in the QST would be related to reading instead of to other cognitive processes. As QST is posited as a measure of OP, I expected that OP would be the strongest predictor, relative to the other three predictors.

Finally, the fourth research question was related to the four categories of letter strings and their relative degree of orthographic structure. Previous studies have repeatedly stated that readers recognize real words and pseudowords more easily than the bigrams because of the higher orthographic structure. Therefore I studied whether the four types of letter strings actually differed in orthographic structure, or whether they differed in more ways than simply orthographic structure. To answer this question, I determined which of the four reading predictors best accounted for the difference between the QST categories. I also expected OP to be the strongest predictor of this difference, consistent with the design of the QST.

CHAPTER 2: METHOD

Participants

This study examined the Grade 3 and 4 data from Kirby, Parrila, Deacon, and Wade-Woolley's (2004-2007) longitudinal study, which followed a cohort of children from Grades 3 to 5. In Grade 3, the researchers recruited 92 children from elementary schools in Kingston, Ontario, and 100 children from elementary schools in St. Albert, Alberta. All participants received letters of information and signed consent forms. The only requirement for inclusion in the study was that the child would be able to understand the instructions for the various tasks. The Grade 3 sample included 99 females and 93 males, with ages ranging from 87 months to 120 months (M = 105.15 months, SD = 4.39 months). In Grade 4, 149 of the original sample remained, of which 78 were female and 71 were male. The distribution of ages in the Grade 4 sample did not differ from that of the Grade 3 sample.

Measures

Kirby et al. (2004-2007) administered a battery of various measures to the sample each year. This study included 14 of the measures administered in Grade 3: two IQ measures, two PA measures, one PD measure, two NS measures, three OP measures, the Quick Spell task, and three other reading measures. The examiner administered seven of the measures on a computer using DirectRT (Jarvis, 2000) software. This study also included four measures subsequently administered in Grade 4: the Quick Spell task and the same three reading measures. For the published tests, the reported reliabilities are taken from the respective test manuals.

Intelligence Measures

Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999): Vocabulary Subtest

In this measure of IQ, the examiner presented up to 26 words one-by-one to the child, with each item printed on a separate page of a booklet. The child was to define each word orally, and the examiner assigned scores of 0, 1, or 2 to the responses according to the criteria in the WASI manual. The examiner discontinued the task after five consecutive scores of 0. The reported internal consistency reliability coefficient is .86 for children at age eight (Wechsler, 1999).

WASI (Wechsler, 1999): Block Design Subtest

In this measure of IQ, the examiner presented up to 13 two-dimensional geometric patterns one-by-one to the child, with each pattern printed on a separate page of a booklet. The examiner also provided the child with the appropriate number of two-colour blocks required to produce each pattern. The child had two trials to replicate each design within a designated time limit. The examiner assigned scores of 0, 1, or 2 to the child's replications according to the WASI manual and discontinued the task after three consecutive scores of 0. The reported internal consistency reliability coefficient is .87 for children at age eight (Wechsler, 1999).

Phonological Awareness Measures

Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999): Elision Task

Although the original version of this task uses audio cassettes to administer the stimuli, the current study adapted the task for use with a computer. In this task, the computer presented up to 20 items one-by-one to the child. For each item, the computer program generated a spoken word to which the child listened through headphones (e.g., “cup”). The program first asked the child to repeat the given word into a microphone. Next, the program asked the child to repeat the word again without a specified syllable or phoneme (e.g., “cup” without the “/k/”). The examiner discontinued the task after three consecutive incorrect answers, and the score was the total number of correct responses. The Spearman-Brown split-half coefficient was .90 for the computer version of the task in this study.

CTOPP (Wagner et al., 1999): Word Blending Task

The examiner presented up to 20 items one-by-one to the child. For each item, the examiner pronounced a series of separate sounds and asked the child to combine them into a whole word (e.g., “sh-ē” into “she”). The examiner discontinued the task after three consecutive incorrect answers and the score was the total number of correct responses. The reported Cronbach’s alpha for this task is .79 for children at age eight (Wagner et al., 1999).

Phonological Decoding Measure

Woodcock Reading Mastery Tests – Revised (WRMT-R; Woodcock, 1998; Form H): Word Attack Task

The examiner presented up to 45 nonwords or very-low-frequency words (e.g., “rayed” pronounced “r-ā-d”) in a booklet for the child to pronounce. The examiner discontinued the test after six consecutive errors in pronunciation, and the score was the number of correct pronunciations for the child. The reported Spearman-Brown-corrected split-half reliability coefficient for this subtest is .91 for children in Grade 3 (Woodcock, 1998).

Naming Speed Measures

Reliability was not calculated for the NS measures, as each measure only included one set of stimuli and the speed for each individual item could not be assessed. The closest approximation to a reliability estimate was the correlation between the digit naming task and the letter naming task: $r = .79$.

Rapid Digit Naming Task

Kirby et al. (2004-2007) adapted this measure from the CTOPP (Wagner et al., 1999) Rapid Digit Naming Task. A computer screen presented 50 items simultaneously, with five rows and ten columns of five randomly arranged digits (i.e., 2, 7, 4, 9, and 5). The examiner asked the child to sequentially read all the digits on the screen, with instructions to be as fast as possible without making mistakes. The score was the number of milliseconds the child took to read all the digits on the page in order, as measured by the computer. The number of errors made on this task was low ($M = 0.24$, $SD = 0.79$) and subsequently ignored.

Rapid Letter Naming Task

Kirby et al. (2004-2007) also adapted this task from the CTOPP (Wagner et al., 1999) Rapid Letter Naming Task. Similar to the digit naming measure, a computer screen presented 50 items simultaneously, with five rows and ten columns of five randomly arranged letters (i.e., o, s, a, d, and p). The examiner asked the child to sequentially read all the letters on the screen, with instructions to be as fast as possible without making mistakes. The score was the number of milliseconds the child took to read all the letters on the page in order, as measured by the computer. The number of errors made on this task was also low ($M = 0.25$, $SD = 0.68$) and subsequently ignored.

Orthographic Processing Measures

Orthographic Choice Task

A computer program presented 30 pairs of irregular words and their pseudohomophones (e.g., answer – anser), showing one pair at a time on the computer screen. The examiner indicated two keys on the computer keyboard which corresponded to the two responses on the screen. The child indicated the correctly-spelled word in the pair by pressing the appropriate key, and the score was the total number of correct responses. Kirby et al. (2004-2007) selected the word pairs from the Orthographic Coding Task in Olson, Forsberg, Wise, and Rack's (1994) study. Appendix B includes a full list of the stimuli used in Kirby et al.'s study, and the Cronbach's alpha for the task in this study was .82.

Wordlikeness Task

In this measure of OP, a computer program presented 24 pairs of non-words (e.g., filv-filk), showing one pair at a time on the computer screen. The examiner indicated two keys on the computer keyboard which corresponded to the two responses on the screen. The child indicated the nonword in the pair which more closely resembled a real word, and the score was the total number of correct responses. Kirby et al. (2004-2007) selected the non-word pairs from the three Letter String Choice tasks in Cunningham et al.'s (2001) study. Appendix C includes a full list of the stimuli used in Kirby et al.'s study, and the Cronbach's alpha for the test in this study was .67.

Word Chains Task

This measure of OP presented 11 lines of print to the child on an 8.5-inch by 11-inch white piece of paper. Each line consisted of multiple words presented without the spaces between the words. The examiner asked the child to place vertical slashes indicating the location of the missing spaces (e.g., "toydogcar" should be "toy|dog|car"). The task included a total of 51 words that should be separated by slashes. The score was the number of correct slashes drawn within a one-minute time limit. Kirby et al. (2004-2007) created the stimuli in this study (see Appendix D for an example of the exact stimuli). The Spearman-Brown-corrected split-half coefficient for unequal length between odd and even lines was .85 in this study.

Quick Spell Test

In the QST (Bowers, 1996), the computer randomly presented 40 four-letter strings in lowercase one-by-one to the child for 300 ms. The computer also showed a mask (i.e., "xxxx") before the first target string and immediately after each target string, remaining on the screen until the child was ready for the next target string. Kirby et al. (2004-2007) added this mask to the QST in order to erase any visual afterimage from the target string exposure. As previously described, the 40 target strings (see Appendix A for a list of stimuli) were real words, pseudowords, high frequency (HF) bigrams, or low frequency (LF) bigrams, with 10 strings in each category. The examiner instructed the child that he/she would see four letters quickly on the screen and that the child was to write down what he/she saw after each four-letter string. Although the previous QST studies asked the participants to verbally report the target string, Kirby et al. used written responses in order to eliminate sources of error due to articulation and to produce written records that could be easily stored and managed for future analyses. The score was the total number of correct strings for the child. The QST was administered in both Grades 3 and 4. For the Grade 3 scores, the Cronbach's alpha in this study was .91 for the total score, .72 for real words, .80 for pseudowords, .84 for HF bigrams, and .76 for LF bigrams. In Grade 4, the Cronbach's alpha was .92 for the total score, .77 for real words, .78 for pseudowords, .87 for HF bigrams, and .78 for LF bigrams.

Reading Measures

WRMT-R (Woodcock, 1998; Form H): Word Identification Task

In this measure of reading ability, the examiner presented a booklet containing 106 words listed in order of decreasing frequency (i.e., from "go" as item 1 to "shillelagh" as item 106) to the child. Each page contained one to nine words and the child identified each item. The examiner discontinued the test after six consecutive errors, and the score for the child was the total number of correct responses. The reported split-half reliability coefficient is .97 for children in Grade 3, and this measure was administered in both Grades 3 and 4 (Woodcock, 1998).

WRMT-R (Woodcock, 1998; Form H): Passage Comprehension Task

In this reading measure, the examiner presented up to 68 passages to the child in a booklet, with three or four passages per page. Each passage consisted of one to three sentences and contained a blank line to represent a missing word. The child silently read each passage and identified the appropriate missing word. The examiner discontinued the test after six consecutive errors, and the score for the child was the total number of correct responses. Although this measure was administered in both Grades 3 and 4, the manual reported the split-half reliability coefficient for only Grade 3 (.92; Woodcock, 1998).

Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999; Form A): Sight Word Efficiency Task

In this measure of reading ability, the examiner presented 104 words, listed in order of increasing difficulty (i.e., from "is" as item 1 to "transient" as item 104), on a laminated 8.5-inch by 11-inch white piece of paper. The child identified as many words in order as possible within a 45-second time limit and the score for the child was the total number of correct responses. The reported test-retest reliability for this measure is .97 for children aged 6 to 9. This measure was administered in both Grades 3 and 4 (Torgesen et al., 1999).

Procedure

Kirby et al. (2004-2007) distributed letters of information (Appendix E) and consent forms (Appendix F) to the parents of the participating children in Grade 3, and similar letters and consent forms were re-distributed in Grade 4. The examiner administered the measures in a quiet room at the children's schools. The examiner conducted two testing sessions for each child: a session administering paper-and-pencil tasks and a session administering computer tasks. Each session lasted approximately 45 minutes. The researchers collected the Grade 3 data between February and May of 2005 and the Grade 4 data between February and May of 2006.

CHAPTER 3: RESULTS

Descriptives

Table 2 and Table 3 show the mean, standard deviation, skewness, and kurtosis statistics for all the measures in this study. Although the sample size was large enough for analyses to be robust to violations of normality, I nevertheless performed transformations on all measures which violated normality because of the large number of non-normal distributions and the different directions of skewness found in different measures. As per Tabachnick and Fidell's (2001) guidelines, I used a criterion of .01 on the z distribution to indicate a violation of normality, meaning that I considered a measure non-normal if the skewness and kurtosis value exceeded 3.09 when divided by the appropriate standard error of the statistic. Therefore, I transformed the Grade 3 scores for the Block Design, Rapid Digit Naming, Rapid Letter Naming, Word Attack, Orthographic Choice, Wordlikeness, Wordchains, Word Identification, Passage Comprehension, TOWRE, and Quick Spell Test (QST) measures; and the Grade 4 scores for the Word Identification and QST measures. I windsorized the Grade 4 Passage Comprehension and TOWRE scores instead of applying transformations because the non-normality was caused by several outliers rather than by the distribution shape. Appendix G provides the descriptive statistics for the transformed measures and indicates the specific transformation applied to each measure. The inverse transformation of the Rapid Digit Naming and Rapid Letter Naming tasks changed the scores from indicating naming time to indicating naming speed. Thus, after transformation, higher scores on these tasks indicated faster naming. I conducted the remaining analyses with the transformed scores. Appendix H shows the correlations between all the measures used in this study.

QST performance in both the total score and the four subscale scores improved as the children progressed from Grade 3 to Grade 4 (Table 2). Performance in the four subscales also conformed to Bowers' (1996) design in that the accuracy rates were higher for the stimuli with higher orthographic structure (i.e., real words > pseudowords > high frequency [HF] bigrams > low frequency [LF] bigrams). However, the distributions for many of the QST variables violated normality and thus required transformations. After the appropriate transformations, all variables approximated normal distributions except the real words in both Grades 3 and 4 (Appendix G). The severe abnormality in the real word subscale was mainly due to a ceiling effect: Grade 3 children accurately reported almost all real word stimuli (M = 9.36, *SD* = 1.16), thus leaving little room for improvement in Grade 4 (*M* = 9.64, *SD* = 1.04). Conversely, the LF bigram subscale showed a floor effect, although less severe than the real word subscale's ceiling effect. Grade 3 children reported an average of 1.57 LF bigrams, and the standard deviation of 1.99 indicates that a portion of the participants did not report any LF bigrams accurately. However, unlike the real words, the LF bigram variables approximated normal distributions after applying transformations (Appendix G).

Table 2

Raw Score Descriptive Statistics of Quick Spell Measures

Measure	**N**	**Mean**	**SD**	**Skewness**	**SE**[a]	**Kurtosis**	**SE**[b]
Grade 3	186						
Total		22.35	7.11	-0.34	.18	-0.42	.36
Real Words		9.36	1.16	-2.85	.18	10.18	.36
Pseudowords		7.33	2.54	-1.09	.18	0.62	.36
HF Bigrams		4.11	2.98	0.05	.18	-1.28	.36
LF Bigrams		1.57	1.99	1.47	.18	1.65	.36
Grade 4	149						
Total		26.55	7.17	-0.81	.20	0.45	.40
Real Words		9.64	1.05	-5.30	.20	35.76	.40
Pseudowords		8.49	2.05	-1.86	.20	3.31	.40
HF Bigrams		5.78	3.11	-0.65	.20	-0.90	.40
LF Bigrams		2.64	2.42	0.80	.20	-0.13	.40

Note: The maximum score is 40 for the QST total and 10 for each QST subscale.
[a] Standard error of skewness statistic.
[b] Standard error of kurtosis statistic.

Table 3

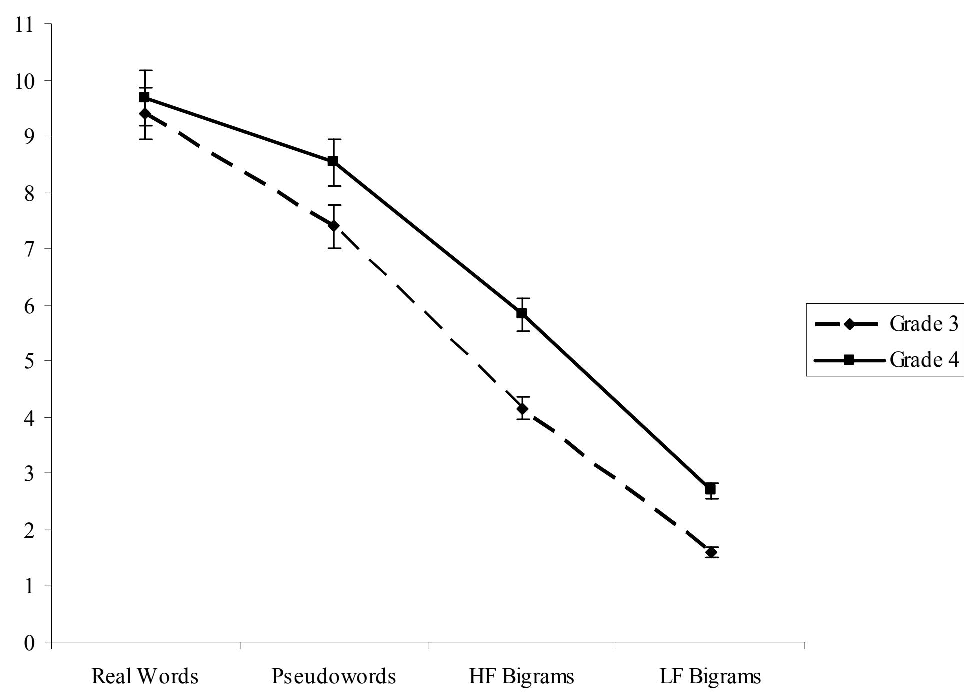
Raw Score Descriptive Statistics of Other Reading Measures

Measure	N	Mean	SD	Skewness	SE[a]	Kurtosis	SE[b]
Grade 3							
Vocabulary	191	33.60	7.53	-0.22	.18	-0.11	.35
Block Design	191	19.76	12.05	0.99	.18	0.24	.35
Elision	190	20.26	5.39	-0.35	.18	-0.50	.35
Word Blending	191	14.58	3.38	-0.54	.18	-0.26	.35
Digit Naming	186	31.21	8.58	2.70	.18	12.64	.36
Letter Naming	186	31.05	8.25	3.01	.18	16.21	.36
Orthographic Choice	190	26.59	5.50	-3.47	.18	15.92	.35
Wordlikeness	189	17.63	5.37	-1.41	.18	2.02	.35
Word Chains	191	24.02	10.63	-1.43	.18	4.27	.35
Word Attack	190	26.34	9.20	-0.79	.18	-0.08	.35
Word Identification	191	63.41	11.62	-0.83	.18	1.12	.35
Passage Comprehension	191	33.80	7.22	-1.28	.18	3.77	.35
TOWRE	191	59.30	13.83	-1.24	.18	1.90	.35
Grade 4							
Word Identification	149	70.15	11.92	-1.09	.20	2.19	.40
Passage Comprehension	148	37.55	7.27	-0.68	.20	2.40	.40
TOWRE	149	66.78	11.43	-0.93	.20	2.47	.40

[a] Standard error of skewness statistic.
[b] Standard error of kurtosis statistic.

Furthermore, I conducted a 2 (Grade) x 4 (QST Subscale) repeated-measures ANOVA to examine whether performance across the two grades improved differently for each subscale. For this analysis, I used the untransformed QST variables instead of the transformed scores to keep the variables on the same scale (i.e., out of 10). The ANOVA showed a significant main effect of both grade, Wilks' $\Lambda = .44$, $F(1, 144) = 180.61$, $p < .01$; and subscale, Wilks' $\Lambda = .06$, $F(3, 432) = 751.87$, $p < .01$. The interaction term was also significant, Wilks' $\Lambda = .68$, $F(3, 432) = 22.44$, $p < .01$. This significant interaction appeared to be due to the ceiling effect on the real word subscales (Figure 1), as the difference in accuracy between the two grades was similar for the other subscales. Therefore, children appeared to improve similarly on every QST subscale as they matured, and the relative accuracy of the four subscales (i.e., real words > pseudowords > HF bigrams > LF bigrams) did not seem to change as the children developed.

Figure 1. *Quick Spell Test Raw Scores: Interaction between Grade and Subscale*



Note: Errors bars indicate the 95% confidence interval around the mean. The maximum score for any subscale is 10.

Factor Analysis

Prior to examining the relationship between the QST and phonological awareness (PA), naming speed (NS), and orthographic processing (OP), I determined whether the measures of these three reading predictors would form distinct factors as predicted by the literature. I entered the Grade 3 scores for the Elision (PA), Word Blending (PA), Rapid Digit Naming (NS), Rapid Letter Naming (NS), Word Chains (OP), Orthographic Choice (OP), and Wordlikeness (OP) tasks into a principal-axis factor analysis. Appendix I shows the correlations between these measures. On the basis of theory, the eigenvalues-greater-than-1.0 guideline, and the scree plot, I extracted three factors and rotated the factors using the direct oblimin criterion. I used an oblique rotation as the literature has shown the three constructs to be correlated with each other (e.g., Swanson et al., 2003; Wolf & Bowers, 1999). The eigenvalues for the three rotated factors were 2.03, 2.06, and 1.04, and the factors together accounted for 57.13% of the variance in the measures. Table 4 shows the rotated factor loadings from the pattern matrix and the correlations between the three extracted factors. As expected, the Word Chains, Orthographic Choice, and Wordlikeness tasks formed an OP factor, the Digit and Letter Naming measures formed a NS factor, and the Elision and Word Blending tasks formed a third PA factor. Only the Elision task had loadings greater than .30 on more than one factor (i.e., .59 on the PA factor and .36 on the OP factor); however its PA factor loading was clearly stronger than its OP factor loading.

Z-Score Composites

On the basis of the factor analysis results, I created three *z*-score composites for the six Grade 3 PA, NS, and OP measures. More specifically, I transformed the Elision and Word Blending scores into *z*-scores separately, and subsequently averaged the two *z*-scores to create a PA composite. Similarly, I created a NS composite from the Digit and Letter Naming tasks and an OP composite from the Orthographic Choice, Wordlikeness, and Word Chains measures. Table 5 shows the descriptive statistics for the three computed composites. I used these three composite variables in lieu of the individual measures in all remaining analyses.

Table 4

Factor Analysis of Grade 3 Phonological Awareness, Naming Speed, and Orthographic Processing Measures (Pattern Matrix; N=186)

Measure	**Orthographic Processing**	**Naming Speed**	**Phonological Awareness**
Elision	.36	-.09	**.59**
Word Blending	.12	.09	**.55**
Digit Naming	-.02	**.91**	.04
Letter Naming	.14	**.81**	.02
Orthographic Choice	**.78**	-.04	-.05
Wordlikeness	**.62**	.20	-.08
Word Chains	**.58**	.10	.18
Correlations			
Orthographic Processing.	-		
Naming Speed	.40	-	
Phonological Awareness.	.25	.28	-

Note: Bolded font indicates the factor onto which the measure loaded.

Table 5

Descriptive Statistics of Grade 3 Z-Score Composites

Measure	**N**	**Mean**	**SD**	**Skewness**	**SE[a]**	**Kurtosis**	**SE[b]**
Phonological Awareness	191	0.00	0.82	-0.46	.18	0.15	.35
Naming Speed	186	0.00	0.95	-0.13	.18	0.75	.36
Orthographic Processing	191	-0.00	0.82	-0.86	.18	0.66	.35

[a] Standard error of skewness statistic.
[b] Standard error of kurtosis statistic.

Regression Analyses

To address the first research question of whether readers with good OP skills benefited more from orthographic structure than readers with poor OP skills, I conducted several regression analyses. First, I used the Grade 3 OP composite to predict each of the Grade 3 QST subscales. According to Bowers' (1996; 2001) claim, the difference between the subscales should be greater at the higher levels of OP than at the lower levels. In other words, the regression lines should follow a fan-shaped pattern. Figure 2 shows the regression lines of the Grade 3 OP analyses. The observed patterns partially supported Bowers' (1996) claim, as the slope was greater for pseudowords ($B = 2.14$, $p < .01$) and HF bigrams ($B = 2.34$, $p < .01$) than for LF bigrams ($B = 0.93$, $p < .01$). Therefore, readers with good OP skills benefited more from the higher orthographic structure of the pseudowords and HF bigrams than the readers with low OP skills. However, the slopes of the pseudowords and HF bigrams were similar; thus the higher orthographic structure of the pseudowords benefited all readers similarly. Moreover, the slope was smaller for the real words ($B = 0.72$, $p < .01$) than for the other three subscales, indicating that readers with high OP skills benefited less from orthographic structure than readers with low OP skills. However, the small slope could have been due to the ceiling effect in the real word subscales; thus the interpretation of the real word slope is unclear.

Similarly, I used the Grade 3 OP composite to predict each of the Grade 4 QST subscales to determine whether the trends changed in Grade 4 (Figure 3). In Grade 4, the slope of the HF bigrams ($B = 2.59$, $p < .01$) was also greater than that of the LF bigrams ($B = 1.31$, $p < .01$). However, the pseudoword slope ($B = 1.57$, $p < .01$) was smaller than the HF bigram slope, and the real word slope ($B = 0.54$, $p < .01$) was smaller than the pseudoword slope. Due to the ceiling effect in the real word subscale, and the potential ceiling effect in the pseudoword subscale, these smaller slopes are also difficult to interpret.

Figure 2. *Regression Lines of OP Predicting QST subscales in Grade 3*

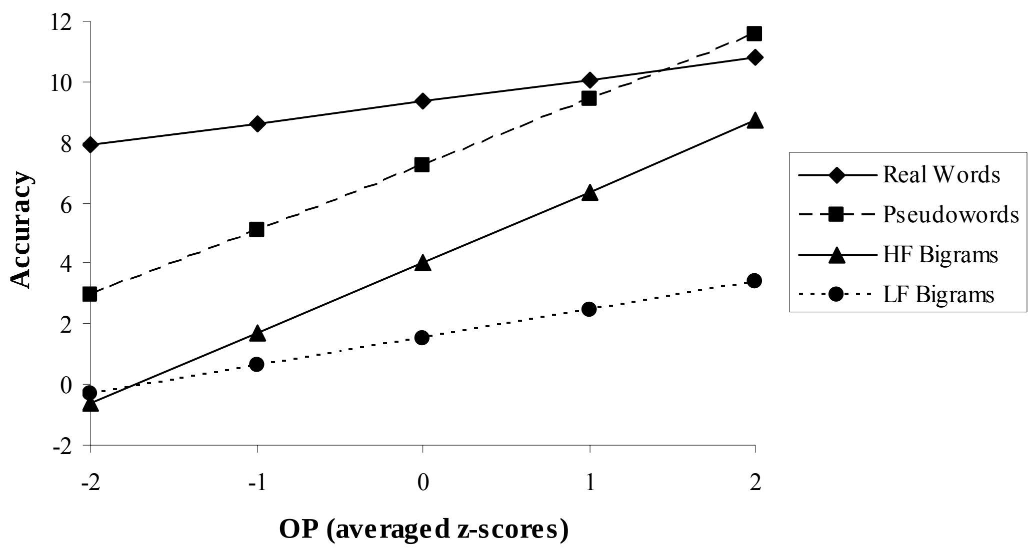
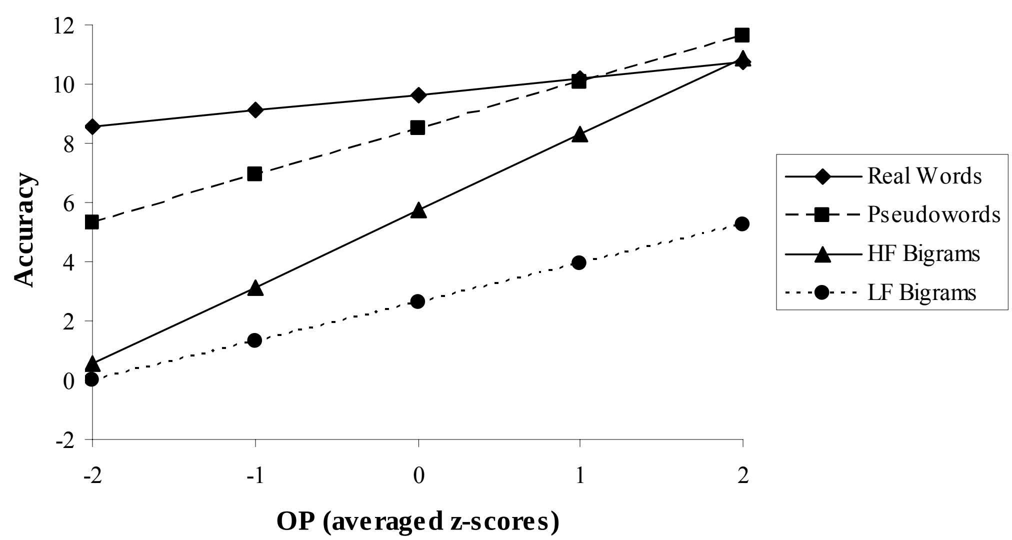


Figure 3. *Regression Lines of Grade 3 OP Predicting Grade 4 QST Subscales*



Next, I conducted similar analyses with Grade 3 NS predicting the QST subscales in both Grades 3 and 4. Figure 4 shows the Grade 3 analyses and Figure 5 shows the Grade 4 analyses. Similar to the OP analyses, the pseudoword ($B = 1.33$, $p < .01$) and HF bigram ($B = 1.55$, $p < .01$) slopes were similar to each other but were greater than the LF bigram slope ($B = 0.78$, $p < .01$) in Grade 3. The real word slope ($B = 0.46$, $p < .01$) was also smaller than the other three subscale slopes, and was likely to have been due to a ceiling effect. In Grade 4, the pseudoword ($B = 1.01$, $p < .01$) and HF bigram slopes ($B = 1.58$, $p < .01$) were also greater than the LF bigram slope ($B = 0.89$, $p < .01$) but similar to each other. The real word slope ($B = 0.41$, $p < .01$) was again smaller than the other subscale slopes. Because the NS analyses were similar to the OP analyses, they provide evidence that NS may be related to OP.

Figure 4. *Regression Lines of NS Predicting QST Subscales in Grade 3*

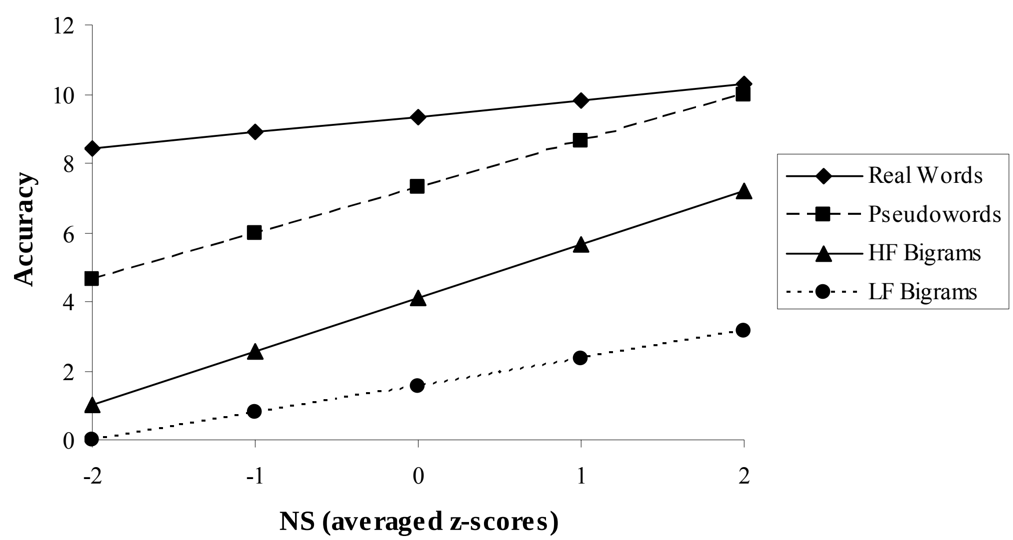
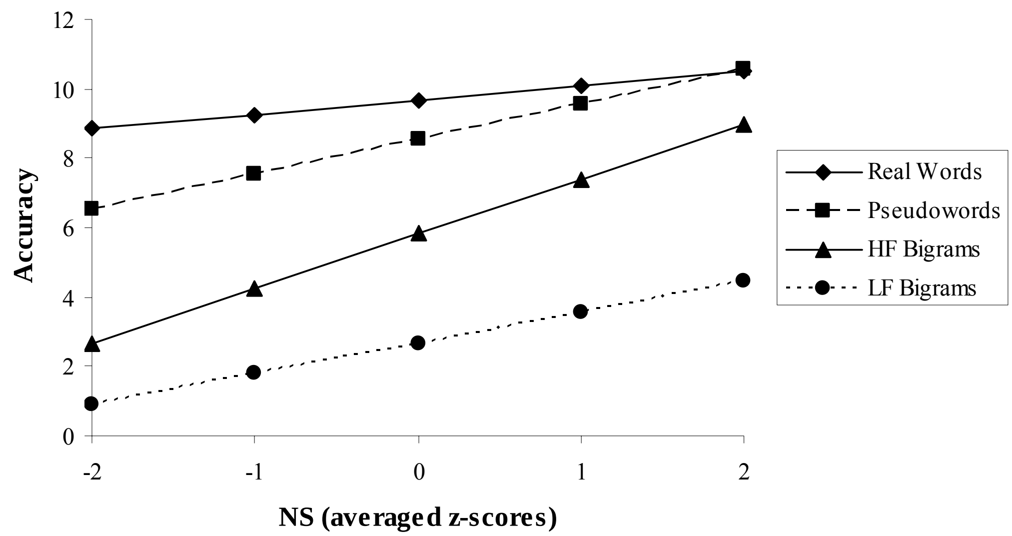


Figure 5. *Regression Lines of Grade 3 NS Predicting Grade 4 QST Subscales*



Hierarchical Regression Analyses

To answer the remaining research questions, I conducted numerous hierarchical regression analyses to examine the relationship between the QST and various reading measures. For each analysis, I entered age and the two IQ measures into the first step to control for the effects of these variables. The sections below outline the analyses examining the relationship between Grade 3 QST scores and Grade 3 reading ability, the relationship between Grade 4 QST scores and Grade 4 reading ability, and the predictive ability of Grade 3 QST scores in Grade 4 reading ability.

Concurrent Analyses

Grade 3

Analyses Involving Quick Spell Test Total Score. To answer the second research question concerning the relationship between the QST and reading ability, I conducted three hierarchical regression analyses with Grade 3 QST predicting the three Grade 3 reading measures. Controlling for IQ and age, the QST total score significantly predicted

29% of the variance in Word Identification, 14% of the variance in Passage Comprehension, and 34% of the variance in TOWRE (analysis a in Table 6). The larger shared variance between the QST and TOWRE was likely due to a common speed component in both measures.

Next, to address the third question of whether the QST was a measure of OP, I examined which of the four reading predictors (i.e., PA, NS, PD, and OP) could best account for the shared variance between QST and reading. In other words, I determined which of the four predictors would weaken the predictive ability of the QST on the three reading measures. Thus I entered each of the four reading predictors into the second step and the QST total score into the third step of the regression analyses. Table 6 provides a summary of these analyses and Appendix J contains the details of each individual analysis. OP and Word Attack (i.e., PD) reduced the QST effect on the reading measures more than PA and NS. After adding PA into the second step, the QST predicted 20% of the variance in Word Identification, 9% of the variance in Passage Comprehension, and 27% of the variance in TOWRE (analysis b in Table 6). Therefore, the addition of PA reduced the QST's predictive ability by 9% in Word Identification, 5% in Passage Comprehension, and 7% in TOWRE. Adding NS into the second step reduced the QST effect by 12% in Word Identification, 7% in Passage Comprehension, and 23% in TOWRE (analysis c). The large amount of shared variance between NS and TOWRE is not surprising, given that both the NS measures and the TOWRE required speed whereas Word Identification and Passage Comprehension did not. On the other hand, controlling for OP reduced the QST effect by 22% in Word Identification, 11% in Passage Comprehension, and 23% in TOWRE (analysis d). Similarly, adding PD reduced the

QST effect by 28% in Word Identification, 13% in Passage Comprehension, and 27% in TOWRE (analysis e).

Because OP and PD reduced the effect of the QST total the most, I examined the predictive ability of these measures in further detail. Table 6 and Appendix J also provide the results of these analyses. The addition of *both* OP and PD into the second step reduced the QST effect by 31% in TOWRE and to nonsignificance in Word Identification and Passage Comprehension (analysis f). I decomposed the unique and shared variances of OP and PD that are related to the shared variance between the QST and the reading measures. In other words, I examined whether the unique variance of either OP or PD or the covariance between OP and PD would reduce the effect of the QST the most. First, I regressed the OP composite and Word Attack variables onto each other and created two residual variables: an OP composite residual consisting of the variance in the OP composite that Word Attack did not explain, and a Word Attack residual consisting of the variance that the OP composite did not explain. These residual variables represented the unique variances of OP and PD.

The OP composite residual did not significantly predict any of the reading measures, and reduced the QST effect by only 1% in Word Identification and Passage Comprehension, and 3% in TOWRE (analysis g). Thus the OP residual played a negligible role in weakening the QST's predictive ability. In contrast, the Word Attack residual reduced the QST effect by 13% in Word Identification, 6% in Passage Comprehension, and 10% in TOWRE. Hence, the Word Attack residual performed very similarly to PA and NS in weakening the QST effect. The ability of the OP-PD covariance to reduce the QST effect can also be deduced from these analyses. For example, the QST predicted 29% of the variance in Word Identification after controlling for the background variables (analysis a), but it did not predict any unique variance in Word Identification after controlling for both OP and PD (analysis f). Therefore OP and PD together accounted for all the shared variance between the QST and Word Identification. The unique OP variance (i.e., the OP composite residual) reduced the QST effect by 1% (analysis g), and the unique PD variance (i.e., the Word Attack residual) reduced the QST effect by 13% (analysis h). Hence, the OP-PD covariance would reduce the QST effect by 15% in Word Identification (i.e., 29% − 1% − 13%).

By similar subtraction methods, the OP-PD covariance would also reduce the QST effect by 7% in Passage Comprehension (i.e., 14% − 1% − 6%). Because the QST predicted 34% of the variance in TOWRE after controlling for the background variables and 3% of the variance in TOWRE after also controlling for both OP and PD, the shared variance between the QST and TOWRE explained by OP and PD together was 31% (i.e., 34% − 3%). Accordingly, the OP-PD covariance would reduce the QST effect by 18% in TOWRE (i.e., 31% − 3% − 10%). Overall, the OP-PD covariance weakened the QST effect to a similar extent as the Word Attack residual. Therefore, both the Word Attack residual and the OP-PD covariance best accounted for the shared variance between QST and reading ability.

Table 6

Summary of ΔR^2 Values in Grade 3 Hierarchical Regression Analyses with the Quick Spell Test Total Predicting Reading Measures

Step/Predictor	**Word Identification**	**Passage Comprehension**	**TOWRE**
1. Background Variables [a]	.21 - .22***	.30 - .31***	.20 - .21***
2a. QST Total	.29***	.14***	.34***
2b. PA Composite	.15***	.11***	.10***
3b. QST Total	.20***	.09***	.27***
2c. NS Composite	.13***	.09***	.35***
3c. QST Total	.17***	.07***	.11***
2d. OP Composite	.28***	.14***	.26***
3d. QST Total	.07***	.03**	.11***
2e. Word Attack	.47***	.22***	.33***
3e. QST Total	.01*	.01	.07***
2f. OP Composite Word Attack	.50***	.23***	.38***
2f. QST Total	.00	.00	.03***
2g. OP Composite Residual	.01	.01	.02*
3g. QST Total	.28***	.13***	.31***
2h. Word Attack Residual	.21***	.09***	.11***
3h. QST Total	.16***	.08***	.24***

Note: $N = 184$ for analyses a, b, & d; $N = 183$ for analyses e, f, g, & h; $N = 182$ for analysis c.
[a] Background variables are age, vocabulary and block design.
* $p < .05$. ** $p < .01$. *** $p < .001$

Analyses Involving Quick Spell Test Subscale Scores. Next, I conducted similar analyses using the four Grade 3 subscale scores: real words, pseudowords, high frequency (HF) bigrams, and low frequency (LF) bigrams. Table 7 shows the predictive ability of the four subscales in the three reading measures, after controlling for age and IQ. The four subscales together added 30% unique variance in Word Identification, 15% in Passage Comprehension, and 33% in TOWRE, similar to the variance added by the QST total score. However, only pseudowords (β = .21 to .25) and HF bigrams (β = .24 to .40) were consistently significant individual predictors of the three reading measures.

Table 7

Summary of Hierarchical Regression Analyses with Quick Spell Test Subscales Together Predicting Reading Measures in Grade 3 (N = 184)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.09		-.06	-.08		-.07	-.10
Vocabulary		.39***	.18**		.48***	.34***		.37***	.16**
Block Design		.17*	.10		.17*	.13*		.18*	.09
Step 2	.30***			.15***			.33***		
Real Words			.07			.11			.13*
Pseudowords			.25**			.23**			.21**
HF Bigrams			.40***			.24*			.37***
LF Bigrams			-.05			-.10			.04

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

On the other hand, the four subscales each added unique variance to the prediction of the reading measures when entered separately into the regression analyses. Table 8 shows the predictive ability of the four subscales in the three reading measures after controlling for age and IQ (see Appendix K for details of each analysis). Consistent with the previous analysis, real words (ΔR^2 = .05 to .09) and LF bigrams (ΔR^2 = .04 to .15) added less variance to the reading measures than pseudowords (ΔR^2 = .13 to .24) and HF bigrams (ΔR^2 = .11 to .28).

Table 8

Summary of ΔR^2 Values in Grade 3 Hierarchical Regression Analyses with the Quick Spell Test Subscales Individually Predicting Reading Measures (N = 184)

Step/Predictor	**Word Identification**	**Passage Comprehension**	**TOWRE**
1. Background Variables [a]	$.21^{***}$	$.30^{***}$	$.20^{***}$
2a. Real Words	$.07^{***}$	$.05^{***}$	$.09^{***}$
2b. Pseudowords	$.24^{***}$	$.13^{***}$	$.24^{***}$
2c. HF Bigrams	$.26^{***}$	$.11^{***}$	$.28^{***}$
2d. LF Bigrams	$.11^{***}$	$.04^{**}$	$.15^{***}$

[a] Background variables are age, vocabulary and block design.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Next, I addressed the fourth research question of whether the four QST subscales differed in the level of orthographic structure. Bowers (1996) designed the QST such that real words had the highest orthographic structure and the LF bigrams had the lowest orthographic structure. If the four subscales differed in orthographic structure, then OP should best account for the difference between any two subscales. For example, if real words had more orthographic structure than pseudowords, the variance added by real words after controlling for pseudowords (i.e., the difference between the subscales) would be best accounted for by OP. In other words, also controlling for OP should weaken the effect of the real words more than controlling for the other predictors. First, I examined the predictive power of real words after controlling for age, IQ, and pseudowords (Table 9). The real words did not add any significant variance to Word Identification and Passage Comprehension and only added 1% unique variance in TOWRE. Because real words predicted only a small amount of variance, the impact of the different reading predictors would therefore be difficult to differentiate; therefore, I did not conduct further analyses with the real words.

Table 9

Summary of ΔR^2 Values in Grade 3 Hierarchical Regression Analyses with Quick Spell Test Real Words Predicting Reading Measures (N = 184)

Step/Predictor	**Word Identification**	**Passage Comprehension**	**TOWRE**
1. Background Variables [a]	$.21^{***}$	$.30^{***}$	$.20 - .21^{***}$
2. Pseudowords	$.24^{***}$	$.13^{***}$	$.23 - .24^{***}$
3a. Real Words	.00	.01	$.01^{*}$

[a] Background variables include age, vocabulary and block design.
$^{*} p < .05.$ $^{**} p < .01.$ $^{***} p < .001$

Similarly, I examined the predictive power of pseudowords after controlling for age, IQ, and HF bigrams. Pseudowords uniquely predicted 4% of the variance in Word Identification and 3% of the variance in Passage Comprehension and TOWRE (Table 10). Next I entered the four reading predictors into an intermediate step to determine whether the variance added by the pseudowords (i.e., the difference between pseudowords and HF bigrams) was best explained by OP. Table 10 provides a summary of these analyses and Appendix M contains the details of each analysis. Similar to the analyses involving the QST total, OP and PD reduced the effect of pseudowords more than PA and NS. Entering either OP or PD into the third step reduced the pseudoword effect by 4% in Word Identification and 2% in Passage Comprehension and TOWRE (analyses d & e in Table 10). Entering OP and PD together reduced the pseudoword effect by 4% in Word Identification and 3% in Passage Comprehension and TOWRE (analysis f). Similar to previous analyses, I also examined the ability of the unique and shared variances of OP and PD in weakening the pseudoword effect. The OP composite residual (i.e., unique OP variance) did not significantly predict any of the reading measures and did not reduce the pseudoword effect (analysis g). On the other hand, although the Word Attack residual significantly predicted all three reading measures, it did not weaken the pseudoword effect in Passage Comprehension and TOWRE and only reduced the pseudoword effect by 1% in Word Identification (analysis h). Using similar subtraction methods as before, the OP-PD covariance would reduce the pseudoword effect by 3% in each reading measure. However, in contrast to the analyses involving the QST total, the OP-PD covariance weakened the pseudoword effect more than either of the OP or PD residuals. Therefore, this covariance appeared to best account for the difference between pseudowords and HF bigrams.

Table 10

Summary of ΔR^2 Values in Grade 3 Hierarchical Regression Analyses with Quick Spell Test Pseudowords Predicting Reading Measures

Step/Predictor	Word Identification	Passage Comprehension	TOWRE
1. Background Variables [a]	.21 - .22***	.30 - .31***	.20 - .21***
2. HF Bigrams	.26 - .27***	.10 - .11***	.27 - .28***
3a. Pseudowords	.04***	.03**	.03**
3b. PA Composite	.07***	.06***	.03**
4b. Pseudowords	.03**	.02**	.03**
3c. NS Composite	.02**	.03**	.16***
4c. Pseudowords	.02**	.02**	.01*
3d. OP Composite	.08***	.05***	.07***
4d. Pseudowords	.01	.01*	.01
3e. Word Attack	.23***	.11***	.11***
4e. Pseudowords	.00	.01	.01
3f. OP Composite	.25***	.13***	.13***
Word Attack			
4f. Pseudowords	.00	.00	.00
3g. OP Composite Residual	.00	.00	.00
4g. Pseudowords	.04***	.03**	.03**
3h. Word Attack Residual	.09***	.04***	.03**
4h. Pseudowords	.03**	.03**	.03**

Note: $N = 184$ for analyses a, b, & d; $N = 183$ for analyses e, f, g, & h; $N = 182$ for analysis c.
[a] Background variables include age, vocabulary and block design.
* $p < .05$. ** $p < .01$. *** $p < .001$

Finally, I studied the predictive ability of the HF bigrams after controlling for age, IQ, and LF bigrams. The HF bigrams uniquely predicted 15% of the variance in Word Identification, 7% of the variance in Passage Comprehension, and 13% of the variance in TOWRE (Table 11). I also entered the four reading predictors into an intermediate step to test whether OP could best account for the variance added by the HF bigrams (i.e., the difference between the HF and LF bigrams). Table 11 contains a summary of these analyses and Appendix N includes the details of each individual analysis. Similar to previous analyses, OP and PD had the greatest impact on reducing the HF bigrams' predictive ability. Adding OP reduced the HF bigram effect by 10% in Word Identification, 5% in Passage Comprehension, and 9% in TOWRE (analysis d in Table 11). Adding Word Attack reduced the HF bigrams effect by 13% in Word Identification, 6% in Passage Comprehension, and 10% in TOWRE (analysis e).

Examining the shared and unique variances of OP and PD, I found that controlling for OP and PD together reduced the HF bigram by 14% in Word Identification, 7% in Passage Comprehension, and 12% in TOWRE (analysis f). The OP composite residual did not predict any reading measure and reduced the HF bigram effect by only 1% in Word Identification and 2% in TOWRE (analysis g). The Word Attack residual weakened the HF bigram effect by 5% in Word Identification, 2% in Passage Comprehension, and 3% in TOWRE (analysis h). Therefore, the OP-PD covariance reduced the HF bigram effect by 8% in Word Identification, 5% in Passage Comprehension, and 7% in TOWRE. Overall, the OP-PD covariance accounted for the difference between the HF and LF bigrams better than either residual variable, although it performed more similarly to the Word Attack residual than to the OP composite residual.

Table 11

Summary of ΔR^2 Values in Grade 3 Hierarchical Regression Analyses with Quick Spell Test High Frequency Bigrams Predicting Reading Measures

Step/Predictor	Word Identification	Passage Comprehension	TOWRE
1. Background Variables [a]	.21 - .22***	.30 - .31***	.20 - .21***
2. LF Bigrams	.11 - .12***	.03 - .04**	.15 - .16***
3a. HF Bigrams	.15***	.07***	.13***
3b. PA Composite	.11***	.09***	.06***
4b. HF Bigrams	.11***	.05***	.10***
3c. NS Composite	.07***	.06***	.23***
4c. HF Bigrams	.11***	.05***	.05***
3d. OP Composite	.19***	.10***	.15***
4d. HF Bigrams	.05***	.02*	.04***
3e. Word Attack	.37***	.18***	.20***
4e. HF Bigrams	.02**	.01	.03***
3f. OP Composite Word Attack	.39***	.20***	.24***
4f. HF Bigrams	.01*	.00	.01*
3g. OP Composite Residual	.00	.00	.02*
4g. HF Bigrams	.14***	.07***	.11***
3h. Word Attack Residual	.14***	.06***	.05***
4h. HF Bigrams	.10***	.05***	.10***

Note: N = 184 for analyses a, b, & d; N = 183 for analyses e, f, g, & h; N = 182 for analysis c.
[a] Background variables include age, vocabulary and block design.
* $p < .05$. ** $p < .01$. *** $p < .001$

Summary of Grade 3 Analyses

In Grade 3, the QST total score significantly predicted 14 to 34% of the variance in the three reading measures, after controlling for IQ and age. OP and PD both weakened the predictive power of the QST total score more than PA and NS, indicating that OP and PD could best explain the relationship between the QST total score and reading ability. To further analyze the ability of OP and PD to account for the shared variance between QST and reading, I examined whether the unique variances of either OP or PD or the covariance between the two better accounted for the QST-reading ability shared variance. Overall, I found that the OP-PD covariance and the unique PD variance (i.e., the Word Attack residual) explained the QST-reading ability shared variance better than the unique OP variance (i.e., the OP composite residual).

Examining the four subscales, I found that each subscale significantly predicted all three reading measures when entered individually, after controlling for age and IQ. However, when all four subscales were entered together, only the pseudowords and HF bigrams remained individual significant predictors of the reading measures. Subsequently, I investigated whether the four subscales differed in the level of orthographic structure by determining whether OP could best account for the difference between any two subscales. Similar to the analyses involving the QST total score, OP and PD best explained the difference between pseudowords and HF bigrams and between HF bigrams and LF bigrams. Unlike the analyses involving the QST total, I found that the OP-PD covariance explained these differences better than either of the unique variances. However, the Word Attack and OP composite residuals both performed similarly in explaining the difference between pseudowords and HF bigrams, whereas the Word

Attack residual explained the difference between HF and LF bigrams better than the OP composite residual. Because the real words barely added any variance to the reading measures after controlling for age, IQ, and pseudowords, I could not adequately examine the difference between these two subscales.

Grade 4

Analyses Involving Quick Spell Test Total Score. Similar to the Grade 3 analyses, I examined the ability of the Grade 4 QST total score to predict the three Grade 4 reading measures after controlling for age and IQ. I subsequently added the four Grade 3 reading predictors into an intermediate step to determine the extent to which each of the predictors can weaken the QST effect on the reading measures. I also controlled for the same Grade 3 reading predictors in the Grade 4 analyses as in the Grade 3 analyses so that the results would be comparable. Table 12 shows a summary of these analyses and Appendix O contains the details of each analysis.

After controlling for age and IQ, the QST total score significantly added 22% of the variance in Word Identification, 11% of the variance in Passage Comprehension, and 27% of the variance in TOWRE. Moreover, OP and PD also weakened the QST effect the most: OP reduced the QST effect by 17% in Word Identification, 9% in Passage Comprehension, and 18% in TOWRE (analysis d in Table 12); and Word Attack reduced the effect by 21% in Word Identification and TOWRE, and 10% in Passage Comprehension (analysis e). Adding OP and PD together eliminated the QST effect in Word Identification and Passage Comprehension, and reduced the effect by 23% in TOWRE (analysis f). The OP composite residual also was not a significant predictor of the reading measures, only reducing the QST effect by 1% in Passage Comprehension and 2% in TOWRE (analysis g); whereas the Word Attack residual reduced the QST effect by 9% in Word Identification, 4% in Passage Comprehension, and 7% in TOWRE (analysis h). Therefore the PD-OP covariance would reduce the QST effect by 13% in Word Identification, 6% in Passage Comprehension, and 14% in TOWRE, thereby weakening the QST effect more than either of the residuals. Also, the covariance performed more similarly to the Word Attack residual than the OP composite residual.

Analyses Involving Quick Spell Test Subscale Scores. Next, I also examined the predictive ability of the four Grade 4 QST subscales in the Grade 4 reading measures. As shown in Table 13, the four subscales together uniquely predicted 23% of the variance in Word Identification, 10% of the variance in Passage Comprehension, and 29% of the variance in TOWRE, after controlling for age and IQ. However, only the HF bigram subscale was a significant individual predictor of Word Identification (β = .40) and TOWRE (β = .43). None of the four subscales were significant individual predictors of Passage Comprehension. However, when entered separately into the analysis, all four QST subscales significantly predicted the reading measures. The only exception is that the real words did not significantly predict Passage Comprehension. Table 14 shows a summary of these analyses and Appendix P provides the details of each analysis.

Table 12

Summary of ΔR^2 Values in Grade 4 Hierarchical Regression Analyses with Quick Spell Total Predicting Reading Measures

Step/Predictor	Gr. 4 Word Identification[a]	Gr. 4 Passage Comprehension[b]	Gr. 4 TOWRE[a]
1. Background Variables[c]	.21 - .22***	.26 - .28***	.14 - .15***
2a. Gr. 4 QST Total	.22***	.11***	.27***
2b. Gr. 3 PA Composite	.21***	.08***	.11***
3b. Gr. 4 QST Total	.14***	.08***	.20***
2c. Gr. 3 NS Composite	.13***	.08***	.37***
3c. Gr. 4 QST Total	.12***	.06***	.07***
2d. Gr. 3 OP Composite	.23***	.14***	.20***
3d. Gr. 4 QST Total	.05***	.02*	.09***
2e. Gr. 3 Word Attack	.46***	.21***	.29***
3e. Gr. 4 QST Total	.01*	.01	.06***
2f. Gr. 3 OP Composite	.47***	.22***	.31***
Gr. 3 Word Attack			
3f. Gr. 4 QST Total	.00	.00	.04**
2g. Gr. 3 OP Composite Residual	.00	.01	.01
3g. Gr. 4 QST Total	.23***	.10***	.25***
2h. Gr. 3 Word Attack Residual	.24***	.08***	.11***
3h. Gr. 4 QST Total	.13***	.07***	.20***

[a] $N = 146$ for analyses a, b, & d; $N = 145$ for analyses e, f, g, & h; $N = 144$ for analysis c.
[b] $N = 145$ for analyses a, b, & d; $N = 144$ for analyses e, f, g, & h; $N = 143$ for analysis c.
[c] Background variables include age, vocabulary and block design.
* $p < .05$. ** $p < .01$. *** $p < .001$

Table 13

Summary of Hierarchical Regression Analyses with Quick Spell Subscales Together Predicting Reading Measures in Grade 4

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.26^{***}$			$.14^{***}$		
Age		-.06	-.08		.05	.02		.04	.02
Vocabulary		$.37^{***}$	$.24^{**}$		$.43^{***}$	$.35^{***}$		$.32^{***}$	$.18^{*}$
Block Design		$.20^{*}$	.11		$.16^{*}$	.09		.11	-.00
Step 2	$.23^{***}$			$.10^{***}$			$.29^{***}$		
Real Words			.05			.02			.10
Pseudowords			.11			-.01			.15
HF Bigrams			$.40^{**}$			.22			$.43^{***}$
LF Bigrams			.00			.15			-.02

Note: *N* = 146 for analyses involving Grade 4 Word Identification and TOWRE; *N* = 145 for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Table 14

Summary of ΔR^2 Values in Grade 4 Hierarchical Regression Analyses with the Quick Spell Test Subscales Individually Predicting Reading Measures

Step/Predictor	**Word Identification**	**Passage Comprehension**	**TOWRE**
1. Background Variables [a]	$.21^{***}$	$.26^{***}$	$.14^{***}$
2a. Real Words	$.05^{**}$	.02	$.08^{***}$
2b. Pseudowords	$.15^{***}$	$.04^{**}$	$.19^{***}$
2c. HF Bigrams	$.22^{***}$	$.09^{***}$	$.27^{***}$
2d. LF Bigrams	$.10^{***}$	$.08^{**}$	$.11^{***}$

Note: N = 146 for analyses involving Grade 4 Word Identification and TOWRE; N = 145 for analyses involving Grade 4 Passage Comprehension.
[a] Background variables include age, vocabulary and block design.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Similar to the Grade 3 analyses, I subsequently conducted regression analyses to determine whether orthographic structure would best account for the difference between the types of QST strings in Grade 4. First, I examined the ability of real words to predict the reading measures after controlling for age, IQ, and pseudowords (Table 15). Real words did not significantly predict Word Identification and Passage Comprehension; thus I did not conduct further analyses with real words predicting these measures. Real words uniquely predicted 3% of the variance in TOWRE; hence I entered the four Grade 3 reading predictors into an intermediate step to see if OP would best explain the variance added by real words (i.e., the difference between real words and pseudowords). Table 15 provides a summary of these analyses and Appendix Q contains the details of each analysis. Adding PA reduced the real word effect by 1% (analysis b in Table 15), adding

NS reduced the effect by 3% (i.e., eliminated the real word effect; analysis c), and adding either OP or PD reduced the effect by 2% (analyses d & e). Adding OP and PD together also reduced the real word effect by 2% (analysis f), and adding either the OP or Word Attack residual did not reduce the effect (analyses g & h); thus the OP-PD covariance also reduced the real word effect by 2% and explained the difference between real words and pseudowords better than either residual.

Next, I also evaluated the predictive ability of the Grade 4 pseudowords in the Grade 4 reading measures, after controlling for age, IQ, and the Grade 4 HF bigrams. The results showed that the pseudoword subscale did not add significant variance to any of the three reading measures (Table 16). Therefore, I did not carry out further analyses with the pseudoword subscale.

Table 15

Summary of ΔR^2 Values in Grade 4 Hierarchical Regression Analyses with Quick Spell Test Real Words Predicting Reading Measures

Step/Predictor	Gr. 4 Word Identification	Gr. 4 Passage Comprehension	Gr. 4 TOWRE
1. Background Variables [a]	$.21^{***}$	$.26^{***}$	$.14 - .15^{***}$
2. Gr. 4 Pseudowords	$.15^{***}$	$.04^{**}$	$.19 - .20^{***}$
3a. Gr. 4 Real Words	.01	.01	$.03^{*}$
3b. Gr. 3 PA Composite			$.06^{**}$
4b. Gr. 4 Real Words			$.02^{*}$
3c. Gr. 3 NS Composite			$.23^{***}$
4c. Gr. 4 Real Words			.00
3d. Gr. 3 OP Composite			$.07^{***}$
4d. Gr. 4 Real Words			.01
3e. Gr. 3 Word Attack			$.15^{***}$
4e. Gr. 4 Real Words			.01
3f. Gr. 3 OP Composite			$.15^{***}$
Gr. 3 Word Attack			
4f. Gr. 4 Real Words			.01
3g. Gr. 3 OP Composite Residual			.00
4g. Gr. 4 Real Words			$.03^{*}$
3h. Gr. 3 Word Attack Residual			$.07^{***}$
4h. Gr. 4 Real Words			$.03^{*}$

Note: $N = 146$ for Word Identification analysis; $N = 145$ for Passage Comprehension analysis.
Note: For TOWRE analyses, $N = 184$ for analyses a, b, & d; $N = 183$ for analyses e, f, g, & h; $N = 182$ for analysis c.
[a] Background variables include age, vocabulary and block design.
$^{*} p < .05$. $^{**} p < .01$. $^{***} p < .001$

Table 16

Summary of Hierarchical Regression Analyses with Quick Spell Test Pseudowords Predicting Reading Measures in Grade 4

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.07		.05	.04		.04	.02
Vocabulary		.37***	.25***		.43***	.36***		.32***	.19**
Block Design		.20**	.10		.16*	.10		.11	.00
Step 2	.22***			.09***			.27***		
HF Bigrams		.50***	.42***		.32***	.33**		.55***	.44***
Step 3	.01			.00			.01		
Pseudowords			.12			-.01			.16

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

Finally, I examined whether the Grade 4 HF bigrams predicted the Grade 4 reading measures after controlling for age, IQ, and the Grade 4 LF bigrams. As with previous analyses, I also conducted analyses with the four Grade 3 reading predictors entered separately into an intermediate step to determine whether OP would best account for the variance added by the HF bigrams (i.e., the difference between the HF and LF bigrams). Table 17 shows the summary of these analyses and Appendix R provides the details of each separate analysis. The HF bigrams significantly added 12% of the variance in Word Identification, 2% of the variance in Passage Comprehension, and 16% of the variance in TOWRE, after controlling for age, IQ, and the LF bigrams. Similar to the Grade 3 results, OP and PD greatly reduced the amount of unique variance predicted by the HF bigrams. The addition of OP reduced the HF bigram effect by 9% in Word Identification, 2% in Passage Comprehension, and 8% in TOWRE (analysis d in Table 17). Adding Word Attack reduced the HF bigram effect by 10% in Word Identification and TOWRE, and 2% in Passage Comprehension (analysis e).

Moreover, adding OP and PD together reduced the effect by 11% in Word Identification, 2% in Passage Comprehension, and 12% in TOWRE (analysis f). The OP composite residual reduced the HF bigram effect by only 1% in TOWRE (analysis g), whereas the Word Attack residual reduced the effect by 3% in Word Identification and TOWRE, and by 1% in Passage Comprehension (analysis h). Accordingly, the OP-PD covariance was also the best predictor of the difference between the HF and LF bigrams, as it reduced the HF bigram effect by 8% in Word Identification and TOWRE, and by 1% in Passage Comprehension.

Table 17

Summary of ΔR^2 Values in Hierarchical Regression Analyses with Quick Spell Test High Frequency Bigrams Predicting Reading Measures in Grade 4

Step/Predictor	**Gr. 4 Word Identification** [a]	**Gr. 4 Passage Comprehension** [b]	**Gr. 4 TOWRE**
1. Background Variables [c]	.21 - .22***	.26 - .28***	.14 - .15***
2. Gr. 4 LF Bigrams	.10***	.08***	.11***
3a. Gr. 4 HF Bigrams	.12***	.02*	.16***
3b. Gr. 3 PA Composite	.17***	.06***	.08***
4b. Gr. 4 HF Bigrams	.07***	.01	.12***
3c. Gr. 3 NS Composite	.08***	.04***	.28***
4c. Gr. 4 HF Bigrams	.08***	.01	.06***
3d. Gr. 3 OP Composite	.15***	.08***	.12***
4d. Gr. 4 HF Bigrams	.03**	.00	.08***
3e. Gr. 3 Word Attack	.37***	.14***	.19***
4e. Gr. 4 HF Bigrams	.02**	.00	.06***
3f. Gr. 3 OP Composite	.37***	.15***	.21***
Gr. 3 Word Attack			
4f. Gr. 4 HF Bigrams	.01*	.00	.04**
3g. Gr. 3 OP Composite Residual	.00	.00	.00
4g. Gr. 4 HF Bigrams	.13***	.02	.15***
3h. Gr. 3 Word Attack Residual	.19***	.06***	.07***
4h. Gr. 4 HF Bigrams	.09***	.01	.13***

[a] *N* = 146 for analyses a, b, & d; *N* = 145 for analyses e, f, g, & h; *N* = 144 for analysis c.
[b] *N* = 145 for analyses a, b, & d; *N* = 144 for analyses e, f, g, & h; *N* = 143 for analysis c.
[c] Background variables include age, vocabulary and block design.
* $p < .05$. ** $p < .01$. *** $p < .001$

Summary of Grade 4 Analyses

In general, the Grade 4 results paralleled those in Grade 3. The QST total uniquely predicted 11 to 27% of the variance in the three reading measures, after controlling for age and IQ. These percentages are slightly lower than those in Grade 3. Amongst the Grade 3 reading predictors (i.e., PA, NS, PD, and OP), OP and PD were also the best predictors of the relationship between QST and the reading measures in Grade 4. However, the OP-PD covariance explained the QST-reading ability relationship better than either the OP or Word Attack residuals in Grade 4; whereas the covariance performed similarly to the Word Attack residual in Grade 3. Moreover, similar to Grade 3, the four subscales each predicted the three reading measures when entered separately, after controlling for age and IQ. However, when entered together, only HF bigrams significantly predicted Word Identification and TOWRE, and none of the four subscales was a significant individual predictor of Passage Comprehension.

When examining whether OP could account for the difference between subscales, I found that the real words predicted a significant amount of variance only in the TOWRE, after controlling for age, IQ, and pseudowords. Contrary to other analyses, adding OP, PD, *or* NS into an intermediate step reduced the real word effect to nonsignificance. In other words, any of the three predictors could account for the variance added by the real words (i.e., the difference between real words and pseudowords). Considering that NS and TOWRE share a common speed component, however, NS's ability to weaken the real word effect on TOWRE is unsurprising. On the other hand, after controlling for age, IQ, and LF bigrams, OP and PD weakened the effect of HF bigrams more than PA and NS; therefore OP and PD were consistently the strongest predictors of the difference between the real words and pseudowords and between the HF and LF bigrams. Moreover, the OP-PD covariance explained the difference both between the real words and pseudowords and between the HF and LF bigrams better than either residual, echoing the results of previous analyses. Unfortunately, I could not examine the difference between the Grade 3 real words and pseudowords and the difference between the Grade 4 pseudowords and HF bigrams adequately, due to the lack of variance added by the latter subscale in each pair. Therefore, direct comparisons cannot be made across grades for these two pairs of subscales.

Longitudinal Analyses

To address the latter part of my second research question, I investigated whether the Grade 3 QST measures could predict growth in reading ability. Therefore, I used the Grade 3 QST scores to predict the Grade 4 reading measures, controlling for age and IQ. For each Grade 4 reading measure, I also controlled for its Grade 3 counterpart as an autoregressor (e.g., control for Grade 3 Word Identification when predicting Grade 4 Word Identification).

Analyses Involving Quick Spell Test Total Score

The first set of analyses examined the ability of the Grade 3 QST total score to predict Grade 4 reading ability. As Table 18 shows, after controlling for age, IQ, and Grade 3 reading ability, the Grade 3 QST total score significantly predicted 1% of the variance in Grade 4 Word Identification and Passage Comprehension, and did not add unique variance in Grade 4 TOWRE. Appendix S provides more details about each analysis. Because the Grade 3 QST total score added little unique variance, I did not conduct any further longitudinal analyses with the measure.

Table 18

Summary of ΔR^2 Values in Hierarchical Regression Analyses with Grade 3 Quick Spell Total Predicting Grade 4 Reading Measures (N = 143)

Step/Predictor	**Gr. 4 Word Identification**	**Gr. 4 Passage Comprehension**	**Gr. 4 TOWRE**
1. Background Variables [a]	$.21^{***}$	$.25^{***}$	$.14^{***}$
2a. Gr. 3 Word Identification	$.55^{***}$		
3a. Gr. 3 QST Total	$.01^{*}$		
2b. Gr. 3 Passage Comprehension		$.37^{***}$	
3b. Gr. 3 QST Total		$.01^{*}$	
2c. Gr. 3 TOWRE			$.49^{***}$
3c. Gr. 3 QST Total			.01

[a] Background variables include age, vocabulary and block design.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Analyses Involving Quick Spell Test Subscale Scores

In similar analyses, I evaluated the predictive ability of the Grade 3 QST subscales in the Grade 4 reading measures, controlling for age, IQ, and Grade 3 reading ability. Table 19 provides a summary of these analyses and Appendix T contains the details of each analysis. When entered together, the Grade 3 QST subscales also did not add unique variance to any of the Grade 4 reading measures; thus I did not conduct further longitudinal analyses with these subscales.

Table 19

Summary of ΔR^2 Values in Hierarchical Regression Analyses with Grade 3 Quick Spell Subscales Predicting Grade 4 Reading Measures (N = 143)

Step/Predictor	**Gr. 4 Word Identification**	**Gr. 4 Passage Comprehension**	**Gr. 4 TOWRE**
1. Background Variables [a]	$.21^{***}$	$.25^{***}$	$.14^{***}$
2a. Gr. 3 Word Identification	$.55^{***}$		
3a. Gr. 3 QST Subscales [b]	.01		
2b. Gr. 3 Passage Comprehension		$.37^{***}$	
3b. Gr. 3 QST Subscales [b]		.02	
2c. Gr. 3 TOWRE			$.49^{***}$
3c. Gr. 3 QST Subscales [b]			.01

[a] Background variables include age, vocabulary and block design.
[b] Quick Spell subscales include Real Words, Pseudowords, High Frequency Bigrams, and Low Frequency Bigrams.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

CHAPTER 4: DISCUSSION

In this study, I examined four research questions regarding the Quick Spell Test (QST) and its relationship with various other reading variables. The first question evaluated Bowers' (1996; 2001) proposition that readers with good orthographic processing (OP) skills should benefit more from orthographic structure than readers with poor OP skills, and therefore readers with good OP should show a greater difference between strings with more orthographic structure and strings with less orthographic structure than readers with low OP. The second question examined the relationship between QST and other reading measures; and the third question investigated whether QST is indeed a measure of OP, as designed by Bowers (1996). Finally, the fourth question pertained to the relative degree of orthographic structure in the four QST subscales. The sections below interpret the results of the current study in the context of these four research issues and discuss the overall implications of this study on the QST literature and educational practices.

Research Question 1: The Benefit of Orthographic Structure

To evaluate Bowers' (1996; 2001) hypothesis that readers with good OP skills would benefit more from orthographic structure than readers with poor OP skills, I examined the regression slopes between OP and the QST. The regression slope patterns only partially supported Bowers' hypothesis. In Grade 3, the difference between the pseudoword and low frequency (LF) bigram regression lines and the difference between the high frequency (HF) bigram and LF bigram regression lines were greater at the higher end of the OP axis than at the lower end (Figure 2). This trend indicates that the higher orthographic structure of pseudowords and HF bigrams, relative to the LF bigrams, benefited readers with good OP more than readers with low OP. However, the higher orthographic structure of pseudowords compared to HF bigrams benefited all readers similarly, as the difference between the two regression lines was similar across different OP levels. In contrast, the difference between the real word and pseudoword regression lines was higher at the lower levels of OP than at the higher levels. This pattern could indicate that the higher orthographic structure of real words compared to pseudowords benefited readers with low OP more than readers with high OP. It is possible that readers with good OP skills perceive the same orthographic structure in both real words and pseudowords, as both string types allow for unitization of the four letters. Thus, in Grades 3 and 4, only readers with low OP would benefit from the higher orthographic structure of the real words, compared to the pseudowords. On the other hand, this finding could have been merely due to the ceiling effect in the real word subscale, which may have limited the slope of the regression line.

A similar pattern was observed in Grade 4 (Figure 3) with the difference between the HF and LF bigram regression lines and the difference between the pseudoword and LF bigram regression lines being smaller at low OP levels and greater at high OP levels. Thus readers with higher OP benefited more from the higher orthographic structure of the pseudowords and HF bigrams, relative to the LF bigrams, than readers with poor OP. Conversely, the difference between the real word and pseudoword regression lines and the difference between the pseudoword and HF bigram regression lines were greater at the low end of OP and smaller at the high end of OP. As previously discussed, this finding could indicate that readers with low OP benefited more from the higher orthographic structure of the real words, compared to the pseudowords, and the higher orthographic structure of the pseudowords, compared to the HF bigrams, than readers with high OP. On the other hand, these results could have also been due to the ceiling effect of the real word and pseudoword subscales.

Furthermore, I also explored the regression slopes between naming speed (NS) and the QST to examine whether they were similar to those between OP and the QST. The results in both Grades 3 (Figure 4) and 4 (Figure 5) paralleled the OP analyses, thus providing support for the hypothesized link between NS and OP (e.g., Wolf & Bowers, 1999). However, considering the problematic distributions of the real word and pseudoword subscales, future studies should replicate these analyses with a refined version of the QST.

Research Question 2: The QST and Reading Ability

Many of the previous QST studies did not specifically examine the measure's ability to predict other established measures of reading ability. In the current study, I used regression analyses to determine how well the QST related to Word Identification, Passage Comprehension, and TOWRE (i.e., speeded word identification). Concurrently, the QST total score predicted 14 to 34% of the variance in these three reading measures in Grade 3 (Table 6), and predicted 11 to 27% of the variance in Grade 4 (Table 12), after controlling for age and IQ. In the longitudinal analyses, the Grade 3 QST total score also predicted Grade 4 reading ability when controlling for age and IQ. However, the Grade 3 QST total became a nonsignificant predictor of Grade 4 reading when also controlling for previous reading ability (i.e., Grade 3 reading ability; Table 18). Therefore, the QST total score was a good predictor of concurrent reading ability, even when taking age and IQ into account; but could not predict growth in reading ability from Grade 3 to Grade 4.

The four QST subscales also predicted concurrent reading ability in both Grades 3 (Table 8) and 4 (Table 14) when entered separately and controlling for age and IQ. However, real words and LF bigrams were weaker predictors than pseudowords and HF bigrams in Grade 3, and became nonsignificant individual predictors when all four subscales were entered together after age and IQ (Table 7). A similar pattern was observed in Grade 4, except that only the HF bigrams remained a significant individual predictor of the reading measures when all four subscales were entered together after age and IQ (Table 13). The weaker predictive power of some subscales may have been due to the distributional properties of those subscales. For example, the real word subscales portrayed a ceiling effect in both Grades 3 and 4, resulting in reduced variance. The LF bigram subscale, in contrast, showed a slight floor effect in both grades, which also led to reduced variance. The reduced variance in these subscales could have led to their weak predictive powers. This explanation is further supported by the trend observed in the pseudoword subscale: its predictive power decreased from Grade 3 to Grade 4, as the distribution moved towards a slight ceiling effect in Grade 4. The distributional problems with some of the QST subscales suggest that future studies should explore the possibility of creating more stimuli or varying the exposure time, in order to maximize the variance and predictive ability of each subscale.

Research Question 3: Is the QST an Orthographic Measure?

Because the QST is proposed to be a measure of OP, I evaluated whether OP would be the best predictor of the QST. More specifically, I examined whether OP could best explain the variance shared between the QST and other reading measures, as I was interested in examining the QST variance that was related to reading ability. The results generally showed that OP accounted for more of the shared variance between the QST total score and the reading measures than phonological awareness (PA) or NS in both Grades 3 (Table 6) and 4 (Table 12), supporting the hypothesis. One caveat, however, was that the phonological decoding (PD) measure also explained as much of the QST-reading ability variance as OP. Thus, one could argue that the QST measured both OP and PD.

Because OP and PD were consistently the strongest predictors of the shared variance between the QST and the reading measures, I analyzed the shared and unique variances associated with OP and PD. By entering OP and PD together into the intermediate step of the regression analyses, I examined their joint ability to explain the variance in reading ability that the QST predicted. By creating residual variables, I examined their unique ability to account for the same QST-reading ability shared variance. Finally, by subtraction methods, I deduced the ability of the OP-PD covariance to explain the QST-reading ability relationship. In Grade 3, I found that both the OP-PD covariance and the PD residual accounted for the shared variance between the QST and reading ability better than the OP residual (Table 6). However, in Grade 4, the OP-PD covariance explained the QST-reading ability relationship better than either of the residuals (Table 12). Because of the mixed findings, the issue of whether the QST is mainly an OP task is yet unresolved.

The nature of the OP-PD covariance is ambiguous and allows for several possible interpretations. One possibility is that the covariance is orthographic in nature. Although Word Attack is a measure of PD and therefore meant to measure phonological processes, it correlated with both PA (r = .56) and OP (r = .62) and appeared to be equally influenced by both phonological and orthographic processes (Appendix H). By creating the Word Attack residual, I eliminated the variance in Word Attack that was predicted by OP, thereby leaving the Word Attack variance that is more phonological in nature. Thus the covariance may reflect the Word Attack variance that is more orthographic in nature. The finding that the OP residual did not significantly predict any reading measure also supports this interpretation, as it may indicate that most of the OP variance was shared with Word Attack. If Word Attack was correlated with PA more than with OP (i.e., if Word Attack was more phonological in nature), OP and PD should not share so much variance as to create an OP residual that lacked predictive ability. Therefore, the OP-PD covariance may reflect mainly orthographic processes.

On the other hand, the PD-OP covariance could reflect an ability to integrate OP and PD processes that is essential for proficient reading. Skilled readers are often able to integrate immediate holistic recognition of familiar words with rapid grapheme-phoneme conversion of unfamiliar words. Moreover, according to Share (1995), readers acquire OP skills through successful PD experiences; in other words, PD is a necessary route for OP development. Hence, the strong predictive ability of the OP-PD covariance found in these results may reflect this OP development. The difference between the Grade 3 and 4 results supports this explanation. In Grade 3, both the covariance and the PD residual explained the QST-reading ability relationship better than the OP residual; whereas the covariance explained the relationship better than both the residuals in Grade 4. This trend across the grades could indicate that the children were more adept at integrating the OP and PD skills in Grade 4.

Research Question 4: Degree of Orthographic Structure in the QST Subscales

Although Bowers (1996) designed the four QST subscales to differ in the degree of orthographic structure, few, if any, studies have examined whether the subscales conform to the proposed ranking of orthographic structure (i.e., real words > pseudowords > HF bigrams > LF bigrams). Thus I explored this issue by determining whether OP could best account for the difference between pairs of subscales. In Grade 3, I examined the difference between pseudowords and HF bigrams by analyzing the variance added by the pseudowords after controlling for age, IQ, and HF bigrams (Table 10). Similarly, I examined the difference between HF and LF bigrams by analyzing the variance added by the HF bigrams after controlling for age, IQ, and LF bigrams (Table 11). For both sets of analyses, OP and PD explained the difference between the subscales better than PA and NS. In Grade 4, I also found that OP and PD explained the difference between real words and pseudowords (Table 15) and the difference between HF and LF bigrams (Table 17) better than PA and NS. Therefore, as with the analyses involving the QST total scores, the hypothesis that the subscales differed mainly in orthographic structure is partially supported.

As with the previous analyses, I also investigated whether the common or unique variances associated with OP and PD would be the better predictor of these differences. In both Grades 3 and 4, I found that the OP-PD covariance could account for the difference between the pairs of subscales better than either residual. These results are consistent with those involving the QST total. Because the nature of this covariance is ambiguous, for reasons outlined in the previous section, it is still unclear as to whether the subscales actually differ in orthographic structure or whether they also differ in other dimensions. Moreover, because the real words did not predict the reading measures after controlling for pseudowords in Grade 3, and the pseudowords did not add unique variance after controlling for HF bigrams in Grade 4, the subscale analyses could not be easily compared across the two grades. Therefore, future studies should investigate this issue with a refined QST measure that would eliminate the distributional problems with the real word and pseudoword subscales, thereby creating adequate variance to analyze.

Implications

Theoretical Implications

The results of the current study contribute to both the QST literature and the general reading literature. In the QST literature, most studies have examined group differences between readers with PA or NS deficits (or both) and readers with adequate skills. This study, however, examined a group of Grade 3 and 4 children with diverse reading abilities and used regression analyses to preserve maximum power in the continuous measures and to avoid the methodological problems of grouping. Although this study did not parallel the methodology of previous research, it provides more information about the characteristics and predictive ability of the QST in reading. The results in this study showed that the QST is a good concurrent predictor of reading, even after controlling for age and IQ. However, the nature of the QST was not resolved in this study, as OP and PD explained the relationship between the QST and reading equally well. Because this study only included a single PD measure, and it related equally well to both phonological and orthographic processes, future studies should use multiple PD measures to further examine the true nature of the QST. In particular, researchers should strive to use PD measures which rely mostly on phonological processes instead of on orthographic abilities, such as creating nonwords which do not resemble other real words and therefore discourage reading by analogy (e.g., using “ wabi” instead of “wame,” which resembles “came”).

The equal predictive ability of OP and PD in this study also brings a new perspective to the general reading literature. Many researchers regard OP and PD as distinct and separate processes; for example, the dual route theory of reading (e.g., Castles & Coltheart, 1993; Castles, 2006) describes a lexical route in which readers holistically recognize words (i.e., OP) and a nonlexical route in which readers convert text into sound using grapheme-phoneme conversion rules (i.e., PD). However, in the current study, OP and PD shared a large amount of variance (r = .62; Appendix H), such that the OP-PD covariance was a stronger predictor of the QST-reading ability relationship than either of the OP or PD residuals. Therefore, these results may remind us that OP and PD processes often collaborate, rather than compete, in skilled reading.

Moreover, many of the QST subscales had distributional problems in this study, resulting in inadequate variance with which to predict reading. In light of these observations, researchers should better calibrate the QST to match different reading levels in order to provide a more sensitive measure. With both the Grade 3 and 4 samples, the real words, and perhaps also pseudowords, exhibited a ceiling effect whereas the LF bigrams displayed a floor effect. One way to increase the variance may be to create shorter and longer letter strings. For example, varying the string length from three to five letters may eliminate both ceiling and floor effects. Another method may be to simulate the threshold studies in perception research, in which researchers vary the stimulus exposure time by increments to uncover the “threshold” at which participants are able to accurately report most strings. Thus poorer readers would be expected to have higher thresholds (i.e., need longer exposures) than good readers. By improving the distributional properties of the QST subscales, future research can then continue examining issues such as the relative degree of orthographic structure between the subscales (i.e., the fourth research question in this study).

Practical Implications

Because this study has shown the QST to be a good predictor of concurrent reading ability, educators may be able to use the QST as a form of reading assessment. The QST measure is a relatively easy task to create, administer, and score; for example, classroom teachers could use Microsoft PowerPoint to simulate the QST measure and administer the task to an entire class at the same time. Administering the task requires only approximately 10 to 15 minutes, and most children may not suspect that the QST is related to reading assessment and therefore are not likely to object to performing the task. Children are also unlikely to be able to remember all 40 of the stimuli; thus the measure can be re-administered at a later date without great practice effects. Moreover, the current study and previous studies have shown children to consistently perform better on the subscales with more orthographic structure (i.e., real words accuracy > pseudowords accuracy > HF bigrams accuracy > LF bigrams accuracy). Therefore, educators could estimate a child’s OP skills easily by determining which subscale the child has difficulty with. Good readers may only have difficulty with the LF bigram subscale; poorer readers may have difficulty with the HF bigram or pseudoword subscale.

Conclusions

Overall, this study provides valuable information about the predictive ability and the characteristics of the QST. Knowing that QST relates well to concurrent reading ability supports its use in future studies as a reading measure or reading predictor, and discovering the distributional problems with some of the QST subscales allow researchers to ameliorate the psychometric properties of the measure. Because the QST is easy to create, administer, and score, understanding the true nature of this task and resolving its current distributional imperfections will not only allow researchers to conduct more detailed studies with the QST but also allow educators to use the QST for reading assessment purposes.

CHAPTER 5: REFERENCES

Badian, N. A. (2001). Phonological and orthographic processing: Their roles in reading prediction. *Annals of Dyslexia, 51*, 179-202.

Baker, K. (2002). Visual processes and the double deficit hypothesis for reading disability. Unpublished doctoral dissertation.

Barker, T. A., Torgesen, J. K., & Wagner, R. K. (1992). The role of orthographic processing skills on five different reading skills. *Reading Research Quarterly, 27*, 334-345.

Bowers, P. G. (1995). Tracing symbol naming speed's unique contributions to reading disabilities over time. *Reading and Writing: An Interdisciplinary Journal, 7*, 189-216.

Bowers, P. G. (1996). The effects of single and double deficits in phonemic awareness and naming speed on new tests of orthographic knowledge. Paper presented to meetings of the Society for the Scientific Study of Reading.

Bowers, P. G. (2001). Exploration of the basis for rapid naming's relationship to reading. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 41-63). York Press: Timonium, MD.

Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. *Reading and Writing: An Interdisciplinary Journal, 15*, 109-126.

Bowers, P. G., Steffy, R. A., & Swanson, L. B. (1986). Naming speed, memory, and visual processing in reading disability. *Canadian Journal of Behavioural Science, 18*, 209-223.

Bowers, P. G., Sunseth, K., & Golden, J. (1999). The route between rapid naming and reading progress. *Scientific Studies of Reading, 3*, 31-53.

Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability: Multiple measures of a singular process. *Journal of Experimental Child Psychology, 51*, 195-219.

Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal, 5*, 68-85.

Bowey, J. A. (1996). Phonological sensitivity as a proximal contributor to phonological recoding skills in children's reading. *Australian Journal of Psychology, 48*, 113-118.

Bowey, J. A., McGuigan, M., & Ruschena, A. (2005). On the association between serial naming speed for letters and digits and word-reading skill: Towards a developmental account. *Journal of Research in Reading, 28*, 400-422.

Bowey, J. A., Storey, T., & Ferguson, A. N. (2004). The association between continuous naming speed and word reading skill in fourth- to sixth-grade children. *Australian Journal of Psychology, 56*, 155-163.

Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91(3)*, 403-414.

Castles, A. (2006). The dual route model and the developmental dyslexias. *London Review of Education, 4(1)*, 49-61.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149-180.

Cunningham, A. E., Perry, K., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing: An Interdisciplinary Journal, 14*, 549-568.

Denckla, M. B., & Rudel, R. G. (1976). Rapid ‘automatized’ naming (R. A. N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools, 36*, 285-293.

Ishaik, G., (2003). Understanding the interrelations between processes underlying reading accuracy in poor and good reader groups. Unpublished doctoral dissertation.

Jarvis, B. (2000). *DirectRT* (Version 2000.4.9) [Computer software]. New York: Empirisoft.

Jorm, A. F., Share, D. L., Maclean, R., & Matthews, R. G. (1984). Phonological recoding skills and learning to read: A longitudinal study. *Applied Psycholinguistics, 5*, 201-207.

Kail, R., Hall, L. K., & Caskey, B. J. (1999). Processing speed, exposure to print, and naming speed. *Applied Psycholinguistics, 20*, 303-314.

Kirby, J. R., Parrila, R., Deacon, S. H., & Wade-Woolley, L. (2004-2007). Beyond the double deficit: The roles of orthographic and morphological processing in the development of reading and spelling. Social Sciences and Humanities Research Council of Canada standard research grant.

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003) Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 95*, 453-464.

Kyte, C. S., & Johnson, C. J. (2004). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology, 93*, 166-185.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323.

Levy, B. A., Gong, Z., Hessels, S., Evans, M. A., & Jared, D. (2006). Understanding print: Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology, 93*, 63-93.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33*, 325-333.

McDougall, P., Borowsky, R., MacKinnon, G. E., & Hymel, S. (2005). Process dissociation of sight vocabulary and phonetic decoding in reading: A new perspective on surface and phonological dyslexias. *Brain and Language, 92*, 185-203.

Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994) Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243-278). Baltimore, Maryland: Paul H. Brookes Publishing Co..

Parrila, R., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development? *Scientific Studies of Reading, 8*, 3-26.

Plaza, M., & Cohen, H. (2003). The interaction between phonological processing, syntactic awareness, and naming speed in the reading and spelling performance of first-grad children. *Brain and Cognition, 53*, 287-292.

Rueffer, K. A. (2000). An examination of the factors underlying the development of skilled reading. A thesis presented to the University of Waterloo in partial fulfillment of a Masters of Arts degree in Psychology.

Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (2002). Relationship of rapid automatized naming and phonological awareness in early reading development: Implications for the double-deficit hypothesis. *Journal of Learning Disabilities, 35*, 245-256.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten predictors of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96*, 265-282.

Share, D. L. (1995). Phonological recoding and self-teaching: *sine qua non* of reading acquisition. *Cognition, 55*, 151-218.

Sprugevica, I., & Høien, T. (2003). Early phonological skills as a predictor of reading acquisition: A follow-up study from kindergarten to the middle of grade 2. *Scandinavian Journal of Psychology, 44*, 119-124.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402-433.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*, 407-440.

Tabachnick, B. C., & Fidell, L. S. (2001). *Using multivariate statistics* (4^{th} ed.). Needham Heights, MA: Allyn & Bacon.

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

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second-to-fifth-grade children. *Scientific Studies of Reading, 1*, 161-185.

van den Bos, K. P., Zijlstra, B. J. H., & lutje Spelberg, H. C. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading, 6*, 25-49.

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

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processes (CTOPP).* Austin, TX: Pro-Ed.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio: Harcourt Brace & Company.

Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly, 26*, 123-141.

Wolf, M. (1997). A provisional, integrative account of phonological and naming-speed deficits in dyslexia: Implications for diagnosis and intervention. In B. Blachman (Ed.), *Foundations of reading acquisition* (pp. 67-92). Hillsdale, NJ: Erlbaum.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading:

A longitudinal study in average and impaired readers. *Child Development, 57*, 988-1000.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*, 415-438.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33*, 387-407.

Wolf, M., O'Rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 15*, 43-72.

Woodcock, R. W. (1998). *Woodcock Reading Mastery Tests-Revised (WRMT-R): Forms G & H*. USA: American Guidance Service.

CHAPTER 6: APPENDICES

Appendix A: Quick Spell Test Stimuli

Real Words	Pseudowords
come	hool
dust	kile
have	meft
like	nase
look	stin
that	tath
went	thid
when	vole
will	wame
with	whis

High Frequency Bigrams	Low Frequency Bigrams
blbs	dlhw
chbt	lkwn
clnd	lswk
crds	mvhw
drgs	ncdk
glks	nwtl
slmb	tclv
splm	tmht
swls	tnws
twmn	vklh

Appendix B: List of Items in the Orthographic Choice Task

Bolded words indicated correct responses. Words in the left column were on the left side of the computer screen and words in the right column were on the right side of the computer screen.

Left	**Right**
anser	a**nswer**
boal	**bowl**
blak	**black**
buk	**book**
chooze	**choose**
cum	**come**
evry	**every**
fase	**face**
lurn	**learn**
rume	r**oom**
sleap	**sleep**
trane	**train**
wize	**wise**
wurd	**word**
yu	**you**
ghost	goast
goat	gote
green	grene
have	hav
heavy	hevvy
not	nott
please	pleese
rain	rane
red	redd
said	sed
snow	snoe
store	stoar
street	streat
take	taik
wait	wate

Appendix C: List of Items in the Wordlikeness Task

Bolded words indicated correct responses. Words in the left column were on the left side of the computer screen and words in the right column were on the right side of the computer screen.

Left	Right
bnad	**blad**
cnif	**crif**
ckader	**dacker**
ddaled	**dalled**
yb	**ib**
dlun	**lund**
moyi	**moil**
muun	**munt**
nitl	**nilt**
tolz	**tolb**
wolg	**wolt**
yikk	**yinn**
clid	cdil
fant	tanf
filk	filv
hift	hifl
ist	iit
jofy	fojy
miln	milg
moke	moje
nuck	ckun
powl	lowp
vadding	vayying
vadd	vaad

Appendix D: Word Chains Task Stimuli

hatsun

goman

toydogcar

fatrunball

dollbirdhopseehe

boypenonlookname

closeflyatwatchwide

writegivefastliveold

tenwalkeathurtwhichitfar

loseonlymuchkiduslighteight

hurtunderworkafteroldopengold

Appendix E: Letter of Information

Research Project
Reading Development from Grades 3 to 6
Letter of Information

Dear Parent or Guardian,

We are writing to ask you to allow your child to participate in a research project about children's reading, called *Reading Development from Grades 3 to 6*. We are Professors of Education and Psychology at Queen's University, Dalhousie University, and the University of Alberta, and we have been awarded a grant from the Social Sciences and Humanities research Council of Canada to carry out this research**.** This research has been cleared by the ethical review committee of Queen's University, and by your child's school board. Please read the information below. If you agree, please sign the attached sheet ("Consent Form") and return it to your child's teacher.

SOME BACKGROUND

Reading is important for success in school and in later life. In spite of the best efforts of parents and teachers, many children continue to experience difficulty learning to read. Research shows that there are several skills that are important for early reading development: of course it is important to learn letters, and we now know that the ability to process the sounds that make up words is also important in early reading development, as is the ability to name things quickly. Difficulties in these areas can lead to different types of reading difficulty. But these skills do not tell the whole story, particularly as children move from the primary to the junior grades. It is important to study other factors in the development of reading so that we can become more able to detect difficulties early, and so that we can design the best type of instruction to support children's development of reading and to avoid reading difficulties.

ABOUT THE STUDY

Our goals in this study are to trace the development of children's reading, spelling and oral language skills from grade 3 to grade 6, observe whether reading difficulties emerge and if so what types, and determine which measures in grade 3 were the best predictors of the reading success and reading difficulties in the later grades. To do this, we need to select a broad sample of children in grade 3, and ask them to perform a variety of oral language, reasoning, reading and spelling activities each year. Our work will contribute to the improvement of teacher education and curriculum development for schools.

How your child would be involved

If you agree, we will arrange with your child's teacher convenient times when we could visit the school. A trained assistant would ask your child (either alone or with other children) to come to a quiet room, for instance the library, to do the activities. Each session will only begin if your child agrees to take part, and your child may decide to end his or her participation in the study at any time. The activities will take about 120 minutes in January-March, spread over 2 or 3 sessions, depending upon the individual child. Most children find these activities interesting. We will do our best to ensure that your child enjoys the activities. If he or she loses interest, we will end the session and continue another day. The activities are like everyday school activities, though some are shown on a portable computer. We will note the answers your child gives, and in some cases record the time an answer takes. The research assistants will all have had police checks.

In grades 4, 5, and 6, at about this same time, we will again ask your permission for your child to continue participating. You may of course say no at any time.

What we promise

It is important to us to protect the privacy of people who participate in the project.

Here is what we promise:

- *Participation is voluntary, and you may withdraw from the study at any time.*
- *Lack of participation will not have any effect upon your child's grades*
- *We will end any activity if your child loses interest.*
- *No names will be used in any published work.*
- The results will not have any effect upon your child's grades.
- The data will be locked indefinitely in an office at Queen's University.
- Only the researchers and their trained assistants will have access to the data.
- There are no known risks or discomfort involved in any of these activities.

We will provide you with a report every year (in June) describing how the children in our sample have done overall, and describing your child's progress. We will share the overall description with your child's classroom teacher, but we will only provide the details about your child to you – you may show them to your child's teacher if you wish. We will publish our results in academic and professional journals, and the research assistants will use the data in their research papers or theses. In every case, results will be combined across many participants, so no individual will be identifiable.

Thank you very much for considering this request. If you have any questions about the project, please contact Dr. John Kirby at the Faculty of Education, Queen's University (533-6220). If you have any concerns about the project, you may contact Dr. Rosa Bruno-Jofré, Dean of the Faculty of Education (533-6210) or Dr. Joan Stevenson, Chair of the General Research Ethics Board at Queen's University (533-6081).

If you consent to have your child participate, please sign the attached Consent Letter and return it to your child's classroom teacher.

Sincerely,

John Kirby, Ph.D., Professor, Faculty of Education, Queen's University (for the research team)

Lesley Wade-Woolley, Ph.D., Associate Professor, Faculty of Education, Queen's University

S. Hélène Deacon, D. Phil., Assistant Professor, Dept. of Psychology, Dalhousie University

Rauno Parrila, Ph.D., Associate Professor, Dept. of Educational Psychology, University of Alberta

Appendix F: Consent Form

Research Project
Reading Development from Grades 3 to 6
Consent Form

I understand that I am being asked to allow my child to participate in a research project entitled *Reading Development from Grades 3 to 6*. I understand that the purposes of the study are to trace the development of children's reading, spelling and oral language skills from grade 3 to grade 6, observe whether reading difficulties emerge and if so what types, and determine which measures in grade 3 were the best predictors of the reading success and reading difficulties in the later grades. I have read, understood, and retained the Letter of Information, and I have had any questions answered to my satisfaction.

I understand that my child's participation in January-March 2005 will take at most 120 minutes. My child will be asked to complete a variety of oral language, reasoning, reading and spelling activities, which will only begin and continue as long as he or she is willing and interested. My child's answers will be recorded on paper or by a computer. Participation in the study and performance on the tests will have no effect upon my child's grades.

I understand that all data will be kept confidential, that no individual child will be identified in any publication, and that I can withdraw my child from the study at any time without any consequence. I understand that there are no known risks, discomforts or inconveniences involved in the study. I understand that this consent is for January-March 2005 period only.

I am aware that I can contact the researcher (Dr. John Kirby: 533-6220), the Dean of the Faculty of Education (Dr. Rosa Bruno-Jofré: 533-6210), or the Chair of Queen's University's General Research Ethics Board (Dr. Joan Stevenson: 533-6081), with any questions, concerns or complaints that I may have.

Please complete the following and return it to your child's classroom teacher as soon as possible.

_______ I **consent** that my child participate

_______ I do **NOT** consent that my child participate.

Name of child (please print): ____

Child's date of birth (day/month/year): _____________________________________

Parent's or guardian's name (please print): _________________________________

Signature: ____

Date: ______________________________

Appendix G: Descriptive Statistics of Transformed Measures

G1. Descriptive Statistics for Transformed Quick Spell Measures

Measure	Mean	SD	Skewness	SE[a]	Kurtosis	SE[b]
Grade 3						
Real Words [e]	0.79	0.29	-0.75	.18	-1.15	.36
Pseudowords [c]	2.19	0.64	-0.49	.18	-0.55	.36
LF Bigrams [d]	0.30	0.30	0.49	.18	-1.03	.36
Grade 4						
Total [c]	3.32	0.94	-0.16	.20	-0.14	.40
Real Words [e]	0.89	0.24	-1.78	.20	1.63	.40
Pseudowords [e]	0.62	0.34	0.04	.20	-1.63	.40
HF Bigrams [c]	1.82	0.68	-0.30	.20	-1.06	.40
LF Bigrams [c]	1.80	0.63	0.30	.20	-0.95	.40

[a] Standard error of skewness statistic.
[b] Standard error of kurtosis statistic.
[c] This measure underwent a square root transformation.
[d] This measure underwent a logarithmic transformation.
[e] This measure underwent an inverse transformation.

G2. Descriptive Statistics of Transformed Reading Measures

Measure	**Mean**	**SD**	**Skewness**	**SE**[a]	**Kurtosis**	**SE**[b]
Grade 3						
Block Design [c]	4.25	1.30	0.50	.18	-0.60	.35
Digit Naming [e]	0.03	0.01	0.08	.18	0.61	.36
Letter Naming [e]	0.03	0.01	-0.33	.18	0.35	.36
Orthographic Choice [d]	2.16	0.42	-0.46	.18	-0.72	.35
Wordlikeness [c]	3.65	0.94	-0.49	.18	0.09	.35
Word Chains [c]	5.62	1.28	-0.20	.18	0.52	.35
Word Attack [c]	3.55	1.12	-0.23	.18	-0.50	.35
Word Identification [c]	4.35	1.25	0.07	.18	0.30	.35
Passage Comprehension [c]	4.09	0.93	0.04	.18	1.78	.35
TOWRE [c]	4.51	1.26	-0.29	.18	1.21	.35
Grade 4						
Word Identification [c]	4.75	1.36	-0.05	.20	0.18	.40
Passage Comprehension [f]	37.64	6.91	-0.20	.20	0.15	.40
TOWRE [f]	67.22	10.13	-0.15	.20	-0.06	.40

[a] Standard error of skewness statistic.
[b] Standard error of kurtosis statistic.
[c] This measure underwent a square root transformation.
[d] This measure underwent a logarithmic transformation.
[e] This measure underwent an inverse transformation.
[f] This measure was windsorized.

Appendix H: Correlations

Measure	Grade 3																
	Predictors									Outcomes			Quick Spell				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Grade 3																	
1. Age	---	**.18**	-.02	-.04	.01	.09	-.06	**.17**	**-.15**	.01	.08	.02	.08	.10	.12	.04	.10
2. Vocabulary		---	.32	.31	.25	.43	.34	.29	.09	.44	.54	.43	.38	**.19**	.35	.37	.25
3. Block Design			---	.14	.14	.27	.20	.20	.03	.28	.29	.28	.28	.13	**.19**	.22	.28
4. PA Composite				---	.27	.33	.56	-.03	.46	.53	.53	.46	.31	.06	.30	.30	**.19**
5. NS Composite					---	.46	.51	**.17**	.29	.48	.42	.68	.54	.31	.47	.48	.39
6. OP Composite						---	.62	.78	.00	.69	.60	.67	.68	.37	.65	.62	.43
7. Word Attack							---	.00	.78	.82	.65	.71	.68	.35	.61	.64	.49
8. OP Residual								---	-.62	.23	.25	.29	.31	**.18**	.33	.26	**.15**
9. Word Attack Residual									---	.50	.36	.38	.33	**.16**	.27	.34	.29
10. Word Identification										---	.75	.82	.67	.33	.61	.64	.44
11. Passage Comprehension											---	.72	.54	.31	.51	.49	.31
12. TOWRE												---	.70	.37	.61	.65	.49
13. Quick Spell Total													---	.50	.86	.91	.80
14. Quick Spell Real Words														---	.44	.33	.27
15. Quick Spell Pseudowords															---	.71	.55
16. Quick Spell HF Bigrams																---	.71
17. Quick Spell LF Bigrams																	---

Measure	Grade 4 Outcomes 18	19	20	Quick Spell 21	22	23	24	25
Grade 3								
1. Age	.05	.15	.10	.10	.10	.06	.05	.13
2. Vocabulary	.42	.48	.37	.30	.26	.27	.28	.23
3. Block Design	.29	.28	**.20**	.27	**.21**	.22	.26	.27
4. PA Composite	.57	.41	.40	.25	.23	.25	.27	.14
5. NS Composite	.46	.40	.69	.50	.34	.42	.48	.37
6. OP Composite	.62	.54	.56	.64	.43	.53	.64	.45
7. Word Attack	.79	.57	.61	.56	.34	.48	.54	.41
8. OP Residual	.15	.22	**.21**	.36	.28	.29	.36	.24
9. Word Attack Residual	.53	.31	.35	.22	.09	**.19**	**.20**	**.17**
10. Word Identification	.87	.71	.67	.58	.34	.50	.55	.44
11. Passage Comprehension	.69	.77	.61	.48	.33	.39	.42	.39
12. TOWRE	.75	.66	.79	.63	.42	.55	.60	.48
13. Quick Spell Total	.65	.55	.61	.85	.46	.68	.80	.70
14. Quick Spell Real Words	.34	.32	.34	.42	.27	.30	.38	.32
15. Quick Spell Pseudowords	.59	.48	.49	.70	.40	.59	.64	.55
16. Quick Spell HF Bigrams	.61	.51	.58	.79	.42	.65	.79	.63
17. Quick Spell LF Bigrams	.45	.36	.45	.72	.30	.51	.66	.69
Grade 4								
18. Word Identification	---	.67	.66	.59	.36	.50	.59	.42
19. Passage Comprehension		---	.60	.49	.29	.34	.45	.42
20. TOWRE			---	.62	.40	.54	.62	.43
21. Quick Spell Total				---	.54	.76	.92	.87
22. Quick Spell Real Words					---	.39	.47	.35
23. Quick Spell Pseudowords						---	.72	.50
24. Quick Spell HF Bigrams							---	.71
25. Quick Spell LF Bigrams								---

Note: Shaded cell indicates a nonsignificant correlation. Bolded font indicates a correlation significant at the $p < .05$ level. All other correlations are significant at the $p < .01$ level.

Appendix I: Correlations between Grade 3 Measures in Factor Analysis

Measures	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1. Elision	---	$.33^{***}$	$.19^{**}$	$.26^{***}$	$.30^{***}$	$.30^{***}$	$.41^{***}$
2. Word Blending		---	$.22^{**}$	$.16^{*}$	.03	.02	.14
3. Digit Naming			---	$.79^{***}$	$.24^{**}$	$.38^{***}$	$.34^{***}$
4. Letter Naming				---	$.29^{***}$	$.45^{***}$	$.42^{***}$
5. Orthographic Choice					---	$.51^{***}$	$.50^{***}$
6. Wordlikeness						---	$.43^{***}$
7. Word Chains							---

Note: $N = 186$.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Appendix J: Hierarchical Regression Analyses with Quick Spell Test Total Predicting Reading Measures In Grade 3

J1. QST Total Predicting Reading Measures (N = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.10		-.06	-.08		-.07	-.10
Vocabulary		.39***	.20**		.48***	.35***		.37***	.16**
Block Design		.17*	.06		.17*	.10		.18*	.06
Step 2	.29***			.14***			.34***		
QST Total			.59***			.42***			.64***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J2. QST Total and Phonological Awareness Predicting Reading Measures (N = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.05		-.06	-.04		-.07	-.07
Vocabulary		.39***	.14*		.48***	.30***		.37***	.13*
Block Design		.17*	.05		.17*	.09		.18*	.06
Step 2	.15***			.11***			.10***		
PA Composite		.42***	.28***		.35***	.26***		.33***	.18**
Step 3	.20***			.09***			.27***		
QST Total			.52***			.34***			.59***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J3. QST Total and Naming Speed Predicting Reading Measures (N = 182)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.21***		
Age		-.08	-.10		-.07	-.08		-.07	-.08
Vocabulary		.39***	.19**		.48***	.34***		.38***	.14**
Block Design		.17*	.06		.17**	.10		.18**	.08
Step 2	.13***			.09***			.35***		
NS Composite		.38***	.13*		.30***	.15*		.61***	.41***
Step 3	.17***			.07***			.11***		
QST Total			.53***			.34***			.42***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J4. QST Total and Orthographic Processing Predicting Reading Measures (N = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.09		-.06	-.08		-.07	-.09
Vocabulary		.39***	.16**		.48***	.32***		.37***	.13*
Block Design		.17*	.04		.17*	.09		.18*	.05
Step 2	.28***			.14***			.26***		
OP Composite		.58***	.34***		.41***	.24**		.56***	.26***
Step 3	.07***			.03**			.11***		
QST Total			.38***			.26**			.48***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J5. QST Total and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31***			.21***		
Age		-.07	.02		-.06	-.01		-.06	-.03
Vocabulary		.39***	.14**		.48***	.32***		.37***	.13*
Block Design		.18*	.07		.18**	.10		.19**	.08
Step 2	.47***			.22***			.33***		
Word Attack		.74***	.64***		.50***	.42***		.62***	.38***
Step 3	.01*			.01			.07***		
QST Total			.15*			.13			.38***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J6. QST Total, Orthographic Processing, and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31***			.21***		
Age		-.07	.01		-.06	-.01		-.06	-.03
Vocabulary		.39***	.12**		.48***	.30***		.37***	.11*
Block Design		.18*	.06		.18**	.09		.19**	.06
Step 2	.50***			.23***			.38***		
OP Composite		.22***	.20**		.17*	.15*		.29***	.18**
Word Attack		.62***	.59***		.41***	.38***		.46***	.33***
Step 3	.00			.00			.03***		
QST Total			.06			.06			.30***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J7. QST Total and Orthographic Processing Residual Predicting Reading Measures (*N* = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.31^{***}$			$.21^{***}$		
Age		-.07	-.09		-.06	-.08		-.06	-.10
Vocabulary		$.39^{***}$	$.20^{**}$		$.48^{***}$	$.35^{***}$		$.37^{***}$	$.17^{**}$
Block Design		$.18^{*}$	.07		$.18^{**}$	.10		$.19^{**}$	.07
Step 2	.01			.01			$.02^{*}$		
OP Composite Residual		.09	-.04		.08	-.01		$.15^{*}$	.02
Step 3	$.28^{***}$			$.13^{***}$		$.41^{***}$	$.31^{***}$		
QST Total			$.60^{***}$						$.63^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

J8. QST Total and Phonological Decoding Residual Predicting Reading Measures (*N* = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.31^{***}$			$.21^{***}$		
Age		-.07	-.03		-.06	-.04		-.07	-.06
Vocabulary		$.39^{***}$	$.19^{**}$		$.48^{***}$	$.35^{***}$		$.37^{***}$	$.16^{**}$
Block Design		$.18^{*}$	.08		$.18^{**}$	.11		$.19^{**}$	.08
Step 2	$.21^{***}$			$.09^{***}$			$.11^{***}$		
Word Attack Residual		$.47^{***}$	$.32^{***}$		$.31^{***}$	$.20^{**}$		$.34^{***}$	$.16^{**}$
Step 3	$.16^{***}$			$.08^{***}$			$.24^{***}$		
QST Total			$.47^{***}$			$.34^{***}$			$.57^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix K: Hierarchical Regression Analyses with Quick Spell Test Subscales Predicting Reading Measures in Grade 3

K1. QST Real Words Predicting Reading Measures (N = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.10		-.06	-.09		-.07	-.10
Vocabulary		$.39^{***}$	$.35^{***}$		$.48^{***}$	$.45^{***}$		$.37^{***}$	$.33^{***}$
Block Design		$.17^{*}$	$.14^{*}$		$.17^{*}$	$.15^{*}$		$.18^{*}$	$.15^{*}$
Step 2	$.07^{***}$			$.05^{***}$			$.09^{***}$		
Real Words			$.26^{***}$			$.24^{***}$			$.31^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

K2. QST Pseudowords Predicting Reading Measures (N = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.11		-.06	-.09		-.07	-.11
Vocabulary		$.39^{***}$	$.23^{***}$		$.48^{***}$	$.37^{***}$		$.37^{***}$	$.21^{**}$
Block Design		$.17^{*}$	.11		$.17^{*}$	$.13^{*}$		$.18^{*}$	$.13^{*}$
Step 2	$.24^{***}$			$.13^{***}$			$.24^{***}$		
Pseudowords			$.52^{***}$			$.38^{***}$			$.53^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

K3. QST High Frequency Bigrams Predicting Reading Measures ($N = 184$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.07		-.06	-.06		-.07	-.07
Vocabulary		$.39^{***}$	$.21^{**}$		$.48^{***}$	$.37^{***}$		$.37^{***}$	$.18^{**}$
Block Design		$.17^{*}$	.10		$.17^{*}$	$.13^{*}$		$.18^{*}$	.11
Step 2	$.26^{***}$			$.11^{***}$			$.28^{***}$		
HF Bigrams			$.55^{***}$			$.35^{***}$			$.58^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

K4. QST Low Frequency Bigrams Predicting Reading Measures ($N = 184$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.11		-.06	-.08		-.07	-.11
Vocabulary		$.39^{***}$	$.32^{***}$		$.48^{***}$	$.45^{***}$		$.37^{***}$	$.30^{***}$
Block Design		$.17^{*}$	.09		$.17^{*}$	.13		$.18^{*}$	.09
Step 2	$.11^{***}$			$.04^{**}$			$.15^{***}$		
LF Bigrams			$.36^{***}$			$.20^{**}$			$.42^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Appendix L: Grade 3 Hierarchical Regression Analyses with Quick Spell Test Real Word Subscale Predicting TOWRE

L1. QST Real Words Predicting Reading Measures (N = 184)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.12*		-.06	-.10		-.07	-.12*
Vocabulary		.39***	.23***		.48***	.36***		.37***	.21**
Block Design		.17*	.11		.17*	.13*		.18*	.12*
Step 2	.24***			.13***			.24***		
HF Bigrams		.52***	.49***		.38***	.34***		.53***	.47***
Step 3	.00			.01			.01*		
Pseudowords			.08			.11			.13*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L2. QST Real Words and Phonological Awareness Predicting TOWRE (N = 184)

Predictor	ΔR^2	β^a	β^b
Step 1	.20***		
Age		-.07	-.08
Vocabulary		.37***	.16*
Block Design		.18*	.11
Step 2	.24***		
HF Bigrams		.53***	.41***
Step 3	.04**		
PA Composite		.21**	.22***
Step 4	.02*		
Pseudowords			.15*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L3. QST Real Words and Naming Speed Predicting TOWRE (N = 182)

Predictor	**ΔR^2**	β^a	β^b
Step 1	.21***		
Age		-.07	-.08
Vocabulary		.38***	.17**
Block Design		.18*	.12*
Step 2	.24***		
HF Bigrams		.53***	.29***
Step 3	.18***		
NS Composite		.48***	.47***
Step 4	.00		
Pseudowords			.07

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L4. QST Real Words and Orthographic Processing Predicting TOWRE (N = 184)

Predictor	**ΔR^2**	β^a	β^b
Step 1	.20***		
Age		-.07	-.10
Vocabulary		.37***	.16**
Block Design		.18*	.08
Step 2	.24***		
HF Bigrams		.53***	.28***
Step 3	.07***		
OP Composite		.37***	.35***
Step 4	.01		
Pseudowords			.10

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L5. QST Real Words and Phonological Decoding Predicting TOWRE ($N = 183$)

Predictor	ΔR^2	β^a	β^b
Step 1	.21***		
Age		-.06	-.02
Vocabulary		.37***	.15**
Block Design		.19**	.11*
Step 2	.23***		
HF Bigrams		.52***	.22***
Step 3	.13***		
Word Attack		.48***	.47***
Step 4	.00		
Pseudowords			.07

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L6. QST Real Words and OP Composite Residual Predicting TOWRE ($N = 183$)

Predictor	ΔR^2	β^a	β^b
Step 1	.21***		
Age		-.06	-.12*
Vocabulary		.37***	.21**
Block Design		.19**	.13*
Step 2	.23***		
HF Bigrams		.52***	.46***
Step 3	.00		
OP Composite Residual		.03	.02
Step 4	.01		
Pseudowords			.12

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

L7. QST Real Words and Phonological Decoding Residual Predicting TOWRE ($N = 183$)

Predictor	ΔR^2	β^a	β^b
Step 1	$.21^{***}$		
Age		-.06	-.07
Vocabulary		$.37^{***}$	$.20^{**}$
Block Design		$.19^{**}$	$.14^{*}$
Step 2	$.23^{***}$		
HF Bigrams		$.52^{***}$	$.41^{***}$
Step 3	$.04^{***}$		
Word Attack Residual		$.22^{***}$	$.22^{***}$
Step 4	.01		
Pseudowords			.11

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*} p < .05.$ $^{**} p < .01.$ $^{***} p < .001$

Appendix M: Grade 3 Hierarchical Regression Analyses with Quick Spell Test Pseudoword Subscale Predicting Reading Measures

M1. QST Pseudowords Predicting Reading Measures (*N* = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.09		-.06	-.08		-.07	-.09
Vocabulary		$.39^{***}$	$.18^{**}$		$.48^{***}$	$.34^{***}$		$.37^{***}$	$.16^{**}$
Block Design		$.17^{*}$	.09		$.17^{*}$	$.12^{*}$		$.18^{*}$	.10
Step 2	$.26^{***}$			$.11^{***}$			$.28^{***}$		
HF Bigrams		$.55^{***}$	$.36^{***}$		$.35^{***}$	$.17^{*}$		$.57^{***}$	$.39^{***}$
Step 3	$.04^{***}$			$.03^{**}$			$.03^{**}$		
Pseudowords			$.28^{***}$			$.26^{**}$			$.27^{**}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

M2. QST Pseudowords and Phonological Awareness Predicting Reading Measures (*N* = 184)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.04		-.06	-.04		-.07	-.05
Vocabulary		$.39^{***}$	$.13^{*}$		$.48^{***}$	$.29^{***}$		$.37^{***}$	$.13^{*}$
Block Design		$.17^{*}$	.08		$.17^{*}$	.11		$.18^{*}$	.10
Step 2	$.26^{***}$			$.11^{***}$			$.28^{***}$		
HF Bigrams		$.55^{***}$	$.33^{***}$		$.35^{***}$	.14		$.57^{***}$	$.36^{***}$
Step 3	$.07^{***}$			$.06^{***}$			$.03^{**}$		
PA Composite		$.30^{***}$	$.28^{***}$		$.28^{***}$	$.26^{***}$		$.20^{**}$	$.18^{**}$
Step 4	$.03^{**}$			$.02^{**}$			$.03^{**}$		
Pseudowords			$.24^{**}$			$.22^{**}$			$.24^{**}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

M3. QST Pseudowords and Naming Speed Predicting Reading Measures (N = 182)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.21***		
Age		-.08	-.09		-.07	-.08		-.08	-.07
Vocabulary		.39***	.17**		.48***	.33***		.38***	.14**
Block Design		.17*	.09		.17**	.12*		.18*	.10*
Step 2	.27***			.11***			.28***		
HF Bigrams		.57***	.35***		.36***	.14		.57***	.28***
Step 3	.02**			.03**			.16***		
NS Composite		.18**	.14*		.19**	.16*		.46***	.43***
Step 4	.02**			.02**			.01*		
Pseudowords			.23**			.22**			.15*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

M4. QST Pseudowords and Orthographic Processing Predicting Reading Measures (N = 184)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30***			.20***		
Age		-.07	-.08		-.06	-.08		-.07	-.08
Vocabulary		.39***	.15*		.48***	.32***		.37***	.13*
Block Design		.17*	.06		.17*	.10		.18*	.08
Step 2	.26***			.11***			.28***		
HF Bigrams		.55***	.27***		.35***	.11		.57***	.31***
Step 3	.08***			.05***			.07***		
OP Composite		.38***	.33***		.30***	.24**		.34***	.29***
Step 4	.01			.01*			.01		
Pseudowords			.15			.17*			.15

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

M5. QST Pseudowords and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.31^{***}$			$.21^{***}$		
Age		-.07	.02		-.07	-.01		-.06	-.02
Vocabulary		$.39^{***}$	$.14^{**}$		$.48^{***}$	$.31^{***}$		$.37^{***}$	$.13^{*}$
Block Design		$.18^{*}$	.08		$.18^{**}$	$.11^{*}$		$.19^{**}$	$.10^{*}$
Step 2	$.26^{***}$			$.10^{***}$			$.27^{***}$		
HF Bigrams		$.55^{***}$	.11		$.35^{***}$	.00		$.56^{***}$	$.22^{**}$
Step 3	$.23^{***}$			$.11^{***}$			$.11^{***}$		
Word Attack		$.65^{***}$	$.63^{***}$		$.46^{***}$	$.42^{***}$		$.44^{***}$	$.40^{***}$
Step 4	.00			.01			.01		
Pseudowords			.08			.13			.14

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

M6. QST Pseudowords, Orthographic Processing, and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.22^{***}$			$.31^{***}$			$.21^{***}$		
Age		-.07	.01		-.06	-.01		-.06	-.02
Vocabulary		$.39^{***}$	$.12^{*}$		$.48^{***}$	$.30^{***}$		$.37^{***}$	$.11^{*}$
Block Design		$.18^{*}$	.06		$.18^{**}$	.10		$.19^{**}$	.08
Step 2	$.26^{***}$			$.10^{***}$			$.27^{***}$		
HF Bigrams		$.55^{***}$	.07		$.35^{***}$	-.03		$.56^{***}$	$.19^{*}$
Step 3	$.25^{***}$			$.13^{***}$			$.13^{***}$		
OP Composite		$.20^{**}$	$.19^{**}$		$.17^{*}$	.15		$.22^{***}$	$.20^{**}$
Word Attack		$.59^{***}$	$.58^{***}$		$.41^{***}$	$.39^{***}$		$.37^{***}$	$.36^{***}$
Step 4	.00			.00			.00		
Pseudowords			.02			.08			.07

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

M7. QST Pseudowords and Orthographic Processing Residual Predicting Reading Measures ($N = 183$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31***			.21***		
Age		-.07	-.08		-.07	-.08		-.06	-.09
Vocabulary		.39***	.19**		.48***	.35***		.37***	.16**
Block Design		.18*	.11		.18**	.13*		.19**	.11*
Step 2	.26***			.10***			.27***		
HF Bigrams		.55***	.36***		.35***	.17*		.56***	.39***
Step 3	.00			.00			.00		
OP Composite Residual		-.01	-.05		.01	-.02		.06	.02
Step 4	.04***			.03**			.03**		
Pseudowords			.29***			.27**			.26**

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

M8. QST Pseudowords and Phonological Decoding Residual Predicting Reading Measures ($N = 183$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31***			.21***		
Age		-.07	-.03		-.06	-.04		-.06	-.05
Vocabulary		.39***	.18**		.48***	.34***		.37***	.16**
Block Design		.18*	.11*		.18**	.14*		.19**	.12*
Step 2	.26***			.10***			.27***		
HF Bigrams		.55***	.27***		.35***	.11		.56***	.34***
Step 3	.09***			.04***			.03**		
Word Attack Residual		.33***	.32***		.22***	.21**		.18**	.17**
Step 4	.03**			.03**		.24**	.03**		
Pseudowords			.25**						.25**

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Appendix N: Grade 3 Hierarchical Regression Analyses with Quick Spell Test High Frequency Bigram Subscale Predicting Reading Measures

N1. QST HF Bigrams Predicting Reading Measures ($N = 184$)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{***}$			$.20^{***}$		
Age		-.07	-.07		-.06	-.05		-.07	-.07
Vocabulary		$.39^{***}$	$.21^{**}$		$.48^{***}$	$.36^{***}$		$.37^{***}$	$.19^{**}$
Block Design		$.17^{*}$	.10		$.17^{*}$	$.14^{*}$		$.18^{*}$	.10
Step 2	$.11^{***}$			$.04^{**}$			$.15^{***}$		
LF Bigrams		$.36^{***}$	-.03		$.20^{**}$	-.08		$.42^{***}$	.06
Step 3	$.15^{***}$			$.07^{***}$			$.13^{***}$		
HF Bigrams			$.57^{***}$			$.41^{***}$			$.53^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*} p < .05$. $^{**} p < .01$. $^{***} p < .001$

N2. QST HF Bigrams and Phonological Awareness Predicting Reading Measures ($N = 184$)

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	$.21^{***}$			$.30^{**}$			$.20^{***}$		
Age		-.07	-.02		-.06	-.01		-.07	-.04
Vocabulary		$.39^{***}$	$.14^{*}$		$.48^{***}$	$.30^{***}$		$.37^{***}$	$.14^{*}$
Block Design		$.17^{*}$	.09		$.17^{*}$	$.12^{*}$		$.18^{*}$	.09
Step 2	$.11^{***}$			$.04^{**}$			$.15^{***}$		
LF Bigrams		$.36^{***}$	-.03		$.20^{**}$	-.07		$.42^{***}$	.06
Step 3	$.11^{***}$			$.09^{***}$			$.06^{***}$		
PA Composite		$.37^{***}$	$.30^{***}$		$.32^{***}$	$.28^{***}$		$.27^{***}$	$.20^{**}$
Step 4	$.11^{***}$			$.05^{***}$			$.10^{***}$		
HF Bigrams			$.50^{***}$			$.34^{***}$			$.48^{***}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*} p < .05$. $^{**} p < .01$. $^{***} p < .001$

N3. QST HF Bigrams and Naming Speed Predicting Reading Measures ($N = 182$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30**			.21***		
Age		-.08	-.06		-.07	-.05		-.07	-.06
Vocabulary		.39***	.19**		.48***	.34***		.38***	.15**
Block Design		.17*	.10		.17**	.14*		.18*	.10*
Step 2	.12***			.04**			.16***		
LF Bigrams		.37***	-.04		.21**	-.09		.42***	.01
Step 3	.07***			.06***			.23***		
NS Composite		.29***	.18**		.27***	.20**		.53***	.45***
Step 4	.11***			.05***			.05***		
HF Bigrams			.51***			.33***			.36***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

N4. QST HF Bigrams and Orthographic Processing Predicting Reading Measures ($N = 184$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.30**			.20***		
Age		-.07	-.07		-.06	-.06		-.07	-.08
Vocabulary		.39***	.15**		.48***	.32***		.37***	.14*
Block Design		.17*	.06		.17*	.11		.18*	.06
Step 2	.11***			.04**			.15***		
LF Bigrams		.36***	-.02		.20**	-.07		.42***	.07
Step 3	.19***			.10***			.15***		
OP Composite		.51***	.38***		.38***	.30***		.46***	.34***
Step 4	.05***			.02*			.04***		
HF Bigrams			.36***			.24*			.34***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

N5. QST HF Bigrams and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31**			.21***		
Age		-.07	.04		-.06	.02		-.06	.00
Vocabulary		.39***	.13**		.48***	.31***		.37***	.14*
Block Design		.18*	.10*		.18**	.13*		.19**	.10
Step 2	.11***			.03**			.15***		
LF Bigrams		.35***	-.11		.19**	-.13		.41***	.00
Step 3	.37***			.18***			.20***		
Word Attack		.74***	.66***		.53***	.47***		.55***	.44***
Step 4	.02**			.01			.03***		
HF Bigrams			.21**			.15			.29***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

N5. QST HF Bigrams, Orthographic Processing, and Phonological Decoding Predicting Reading Measures (N = 183)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31**			.21***		
Age		-.07	.03		-.06	.01		-.06	-02
Vocabulary		.39***	.11*		.48***	.30***		.37***	.12*
Block Design		.18*	.08		.18**	.12*		.19**	.08
Step 2	.11***			.03**			.15***		
LF Bigrams		.35***	-.09		.19**	-.12		.41***	.02
Step 3	.39***			.20***			.24***		
OP Composite		.22***	.19**		.18**	.16*		.27***	.22**
Word Attack		.63***	.60***		.44***	.42***		.42***	.37***
Step 4	.01*			.00			.01*		
HF Bigrams			.14*			.09			.20*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

N7. QST HF Bigrams and Orthographic Processing Residual Predicting Reading Measures ($N = 183$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31***			.21***		
Age		-.07	-.06		-.06	-.05		-.07	-.08
Vocabulary		.39***	.21**		.48***	.36***		.37***	.18**
Block Design		.17*	.11		.18**	.14*		.19**	.10
Step 2	.11***			.03**			.15***		
LF Bigrams		.35***	-.04		.19**	-.08		.41***	.06
Step 3	.00			.00			.02*		
OP Composite Residual		.07	-.01		.06	.01		.13*	.06
Step 4	.14***			.07***			.11***		
HF Bigrams			.58***			.40***			.51***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

N8. QST HF Bigrams and Phonological Decoding Residual Predicting Reading Measures ($N = 183$)

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.31**			.21***		
Age		-.07	.01		-.06	-.00		-.07	-.03
Vocabulary		.39***	.20**		.48***	.36***		.37***	.18**
Block Design		.18*	.13*		.18**	.16*		.19**	.12*
Step 2	.11***			.03**			.15***		
LF Bigrams		.35***	-.09		.19**	-.11		.41***	.02
Step 3	.14***			.06***			.05***		
Word Attack Residual		.40***	.34***		.27***	.23***		.24***	.18**
Step 4	.10***			.05***			.10***		
HF Bigrams			.49***			.35***			.48***

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix O: Hierarchical Regression Analyses with Quick Spell Test Total Predicting Reading Measures in Grade 4

O1. QST Total Predicting Reading Measures

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.11		.05	.02		.04	-.01
Vocabulary		.37***	.25***		.43***	.35***		.32***	.20**
Block Design		.20**	.11		.16*	.09		.11	.00
Step 2	.22***			.11***			.27***		
QST Total			.51***			.36***			.55***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O2. QST Total and Grade 3 Phonological Awareness Predicting Reading Measures

	Gr. 4 Word Identification			Gr. 4 Passage Comprehension			Gr. 4 TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.26***			.14***		
Age		-.06	-.04		.05	.06		.04	.03
Vocabulary		.37***	.15*		.43***	.29***		.32***	.14
Block Design		.21**	.10		.16*	.09		.11	-.01
Step 2	.21***			.08***			.11***		
Gr. 3 PA Composite		.48***	.38***		.30***	.23**		.35***	.23**
Step 3	.14***			.08***			.20***		
Gr. 4 QST Total			.41***			.31***			.50***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O3. QST Total and Grade 3 Naming Speed Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.28***			.14***		
Age		-.05	-.10		.05	.02		.04	.01
Vocabulary		.37***	.22**		.44***	.34***		.32***	.12*
Block Design		.21**	.12		.18*	.12		.11	.02
Step 2	.13***			.08***			.37***		
Gr. 3 NS Composite		.38***	.19*		.29***	.16*		.63***	.49***
Step 3	.12***			.06***			.07***		
Gr. 4 QST Total			.42***			.28***			.32***

Note: $N = 144$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 143$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O4. QST Total and Grade 3 Orthographic Processing Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.10		.05	.02		.04	-.01
Vocabulary		.37***	.20**		.43***	.30***		.32***	.16*
Block Design		.20**	.09		.16*	.07		.11	-.01
Step 2	.23***			.14***			.20***		
Gr. 3 OP Composite		.53***	.34***		.41***	.30**		.49***	.23**
Step 3	.05***			.02*			.09***		
Gr. 4 QST Total			.30***			.18*			.41***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O5. QST Total and Grade 3 Phonological Decoding Predicting Reading Measures

	Gr. 4 Word Identification			Gr. 4 Passage Comprehension			Gr. 4 TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.27***			.15***		
Age		-.05	.04		.06	.12		.05	.08
Vocabulary		.37***	.17**		.43***	.29***		.32***	.15*
Block Design		.22**	.08		.18*	.08		.13	-.01
Step 2	.46***			.21***			.29***		
Gr. 3 Word Attack		.73***	.66***		.49***	.42***		.57***	.39***
Step 3	.01*			.01			.06***		
Gr. 4 QST Total			.13*			.12			.33***

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O6. QST Total, Grade 3 Orthographic Processing, and Grade 3 Phonological Decoding Predicting Reading Measures

	Gr. 4 Word Identification			Gr. 4 Passage Comprehension			Gr. 4 TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.27***			.15***		
Age		-.05	.04		.06	.11		.05	.07
Vocabulary		.37***	.16**		.43***	.27***		.32***	.14*
Block Design		.22**	.07		.18*	.08		.13	-.01
Step 2	.47***			.22***			.31***		
Gr. 3 OP Composite		.14*	.10		.17*	.15		.21*	.08
Gr. 3 Word Attack		.65***	.63***		.39***	.37***		.46***	.36***
Step 3	.00			.00			.04**		
Gr. 4 QST Total			.09			.06			.29**

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O7. QST Total and Grade 3 Orthographic Processing Residual Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.27***			.15***		
Age		-.05	-.09		.06	.03		.05	.00
Vocabulary		.37***	.27***		.43***	.35***		.32***	.21**
Block Design		.22**	.13		.18*	.11		.13	.02
Step 2	.00			.01			.01		
Gr. 3 OP Composite Residual		.03	-.12		.08	-.02		.11	-.06
Step 3	.23***			.10***			.25***		
Gr. 4 QST Total			.54***			.36***			.56***

Note: N = 145 for analyses involving Grade 4 Word Identification and TOWRE; N = 144 for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

O8. QST Total and Grade 3 Phonological Decoding Residual Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehensio n**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.27***			.15***		
Age		-.05	.01		.06	.08		.05	.05
Vocabulary		.37***	.24***		.43***	.34***		.32***	.19**
Block Design		.22**	.11		.18*	.11		.13	.01
Step 2	.24***			.08***			.11***		
Gr. 3 Word Attack Residual		.50***	.41***		.30***	.23**		.35***	.23**
Step 3	.13***			.07***			.20***		
Gr. 4 QST Total			.40***			.30***			.49***

Note: N = 145 for analyses involving Grade 4 Word Identification and TOWRE; N = 144 for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix P: Hierarchical Regression Analyses with Quick Spell Test Subscales Predicting Reading Measures in Grade 4

P1. QST Real Words Predicting Reading Measures

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-07		.05	.04		.04	.02
Vocabulary		.37***	.32***		.43***	.40***		.32***	.26**
Block Design		.20**	.17*		.16*	.14		.11	.06
Step 2	.24**			.02			.08***		
Real Words			.24***			.13			.30***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

P2. QST Pseudowords Predicting Reading Measures

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.08		.05	.04		.04	.02
Vocabulary		.37***	.28***		.43***	.38***		.32***	.23**
Block Design		.20**	.15*		.16*	.13		.11	.04
Step 2	.15***			.04**			.19***		
Pseudowords			.40***			.21**			.46***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

P3. QST High Frequency Bigrams Predicting Reading Measures

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.07		.05	.04		.04	.03
Vocabulary		.37***	.26***		.43***	.36***		.32***	.20**
Block Design		.20**	.11		.16*	.10		.11	.01
Step 2	.22***			.09***			.27***		
HF Bigrams			.50***			.32***			.55***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

P4. QST Low Frequency Bigrams Predicting Reading Measures

	Word Identification			Passage Comprehension			TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.11		.05	.01		.04	-.01
Vocabulary		.37***	.32***		.43***	.38***		.32***	.27**
Block Design		.20**	.14		.16*	.10		.11	.04
Step 2	.10***			.08**			.11***		
LF Bigrams			.34***			.30**			.36***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix Q: Grade 4 Hierarchical Regression Analyses with Quick Spell Test Real Word Subscale Predicting TOWRE

Q1. QST Real Words Predicting Reading Measures

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.24***			.14***		
Age		-.06	-.08		.05	.04		.04	.01
Vocabulary		.37***	.26***		.43***	.37***		.32***	.20**
Block Design		.20**	.13		.16*	.12		.11	.02
Step 2	.15***			.04**			.19***		
Pseudowords		.40***	.36***		.21**	.18*		.46***	.40***
Step 3	.01			.01			.03*		
Real Words			.13			.08			.18*

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analyses involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Q2. QST Real Words and Grade 3 Phonological Awareness Predicting TOWRE ($N = 146$)

Predictor	ΔR^2	β^a	β^b
Step 1	.14***		
Age		.04	.05
Vocabulary		.32***	.14
Block Design		.11	.02
Step 2	.19***		
Gr. 4 HF Bigrams		.46***	.36***
Step 3	.06**		
Gr. 3 PA Composite		.26**	.24**
Step 4	.02*		
Gr. 4 Pseudowords			.15*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Q3. QST Real Words and Grade 3 Naming Speed Predicting TOWRE ($N = 144$)

Predictor	ΔR^2	β^a	β^b
Step 1	.14***		
Age		.04	.03
Vocabulary		.32***	.12
Block Design		.11	.04
Step 2	.20***		
Gr. 4 HF Bigrams		.47***	.25***
Step 3	.23***		
Gr. 3 NS Composite		.54***	.52***
Step 4	.00		
Gr. 4 Pseudowords			.07

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Q4. QST Real Words and Grade 3 Orthographic Processing Predicting TOWRE ($N =$ 146)

Predictor	ΔR^2	β^a	β^b
Step 1	.14***		
Age		.04	.01
Vocabulary		.32***	.15*
Block Design		.11	-.01
Step 2	.20***		
Gr. 4 HF Bigrams		.46**	.28***
Step 3	.07***		
Gr. 3 OP Composite		.34***	.31***
Step 4	.01		
Gr. 4 Pseudowords			.11

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Q5. QST Real Words and Grade 3 Phonological Decoding Predicting TOWRE ($N = 145$)

Predictor	ΔR^2	β^a	β^b
Step 1	$.15^{***}$		
Age		.05	.10
Vocabulary		$.32^{***}$	$.14^{**}$
Block Design		.13	.00
Step 2	$.19^{***}$		
Gr. 4 HF Bigrams		$.45^{***}$	$.23^{**}$
Step 3	$.15^{***}$		
Gr. 3 Word Attack		$.46^{***}$	$.43^{***}$
Step 4	.01		
Gr. 4 Pseudowords			.11

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .001$

Q6. QST Real Words, Grade 3 Orthographic Processing, and Grade 3 Phonological Decoding Predicting TOWRE ($N = 145$)

Predictor	ΔR^2	β^a	β^b
Step 1	$.15^{***}$		
Age		.05	.09
Vocabulary		$.32^{***}$	.13
Block Design		.13	-.01
Step 2	$.19^{***}$		
Gr. 4 HF Bigrams		$.45^{***}$	$.20^{**}$
Step 3	$.15^{***}$		
Gr. 3 OP Composite		.14	.11
Gr. 3 Word Attack		$.40^{***}$	$.39^{***}$
Step 4	.01		
Gr. 4 Pseudowords			.10

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .001$

Q7. QST Real Words and Grade 3 OP Composite Residual Predicting TOWRE ($N = 145$)

Predictor	ΔR^2	β^a	β^b
Step 1	.15***		
Age		.05	.03
Vocabulary		.32***	.20**
Block Design		.13	.04
Step 2	.19***		
Gr. 4 HF Bigrams		.45***	.40***
Step 3	.00		
Gr. 3 OP Composite Residual		.00	-.03
Step 4	.03*		
Pseudowords			.19*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Q8. QST Real Words and Grade 3 Phonological Decoding Residual Predicting TOWRE ($N = 145$)

Predictor	ΔR^2	β^a	β^b
Step 1	.15***		
Age		.05	.09
Vocabulary		.32***	.18*
Block Design		.13	.03
Step 2	.19***		
Gr. 4 HF Bigrams		.45***	.35***
Step 3	.07***		
Gr. 3 Word Attack Residual		.27***	.26***
Step 4	.03*		
Gr. 4 Pseudowords			.18*

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix R: Grade 4 Hierarchical Regression Analyses with Quick Spell Test High Frequency Bigram Subscale Predicting Reading Measures

R1. QST HF Bigrams Predicting Reading Measures

	Word Identification			**Passage Comprehension**			**TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.07		.05	.02		.04	.03
Vocabulary		.37***	.26***		.43***	.36***		.32***	.20**
Block Design		.20**	.11		.16*	.09		.11	.01
Step 2	.10***			.08***			.11***		
LF Bigrams		.34***	.00		.30***	.15		.36***	.03
Step 3	.12***			.02*			.16***		
HF Bigrams			.50***			.22*			.57***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

R2. QST HF Bigrams and Grade 3 Phonological Awareness Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.01		.05	.06		.04	.06
Vocabulary		.37***	.15*		.43***	.29***		.32***	.14*
Block Design		.20**	.10		.16*	.08		.11	-.00
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.34***	.01		.30***	.16		.36***	-.02
Step 3	.17***			.06***			.08***		
Gr. 3 PA Composite		.45***	.39***		.26***	.24**		.30***	.23**
Step 4	.07***			.01			.12***		
Gr. 4 HF Bigrams			.40***			.15			.51***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

R3. QST HF Bigrams and Grade 3 Naming Speed Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.28***			.14***		
Age		-.06	-.06		.05	.03		.04	.04
Vocabulary		.37***	.22**		.44***	.34***		.32***	.12*
Block Design		.21**	.13		.18*	.11		.11	.03
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.34***	-.02		.29***	.14		.36***	-.06
Step 3	.08***			.04***			.28***		
Gr. 3 NS Composite		.30***	.20**		.22**	.19*		.58***	.49***
Step 4	.08***			.01			.06***		
Gr. 4 HF Bigrams			.43***			.14			.38***

Note: $N = 144$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 143$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

R4. QST HF Bigrams and Grade 3 Orthographic Processing Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.26***			.14***		
Age		-.06	-.08		.05	.02		.04	.02
Vocabulary		.37***	.20**		.43***	.30***		.32***	.16*
Block Design		.20**	.09		.16*	.07		.11	.-.01
Step 2	.10***			.08***			.11***		
Gr. 4 LF Bigrams		.34***	-.00		.30***	.15		.36***	-.03
Step 3	.15***			.08***			.12***		
Gr. 3 OP Composite		.46***	.34***		.34**	.34***		.41***	.23**
Step 4	.03**			.00			.07***		
Gr. 4 HF Bigrams			.30**			.02			.43***

Note: $N = 146$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 145$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

R5. QST HF Bigrams and Grade 3 Phonological Decoding Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.27***			.15***		
Age		-.05	.06		.06	.11		.05	.11
Vocabulary		.37***	.17**		.43***	.30***		.32***	.15*
Block Design		.22**	.08		.18*	.08		.13	-.01
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.33***	-.07		.29***	.12		.36***	-.06
Step 3	.37***			.14***			.19***		
Gr. 3 Word Attack		.71***	.66***		.44***	.44***		.51***	.40***
Step 4	.02**			.00			.06***		
Gr. 4 HF Bigrams			.19**			-.00			.38***

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

R6. QST HF Bigrams, Grade 3 Orthographic Processing, and Grade 3 Phonological Decoding Predicting Reading Measures

	Gr. 4 Word Identification			**Gr. 4 Passage Comprehension**			**Gr. 4 TOWRE**		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.27***			.15***		
Age		-.05	.05		.06	.10		.05	.10
Vocabulary		.37***	.16**		.43***	.28***		.32***	.14*
Block Design		.22**	.08		.18*	.07		.13	-.01
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.33***	-.07		.29***	.12		.36***	-.06
Step 3	.37***			.15***			.21***		
Gr. 3 OP Composite		.13*	.08		.15	.17		.19*	.07
Gr. 3 Word Attack		.65***	.63***		.37***	.38***		.43***	.38***
Step 4	.01*			.00			.04**		
Gr. 4 HF Bigrams			.16*			-.08			.35**

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

R7. QST HF Bigrams and Grade 3 Orthographic Processing Residual Predicting Reading Measures

	Gr. 4 Word Identification			Gr. 4 Passage Comprehension			Gr. 4 TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.21***			.27***			.15***		
Age		-.05	-.04		.06	.03		.05	.04
Vocabulary		.37***	.27***		.43***	.36***		.32***	.21**
Block Design		.22**	.13		.18*	.11		.13	.02
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.32***	-.00		.29***	.17		.36***	-.02
Step 3	.00			.00			.00		
Gr. 3 OP Composite Residual		-.02	-.13		.04	-.00		.05	-.07
Step 4	.13***			.02			.15***		
Gr. 4 HF Bigrams			.53***			.19			.58***

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

R8. QST HF Bigrams and Grade 3 Phonological Decoding Residual Predicting Reading Measures

	Gr. 4 Word Identification			Gr. 4 Passage Comprehension			Gr. 4 TOWRE		
Predictor	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b	ΔR^2	β^a	β^b
Step 1	.22***			.27***			.15***		
Age		-.05	.04		.06	.09		.05	.10
Vocabulary		.37***	.24***		.43***	.35***		.32***	.19**
Block Design		.22**	.12*		.18*	.10		.13	.02
Step 2	.10***			.08**			.11***		
Gr. 4 LF Bigrams		.33***	-.05		.29***	.13		.36***	-.05
Step 3	.19***			.06***			.07***		
Gr. 3 Word Attack Residual		.46***	.43***		.25***	.24**		.29***	.25***
Step 4	.09***			.01			.13***		
Gr. 4 HF Bigrams			.43***			.16			.52***

Note: $N = 145$ for analyses involving Grade 4 Word Identification and TOWRE; $N = 144$ for analysis involving Grade 4 Passage Comprehension.
[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix S: Hierarchical Regression Analyses with Grade 3 Quick Spell Test Total Predicting Grade 4 Reading Measures

S1. Grade 3 QST Total Predicting Grade 4 Word Identification (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	$.21^{***}$		
Age		-.06	-.03
Vocabulary		$.36^{***}$	.06
Block Design		$.21^{*}$	.03
Step 2	$.55^{***}$		
Gr. 3 Word Identification		$.82^{***}$	$.73^{***}$
Step 3	$.01^{*}$		
Gr. 3 QST Total			$.15^{*}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

S2. Grade 3 QST Total Predicting Grade 4 Passage Comprehension (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	$.25^{***}$		
Age		.05	.04
Vocabulary		$.42^{***}$	.09
Block Design		$.16^{*}$	.02
Step 2	$.37^{***}$		
Gr. 3 Passage Comprehension		$.72^{***}$	$.64^{***}$
Step 3	$.01^{*}$		
Gr. 3 QST Total			$.14^{*}$

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

S3. Grade 3 QST Total Predicting Grade 4 TOWRE (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	.14***		
Age		.04	.06
Vocabulary		.32***	.03
Block Design		.11	-.05
Step 2	.49***		
Gr. 3 TOWRE		.78***	.70***
Step 3	.01		
Gr. 3 QST Total			.13

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

Appendix T: Hierarchical Regression Analyses with Grade 3 Quick Spell Test Subscales Predicting Grade 4 Reading Measures

T1. Grade 3 QST Subscales Predicting Grade 4 Word Identification (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	.21***		
Age		-.06	-.04
Vocabulary		.36***	.06
Block Design		.21*	.04
Step 2	.55***		
Gr. 3 Word Identification		.82***	.74***
Step 3	.01		
Gr. 3 QS Real Words			.05
Gr. 3 QS Pseudowords			.09
Gr. 3 QS HF Bigrams			.00
Gr. 3 QS LF Bigrams			.04

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

T2. Grade 3 QST Subscales Predicting Grade 4 Passage Comprehension (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	.25***		
Age		.05	.04
Vocabulary		.42***	.09
Block Design		.16*	.02
Step 2	.37***		
Gr. 3 Passage Comprehension		.72***	.64***
Step 3	.02		
Gr. 3 QS Real Words			.06
Gr. 3 QS Pseudowords			.03
Gr. 3 QS HF Bigrams			.12
Gr. 3 QS LF Bigrams			-.03

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$

T3. Grade 3 QST Subscales Predicting Grade 4 TOWRE (N = 143)

Predictor	ΔR^2	β^a	β^b
Step 1	.14***		
Age		.04	.07
Vocabulary		.32***	.04
Block Design		.11	-.05
Step 2	.49***		
Gr. 3 TOWRE		.78***	.71***
Step 3	.01		
Gr. 3 QS Real Words			.01
Gr. 3 QS Pseudowords			-.06
Gr. 3 QS HF Bigrams			.13
Gr. 3 QS LF Bigrams			.04

[a] Standardized beta coefficient for the step at which the predictor first entered the model.
[b] Standardized beta coefficient for the final step of the model.
* $p < .05$. ** $p < .01$. *** $p < .001$