Does How We Measure Vocabulary Breadth Influence its Relationship to Reading Comprehension in Young Elementary School Students?

by

Olivia Innes

Submitted in partial fulfilment of the requirements
for the degree of Master of Arts in School Psychology

at

Mount Saint Vincent University
Halifax, Nova Scotia
October 2022

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

A large body of research exists supporting the relationship between vocabulary and reading comprehension, and for younger children between vocabulary and phonological awareness. Vocabulary is frequently measured using the Peabody Picture Vocabulary Test (PPVT), and several recent studies have used shortened versions of the PPVT in explorations of oral language contributions to reading comprehension. This study examined whether these shortened versions of the PPVT are as strongly predictive of reading comprehension (in a sample of first and second grade students; $N = 62$) and of phonological awareness (in a sample of kindergarten and first grade students; $N = 45$) as the full test. A second focus of this study was on the relationship between children's reaction times to correctly identifying items on this vocabulary measure and reading comprehension, as it has been suggested that speed of recognition may account for unique variance in children's understanding of texts. It was found that the two shortened forms and the full version of the PPVT-5 were not differentially predictive of phonological awareness (in kindergarten and first grade students) or of reading comprehension in first grade students. However, for the second-grade students, one shortened version (every fourth item included) was not as strongly related to reading comprehension as the full version of the PPVT. Turning to the second focus of this study, reaction times to recognize vocabulary items did not explain unique variance in reading comprehension in equations which accounted for the variance explained by grade, word reading, and vocabulary breadth. Overall, the findings from this study raise questions about substituting shortened versions of the PPVT in examinations of the relationship between vocabulary and other reading measures. Furthermore, individual differences for recognizing vocabulary items may not be meaningfully related to reading comprehension in young first and second grade students; a question to be further explored in future research.

Keywords: vocabulary, reading comprehension, vocabulary reaction time, phonological awareness, PPVT, early elementary

Introduction

Comprehending written text is a skill that is essential to performing many activities of daily living. When children struggle to understand what they read, it becomes very challenging for them to keep up with their peers at school, as demands increase and they are required to learn more and more information through written mediums. Research over decades has investigated factors that influence children's reading comprehension in the hopes of building it. One factor that has considerable research supporting its relationship to reading comprehension is vocabulary knowledge (Farvardin & Koosha, 2011; Gottardo et al., 2018; Ouellette, 2006; Ricketts et al., 2007; Tannenbaum et al., 2006). It has been found that students who know more words have better reading comprehension compared to children who know fewer words (Nation & Snowling, 2004). Building on these and similar findings, researchers, clinicians, and teachers alike have sought to determine ways to increase children's vocabulary knowledge. In order to determine the relative size of a child's receptive vocabulary, researchers and clinicians have used a variety of different measures, but one of the most commonly used measures is the Peabody Picture Vocabulary Test. Knowing the relative size of a child's vocabulary can be helpful in order for researchers and clinicians to determine what interventions might be put into place (if any) to promote children's vocabulary development, which in turn should increase their reading comprehension. The current study aims to further explore the relationship between vocabulary breadth, measured using different instantiations of the Peabody Picture Vocabulary Test, and reading comprehension or reading related skills in kindergarten, first, and second grade students.

Vocabulary Breadth and Depth

Vocabulary knowledge has generally been separated along two dimensions, breadth versus depth and receptive versus expressive (Cervatiuc, 2007; Meara & Wolter, 2004; Read,

2004). Vocabulary breadth (also called vocabulary size) can be defined as the number of words that an individual knows, while vocabulary depth is typically defined as the quality of understanding of a word or how well a word is known (Cervatiuc, 2007; Meara & Wolter, 2004; Read, 2004). Depth of vocabulary knowledge includes three features: precision of meaning, comprehensive word knowledge, and network knowledge (Read, 2004). Precision of meaning involves having a detailed and specific knowledge of the meaning of a word (Read, 2004). Comprehensive word knowledge involves knowledge of a word's semantic, orthographic, phonological, morphological, syntactic, and pragmatic characteristics (Read, 2004). Finally, network knowledge involves incorporating a word into a network in the mental lexicon as well as the ability to associate it with and differentiate it from related words (Read, 2004). By nature of its definition, vocabulary breadth is an easier construct to measure than vocabulary depth, and as such conventional and frequently used vocabulary tests usually measure breadth (Meara & Wolter, 2004; Read, 2004). Along the other dimension, receptive knowledge involves the ability to perceive a word and retrieve its meaning when listening or reading, while productive knowledge involves the ability to use a word in speech or writing (Cervatiuc, 2007).

Typical Sequence of Vocabulary Development in the Elementary Years

Research has shown that on average, children acquire about 860 root word meanings per year from age one through age eight (end of Grade 2), which equates to approximately 2.4 root words per day (Biemiller, 2006). This results in the acquisition of about 6000 root word meanings by the end of Grade 2 and 9000 root words by the end of the elementary years (Biemiller, 2003; Biemiller, 2006). It has been found that there are a lot of commonalities in the words that are acquired across the majority of children, and that most children can learn and retain a maximum average of three words per day (Biemiller, 2003).

For children under 3 years of age, word meanings are largely learned through the direct pairing of the novel word and a physical object, action, or modifier by a more advanced speaker of the language (Biemiller, 2006). Once children reach about 3 years of age, word meanings are acquired with verbal referents in addition to concrete referents (Biemiller, 2006). In the majority of cases, the acquisition of root word meanings requires explicit instruction (overt explanation; Biemiller, 2006). When root words are known, derived words (prefixed or suffixed) and compound words are likely to be inferred largely from context (Biemiller, 2006).

Factors Influencing Vocabulary Size and Vocabulary Acquisition

It has been found that most individual differences in vocabulary knowledge develop before Grade 3, when there is a lot of variability in the rate of word acquisition (Biemiller & Slonim, 2001). A significant proportion of the variability in vocabulary size between children is related to opportunity rather than ability (Biemiller, 2006). This has been demonstrated in a number of studies that have found that by the age of four, the size of a child's vocabulary is largely determined by the number and variety of words spoken by the child's parents (Hart & Risley, 1995; 2003). For example, Hart and Risley (2003) found that 86 to 98% of the words recorded in 42 children's vocabularies over four observation periods when they were 33-36 months in age were also recorded in their parents' vocabularies. They also found that the children's vocabulary size was largely influenced by the vocabulary size of their parents. Additional research with school-aged children has provided support for the influence of the home literacy environment on children's vocabulary. A study by Sénéchal et al. (1998) found that storybook exposure at home explained statistically significant unique variance in kindergarten and Grade 1 children's oral language skills (including their receptive vocabulary), and another study by Sénéchal and LeFevre (2002) found that storybook exposure (at home) in

kindergarten explained a significant 9% of unique variance in children's receptive language (including vocabulary) at the beginning of Grade 1. These findings with preschool- and school-aged children support the claim that the home environment is an important determinant of vocabulary size by the end of Grade 2, which is considered to be the end of the "preliterate" period for the majority of children (Biemiller, 2006).

Consistent with these findings, limited evidence has been shown for school factors affecting vocabulary growth in the early grades, with some research showing that a year of schooling has a negligible effect on vocabulary growth. Morrison et al. (1998) compared the vocabulary sizes of January-born first grade children and December-born second grade children, who are only 1 month apart in age. They found that the vocabulary size of the January-born first grade children was approximately equal to the vocabulary size of the December-born second grade children, even though the December-born children had received one additional year of schooling.

As a result of the differences in children's exposure to vocabulary in their home environments and the findings that schooling has largely had a limited impact on children's vocabulary development in the early elementary years, significant differences in vocabulary size have been found between children by the end of Grade 2 (Biemiller, 2003; Hart & Risley, 2003; Morrison et al., 1998). In one study by Biemiller and Slonim (2001), children in the highest quartile had an average vocabulary size of 7100 root words (equating to the acquisition of over three root words per day since the age of one year), while children in the lowest quartile had an average vocabulary size of 3000 words (equating to the acquisition of one root word per day since the age of one year). This means that children in the lowest quartile were about 2000 words behind average children and about 4000 words behind children in the highest quartile (Biemiller,

2003; Biemiller & Slonim, 2001). By Grade 5, the children in the lowest quartile had still not attained the Grade 2 vocabulary size of the children in the highest quartile (Biemiller, 2003; Biemiller & Slonim, 2001). For the lowest quartile children to catch up with the average children, they would need to learn 3.5 to four root words per day, at which rate it would take them five to six years to catch up to average children (Biemiller, 2003). This is concerning for many reasons, one of them being that the impact of a limited vocabulary on reading comprehension starts to become apparent in the early grades, and it is well known that children who struggle to understand what they read fall behind their peers in the early and later elementary years when reading is required to learn (Biemiller, 2003). Findings from the previously mentioned study by Sénéchal and LeFevre (2002) support this claim, where receptive language skills (including vocabulary) measured at the beginning of Grade 1 accounted for statistically significant unique variance in reading (including comprehension) measured at the end of Grade 3.

The National Institute for Literacy and National Institute of Child Health and Human Development (National Institute for Literacy, 2008; National Institute of Child Health and Human Development, 2000) have suggested that one way to mitigate the negative effects of limited exposure to vocabulary in the home environment and promote vocabulary growth at school is to explicitly teach word meanings to students, and a large number of studies to date have investigated the effects of vocabulary interventions on children's word learning (e.g., Beck & McKeown, 2007; Coyne et al., 2007; Wasik et al., 2006).

Marulis and Neuman (2010) conducted a meta-analysis of 67 studies examining the impact of vocabulary interventions on children's vocabulary development in pre-kindergarten and kindergarten, and they found that vocabulary interventions resulted in an overall increase of

close to one standard deviation on vocabulary measures (average effect size of .87 for standardized and author-created measures combined; Marulis & Neuman, 2010). The authors found larger effects for interventions that included both explicit (providing detailed definitions and examples as well as a follow-up discussion of the words) and implicit (teaching words in the context of an activity) instructional methods, interventions that were provided by trained adults, and measures that were author-created as opposed to standardized. The authors also found that the vocabulary gains of middle- and upper-income at-risk children were significantly larger than those of low-income at-risk children. The authors defined at-risk as those with English second language status, low academic achievement, and/or special needs. This means that although all children showed gains in vocabulary following the interventions, the gains achieved by the low-income and/or at-risk children were not large enough to close the gap in vocabulary size between low-income and higher-income or at-risk and not at-risk children (Marulis & Neuman, 2010).

The disparity in vocabulary knowledge between children of low socioeconomic status (SES) and children of high SES has been well established by research (Hoff, 2003; Moats, 1999). For example, a study by Hoff (2003) found that two-year-old children from high-SES households showed greater gains in their productive vocabularies over a 10-week period compared to two-year-old children from mid-SES households (Hoff, 2003). Furthermore, Moats (1999) estimated that upon school entry, the difference in vocabulary size between children from advantaged and disadvantaged backgrounds is about 15000 words (including derived words), with children from more disadvantaged backgrounds having a vocabulary of approximately 5000 words vs. approximately 20000 words for children from advantaged backgrounds (Moats, 1999).

Biemiller and Slonim (2001) examined root-word vocabulary growth in a normative sample of children with a wide socioeconomic range and in a sample of children from

advantaged backgrounds (Biemiller & Slonim, 2001). A root word is a basic word with no prefixes or suffixes attached to it (McEwan, 2008). Children from all elementary grade levels participated in the study (kindergarten to Grade 6). Similar patterns in root word vocabulary growth were found in the two samples, with a very large gain between Grade 1 and Grade 2 and slower growth between the end of Grade 2 and Grade 6 (Biemiller & Slonim, 2001). The authors found that children from the normative sample acquired root words at an average rate of 2.2 words per day from age one until the end of Grade 2 and 2.9 words per day from Grade 3 to Grade 5 (Biemiller & Slonim, 2001). This resulted in the acquisition of 5200 root words by the end of Grade 2 and 3200 additional root words between Grades 3 and 5 (Biemiller & Slonim, 2001). Meanwhile, the advantaged sample acquired root words at an average rate of 2.4 words per day from age one until the end of Grade 2 and 2.3 words per day from Grade 3 to Grade 5 (Biemiller & Slonim, 2001). This resulted in the acquisition of 6200 root words by the end of Grade 2 and 2500 additional root words between Grades 3 and 5 (Biemiller & Slonim, 2001). These results can help to explain the finding that the advantaged sample had a 20% larger root word vocabulary at the end of Grade 2 but only a 3% larger vocabulary by the end of Grade 5, since the gain in root words per day between Grades 3 and 5 for the normative sample was higher than that of the advantaged sample (Biemiller & Slonim, 2001). In conclusion, these findings show that although children from advantaged and less advantaged backgrounds show similar rates of root word vocabulary growth beyond the second grade of the elementary years, the difference in vocabulary size between the two groups remains.

The Relationship Between Vocabulary and Reading Comprehension

One important component of academic performance related to vocabulary is reading comprehension. Decades of research has shown that there exists a significant, positive

relationship between vocabulary and reading comprehension (e.g., Binder et al., 2017; Chen, 2011; Nation & Snowling, 2004; Proctor et al., 2006). The strength of this relationship has been further supported by the considerable variability across the studies that comprise the evidence base, in terms of the age range of participants, the countries of origin of the participants, the first languages of the participants, the dimensions of vocabulary measured, and the different measures used to assess vocabulary and reading comprehension. However, it should be noted that some studies with younger students have not found this significant relationship between vocabulary and reading comprehension. For example, Fricke et al. (2016) did not find significant correlations between kindergarten vocabulary and Grades 1 and 2 reading comprehension in a sample of German-speaking children. Additionally, both Ouellette and Beers (2010) and Sparks and Metsala (2021) found that after controlling for other reading and oral language skills, vocabulary did not predict reading comprehension in Grade 1 students. It can be difficult to measure reading comprehension in young students as word reading is not yet easy and fluent and may account for the majority of individual differences in reading comprehension (for discussion see Campbell, 2021).

Vocabulary Breadth, Vocabulary Depth, and Reading Comprehension in English as a First Language Speakers

A large body of research has examined the relationship between vocabulary (breadth and depth) and reading comprehension in English as a first language speakers, and some of these studies are discussed next. Both Ouellette (2006) and Ricketts et al. (2007) examined the relationship between vocabulary and reading comprehension in Grade 3 and 4 students after controlling for word reading skills. Ouellette further separated vocabulary into the two dimensions of breadth and depth and used receptive and expressive vocabulary tasks to measure

vocabulary breadth and word definitions and synonym tasks to measure vocabulary depth (Ouellette, 2006). He found that before controlling for pseudoword reading and irregular word reading, both vocabulary breadth and depth accounted for significant variance in reading comprehension. However, after controlling for these word reading skills, depth of vocabulary knowledge accounted for a significant amount of the variance in reading comprehension, while breadth of vocabulary knowledge did not (Ouellette, 2006). On the other hand, Ricketts et al. (2007) found that after controlling for pseudoword, regular word, and exception word reading, vocabulary accounted for a significant amount of the variance in reading comprehension (17.8%). Although it is difficult to directly compare the two studies since they measured vocabulary differently, the findings from both studies provide support for the relationship between vocabulary and reading comprehension. Furthermore, in the Ricketts et al. (2007) study, vocabulary was measured using a task that required the students to verbally define words, with the quality of their responses being rated on a two-point scale. This method of measuring vocabulary taps more into an individual's deeper understanding of a word and its usage (context-based understanding of a word), or their vocabulary depth. Therefore, Ouellette's finding that only depth of vocabulary knowledge (not breadth) accounted for significant variance in reading comprehension after controlling for pseudoword reading and irregular word reading is consistent with Ricketts et al.'s findings.

Tannenbaum et al. (2006) went one step further and examined three dimensions of vocabulary (breadth, depth, and fluency) in a sample of Grade 3 students. They defined fluency as the rate at which an individual accesses the meaning of a word. Two measures were used to assess each of the three dimensions of word knowledge (Tannenbaum et al., 2006). The first fluency measure gave participants one minute to use as many target words as possible in

sentences, and the second measure required participants to name as many items as possible from eight categories, with ten seconds allowed for responding to each category. Depth of word knowledge and fluency of word knowledge were combined into one construct, as confirmatory factor analysis and structural equation modeling showed that a two-factor model of breadth and depth/fluency provided the best fit to the data. Findings showed that both breadth and depth/fluency accounted for a significant amount of the variance in reading comprehension, but that breadth explained more variance in reading comprehension (19%) than did depth/fluency (2%; Tannenbaum et al., 2006). There are a few possible explanations for the differences between Tannenbaum et al.'s (2006) findings and those of Ouellette (2006) and Ricketts et al. (2007). First of all, the latter two authors controlled for other reading skills (such as pseudoword reading, regular word reading, and exception word reading) in their examination of the relationship between vocabulary and reading comprehension. In addition, Tannenbaum et al. (2006) combined depth and fluency into one construct, which could have also contributed to the differences in their findings. Despite the differences between Tannenbaum et al.'s (2006) findings and those of Ouellette (2006) and Ricketts et al. (2007), other researchers have also found greater contributions of vocabulary breadth to reading comprehension. For example, Binder et al. (2017) found that both vocabulary breadth and vocabulary depth explained a significant proportion of the variance in reading comprehension in college-aged students, although vocabulary breadth accounted for a larger proportion of the variance than did vocabulary depth (Binder et al., 2017).

Cross-sectional and longitudinal studies comparing the relationship between vocabulary and reading comprehension across younger and older students have yielded conflicting results. Ouellette and Beers (2010) examined the relationship between vocabulary breadth, vocabulary

depth, and reading comprehension in Grade 1 and Grade 6 students after controlling for other reading and oral language skills. They found that after controlling for phonological awareness, decoding, irregular word reading, and listening comprehension, neither vocabulary breadth nor vocabulary depth explained significant variance in reading comprehension in the Grade 1 students. For the Grade 6 students, they found that after controlling for the same variables, vocabulary breadth explained a significant amount of the variance in reading comprehension (15.3%), while vocabulary depth did not contribute additional unique variance to reading comprehension. In contrast, Nation and Snowling (2004) conducted a longitudinal study on the contribution of vocabulary to reading comprehension and examined whether oral language skills measured when students were 8.5 years old (time 1) predicted reading comprehension 4.5 years later when students were 13 years old (time 2). Results indicated that vocabulary measured at time 1 accounted for a significant amount of the variance in concurrent reading comprehension after controlling for age, nonword reading, and phonological skills (% change in $R^2 = 25.2$). At time 2, vocabulary measured at time 1 accounted for a significant amount of the variance in reading comprehension, even after controlling for age, reading comprehension at time 1, nonword reading, and phonological skills (% change in $R^2 = 4.9$; Nation & Snowling, 2004). In summary, vocabulary accounted for gains in reading comprehension across this time, above and beyond word reading related skills. The time 1 findings reported by Nation and Snowling are consistent with those reported by Ouellette (2006) and Ricketts et al. (2007) with children of similar ages. Nation and Snowling's (2004) findings show the importance of vocabulary to students making gains in reading comprehension. Furthermore, other researchers have shown that the contribution of vocabulary to reading comprehension increases as children get older (Storch & Whitehurst, 2002). Ouellette and Beers' (2010) findings that vocabulary did not

explain significant variance in reading comprehension in Grade 1 students but did explain significant variance in reading comprehension in Grade 6 students are in line with the findings reported by Storch and Whitehurst (2002). Therefore, it can be concluded that vocabulary is an important contributor to gains in reading comprehension, and its impact on reading comprehension increases over time.

The research summarized above indicates that while there exists a considerable amount of evidence supporting a significant relationship between vocabulary and reading comprehension, there also exists conflicting findings regarding the relative contribution of different dimensions of vocabulary to reading comprehension and age-related changes in the contribution of vocabulary to reading comprehension (Ouellette, 2006; Ouellette & Beers, 2010; Ricketts et al., 2007; Tannenbaum et al., 2006). While some studies have found a larger contribution of vocabulary breadth to reading comprehension, others have found a larger contribution of vocabulary depth to reading comprehension (Binder et al., 2017; Ouellette, 2006). Likewise, studies examining children of similar ages have revealed inconsistent findings regarding the relationship between vocabulary and reading comprehension (Ouellette, 2006; Ricketts et al., 2007; Tannenbaum et al., 2006). Therefore, the current study aims to further explore the relationship between vocabulary breadth and reading comprehension specifically in Grade 1 and 2 students, a population that has not been extensively studied in previous research.

Vocabulary Breadth, Vocabulary Depth, and Reading Comprehension in English Language Learners

Although not examined in this study, it is important to take into account findings for students who are learning English as an additional language. Consistent with findings from the studies with English first language speakers described above, research with English language

learners has also found a significant correlation between vocabulary and reading comprehension. For example, in Farvardin and Koosha's study investigating the role of vocabulary knowledge in Iranian English as an additional language university students' reading performance, the authors found that both vocabulary breadth and vocabulary depth accounted for a significant amount of the variance in reading comprehension (Farvardin & Koosha, 2011). Additionally, vocabulary breadth, vocabulary depth, and reading comprehension were all significantly correlated with one another, with the correlation between vocabulary breadth and vocabulary depth being the strongest, followed by vocabulary breadth and reading comprehension. These results are in line with Binder et al.'s findings with native English-speaking college students (Binder et al., 2017). Similarly, Kang et al. (2012) found that vocabulary knowledge accounted for a significant amount of the variance in reading comprehension in Korean high school students. When vocabulary breadth and depth were examined separately, they were both found to be significantly correlated with reading comprehension (Kang et al., 2012). When participants were divided into three groups based on their level of English proficiency (Beginning, Intermediate, and Advanced), both vocabulary breadth and vocabulary depth were significantly correlated with reading comprehension in all three groups (Kang et al., 2012).

An important contribution of some of the research on English language learners is that it has examined the relationship between vocabulary and different components of reading comprehension measures as opposed to overall reading comprehension more generally. For example, Nouri and Zerhouni (2016) examined the relationship between vocabulary size (or breadth), vocabulary depth, and different aspects of reading comprehension in English as an additional language students from Morocco. Findings indicated that vocabulary size was significantly correlated with reading comprehension performance overall and was also

significantly correlated with the recall question on the reading comprehension measure. However, vocabulary size was not significantly correlated with the multiple-choice questions on the reading comprehension measure (Nouri & Zerhouni, 2016). On the other hand, vocabulary depth was found to be significantly correlated with reading comprehension performance and was significantly correlated with both the recall questions and the multiple-choice questions on the reading comprehension measure (Nouri & Zerhouni, 2016). Li and Kirby (2015) also examined the relationship between different dimensions of vocabulary and different aspects of reading comprehension, but with Chinese students. Two tasks were used to measure reading comprehension, the first being passages followed by multiple-choice comprehension questions (to measure general reading comprehension) and the second being summary writing (to measure depth of text understanding). The authors found that vocabulary breadth, but not vocabulary depth, explained a significant amount of the variance in the general reading comprehension measure after controlling for Chinese reading comprehension and English word reading. Conversely, vocabulary depth, but not vocabulary breadth, explained a significant amount of the variance in the summary writing measure after controlling for Chinese reading comprehension and English word reading (Li & Kirby, 2015). These findings give us more information on the relationship between the different dimensions of vocabulary (breadth and depth) and different aspects or measures of reading comprehension, which helps to further tease apart the nature and complexity of the relationship between vocabulary and reading comprehension.

Qualitative information gathered in a study by Chen (2011) also provided valuable insights regarding students' perceptions of the relationship between vocabulary and reading comprehension. When students were asked to explain vocabulary breadth, all eight of the interviewees said that it was the number of words they knew, while six out of the eight

interviewees said that it was related to the surface meaning of words, and six interviewees said that it was related to words that are easy to understand and memorize (Chen, 2011). Students were also asked to explain how vocabulary breadth related to their reading ability. Those students with larger English vocabularies felt that reading was easy and wanted to increase their depth of vocabulary knowledge, while students with more limited English vocabularies tended to struggle with reading. In addition, students with lower levels of English proficiency reported that their limited vocabulary breadth was their main concern related to reading comprehension, whereas students with higher levels of English proficiency said that their limited depth of vocabulary knowledge was their main concern related to reading comprehension (Chen, 2011). Overall, interviewees agreed that vocabulary breadth was important in helping them to understand written text. These findings may also be applicable to younger English-speaking students learning how to read, and the information may be helpful for teachers of the early and middle elementary grades. By knowing that students with larger vocabularies tend to have more success with reading and want to increase the depth of their vocabulary knowledge in order to further promote their reading comprehension, research could investigate the effectiveness of vocabulary interventions that first target the development of vocabulary breadth, followed by vocabulary depth.

Research with Spanish-English bilingual students has allowed researchers to present a more comprehensive view of English reading comprehension that considers the development of both first language and second language literacy skills as well as the potential effects of language of initial literacy instruction (Proctor et al., 2006). In one such study, Proctor et al. (2006) explored the effects of Spanish language alphabetic knowledge, fluency, vocabulary knowledge, and listening comprehension on English reading comprehension after controlling for English oral

language proficiency, decoding skills, and language of initial literacy instruction in Grade 4 Spanish-English bilingual students. The majority of students were first taught to read in Spanish (69%), with the remainder receiving initial literacy instruction in English (31%). Before controlling for the language of initial literacy instruction, the authors did not find a significant correlation between Spanish vocabulary knowledge and English reading comprehension (Proctor et al., 2006). However, after controlling for language of initial literacy instruction and English reading skills (alphabetic knowledge, fluency, vocabulary knowledge, and listening comprehension), results showed that Spanish vocabulary knowledge accounted for a significant amount of the variance in English reading comprehension. This finding provides some evidence of literary skills transferring from the first language to the second language (Proctor et al., 2006). Additional research with Spanish-English bilingual students examining the relationship between English vocabulary knowledge and English reading comprehension has revealed similar findings. For example, Gottardo et al. (2018) found that vocabulary explained significant unique variance in reading comprehension after controlling for word reading in students between nine and 13 years of age.

In summary, findings from studies with English language learners have made important contributions to the evidence base on the relationship between vocabulary and reading comprehension. Of particular relevance to the current study, research with English language learners is consistent with research with English as a first language speakers that has found a significant, positive relationship between vocabulary breadth and reading comprehension (Farvardin & Koosha, 2011). This suggests that the finding of this association with English first language speakers generalizes to English language learners. Moreover, research with English language learners has provided us with additional information on the relationship between

different dimensions of vocabulary (breadth and depth) and different aspects or measures of reading comprehension, which could inform the interpretation of the findings in the current study (Li & Kirby, 2015; Nouri & Zerhouni, 2016). For example, the current study focuses only on vocabulary breadth and a timing component related to this. Future research assessing this timing component in relation to vocabulary depth and how well it predicts reading comprehension will help to form a complete picture.

Measures of English Vocabulary Breadth

The current study examines one frequently used measure of vocabulary breadth, the Peabody Picture Vocabulary Test (PPVT). In this section, I discuss the PPVT in addition to some other measures of vocabulary breadth, including the Vocabulary Levels Test, the Vocabulary Size Test, and the Yes/No Vocabulary Size Test.

The Peabody Picture Vocabulary Test (PPVT) was published in 1959 by Lloyd M. Dunn and Leota M. Dunn and was originally developed as a measure of verbal intelligence (Bochner, 1978; Goriot et al., 2021). The PPVT is an individually administered test of receptive vocabulary in which the examinee is required to point to or verbally indicate the picture (from an array of four) that they think best illustrates the meaning of the stimulus word presented orally by the examiner. The current edition of the test consists of 228 items grouped into 19 sets of 12 items each (Dunn & Dunn, 2018). The items (words) include verbs, adjectives, and nouns that are presented in order of decreasing frequency and belong to one of 20 different content categories (Dunn & Dunn, 2018). Although the PPVT was not developed for English language learners, it has been widely used with English language learners in addition to monolingual English speakers and bilingual children and has also been used in early foreign-language education settings (Goriot et al., 2021).

Some other commonly used measures of receptive vocabulary breadth include the Vocabulary Levels Test, Vocabulary Size Test, and Yes/No Vocabulary Size Test. Despite the fact that these measures were designed for English language learners, they share many elements in common with the PPVT and have been found to be reliable and valid measures of receptive vocabulary. The first commonality between all three of these measures and the PPVT is that the items (or words) are presented in order of decreasing frequency (Cameron, 2002; Coxhead et al., 2015; Karami, 2012; Meara & Buxton, 1987; Stewart, 2014; Webb et al., 2017). Additionally, both the PPVT and Vocabulary Levels Test include words from three classes: nouns, verbs, and adjectives (Kremmel & Schmitt, 2018; Webb et al., 2017). All four measures are considered to be measures of decontextualized receptive vocabulary knowledge, since the words are presented in isolation as opposed to in the context of a passage (Cameron, 2002; Coxhead et al., 2015; Meara & Buxton, 1987; Nation & Coxhead, 2014; Stewart, 2014; Webb et al., 2017). The PPVT, Vocabulary Levels Test, and Vocabulary Size Test use a multiple-choice format to test examinees' vocabulary knowledge, although the PPVT is the only measure that uses pictures to illustrate the meaning of the stimulus word (as opposed to written definitions; Stewart, 2014; Webb et al., 2017).

Recently, another version of the Vocabulary Size Test was created so that the vocabulary size of young native English speakers could be tested (Coxhead et al., 2015; Nation & Coxhead, 2014). Coxhead et al. (2015) conducted one study that examined the use of this version with native English speakers between the ages of 13 and 18. They found an increase in vocabulary size with age, and they also found that age was a more significant predictor of test score than grade (Coxhead et al., 2015). These findings are consistent with research showing that the vocabulary sizes of January-born first grade children and December-born second grade children

(who are only 1 month apart in age) are about the same, even though these two groups of children differ by one year in schooling (Morrison et al., 1998). These findings also provide support for the use of age-based start points as opposed to grade-based start points on the PPVT.

Additional studies using the Vocabulary Size Test with native English speakers have provided further support for the validity and use of the PPVT with this population. Duff and Brydon (2020) used an adapted version of the Vocabulary Size Test with English-speaking children between the ages of 7 and 12 years, in which only half the items were administered (the first five items per frequency bin as opposed to the 10 in the original version). The authors found that participants' scores on the adapted version of the Vocabulary Size Test (which provides an estimate of the number of word families known) were significantly related to their raw scores on the PPVT-4, indicating that the two tests were measuring the same construct. Duff and Brydon (2020) concluded that because of the strong relationship between the scores on the adapted Vocabulary Size Test and the raw scores on the PPVT-4, normative information about the PPVT-4 could be used to generate vocabulary size estimates for children of different ages and vocabulary skill levels.

Validity of Shortened Forms of the Peabody Picture Vocabulary Test

The PPVT is one of the most commonly used measures of vocabulary breadth with English first language speakers and has been used extensively in the research examining the relationship between vocabulary and reading comprehension. Researchers have sometimes constructed different forms of the test so that they have the option of choosing a format that best suits the needs of their study, and in particular time constraints of testing participants. There has been extensive research on the reliability and validity of the full form of the PPVT, with the most recent version (PPVT-5) showing good to excellent reliabilities and validity (Dunn & Dunn,

2018). There has also been preliminary research supporting the validity of researcher derived short forms of the PPVT (Deacon et al., 2013).

Several recent research studies have used researcher created, short forms of the PPVT to save time in lengthy research batteries (Deacon et al., 2013; Deacon et al., 2014; Metsala et al., 2021; Sparks & Metsala, 2021). One of these studies examined the relationship between children's morphological awareness and reading accuracy in a longitudinal study, while controlling for vocabulary, phonological awareness, and nonverbal ability (Deacon et al., 2013). In Grade 2, receptive vocabulary was measured using a modified version of the PPVT-III in which only every fourth item was administered. This modified version maintained the progression of item difficulty, but the researchers reported it reduced testing duration to minimize missed class time (Deacon et al., 2013). The same group of children had completed the full PPVT-III in Grade 1. The researchers evaluated the "developmental sensitivity" of the subset of items administered in the modified version in Grade 2 by calculating the participants' scores on the same subset of items administered within the full version in Grade 1 (Deacon et al., 2013). They found that scores on the full PPVT-III administered in Grade 1 correlated similarly with the scores on the subset of items administered in Grade 1 and the subset of items administered in Grade 2 (0.59 and 0.57, respectively; Deacon et al., 2013), and thus they suggested it was valid to use the shortened form in analyses. Although these correlations show that the shortened form is a reliable measure of vocabulary across time, there is only a moderate correlation between scores on the full version and scores on the shortened version. This suggests that there is variance measured by the full version of the PPVT-III that is not accounted for by the shortened version. This could call into question the validity of using the shortened version when examining unique variance accounted for in a reading measure by different oral language or cognitive skills.

In a study that investigated the relationship between morphological awareness and reading comprehension across Grades 3 and 4 (while controlling for vocabulary, phonological awareness, and nonverbal ability), the same modified version of the PPVT-III was used (Deacon et al., 2014). Like the previous study (Deacon et al., 2013), scores on the full and shortened item sets administered when the participants were in Grade 1 correlated similarly with scores on the subset of items administered in Grade 3 (0.62 and 0.61, respectively) and Grade 4 (0.65 and 0.63, respectively; Deacon et al., 2014). The authors argued that the similarity between these correlations indicated that there was stability in the measurement across time, despite a reduction in the number of items administered (Deacon et al., 2014). However, a shortcoming of validating these measures may be that how well the short form of the PPVT correlated with later reading comprehension was not examined, nor was it compared to the correlation between the full form of the PPVT and later reading comprehension.

A modified version of the PPVT-IV, in which every third item was administered, was used in two recent studies also investigating the relationship between vocabulary, morphological awareness, and reading comprehension (Metsala et al., 2021; Sparks & Metsala, 2021). Metsala et al. (2021) examined the contribution of multiple oral language skills to reading comprehension for students initially in Grades 2 and 3, including vocabulary, syntactic awareness, and morphological awareness. The authors found that vocabulary predicted variance in reading comprehension at both follow-up periods 6 and 18 months later, after controlling for initial reading comprehension and word reading (Metsala et al., 2021). After syntactic awareness was added to the equation, vocabulary accounted for unique variance at the short-term follow-up, but it did not contribute unique variance at the longer-term follow-up. Additionally, vocabulary did not account for unique variance at the short-term or longer-term follow-up after morphological

awareness was also added to the equation (Metsala et al., 2021). Therefore, it was concluded that vocabulary accounts for unique variance in gains in reading comprehension, but once additional oral language skills are added into the equation the variance accounted for by vocabulary may be largely shared. Sparks and Metsala (2021) reported similar findings for a Grade 1 sample using this shortened version of the PPVT-IV. While Deacon and colleagues (2013; 2014) provided some initial support for using a shortened form of the PPVT, there remain questions of the validity insofar as the shortened versions may not be as strongly related to the dependent variable as the full form. This may mean that short forms may not be a good measure for research questions examining the unique contribution of vocabulary and different independent variables to a dependent variable such as reading comprehension. For these studies reported, for example, it can be questioned whether the predictive strength of vocabulary knowledge might have been stronger if the full measure had been used.

The Relationship Between Vocabulary and Phonological Awareness

Reading comprehension can be difficult to measure in younger students, as word reading abilities in young children vary widely and may limit many young students' reading of text (for discussion, see Campbell, 2021). In order to examine the validity of shortened forms of the PPVT-5 in a sample including kindergarten and Grade 1 students, I examined the strength of the relationship with phonological awareness. A substantial body of research has found that vocabulary knowledge contributes unique variance to both concurrent and subsequent phonological awareness (Carroll & Snowling, 2001; Lonigan et al., 2000; Sénéchal & LeFevre, 2002; for review, see Metsala, 2011). For example, in their cross-sectional study of 700 two- to five-year-old children, McDowell et al. (2007) found that vocabulary contributed unique variance to the prediction of phonological awareness. Metsala (1999) found that the correlation

between a phonological awareness composite variable and the PPVT-R raw scores for four- to six-year-old children was .662. Additionally, in an intervention study in which preschool children received training in vocabulary over the course of an academic year, they improved on post-intervention phonological awareness compared to a control group (Lonigan, 2007). Based on this past research, phonological awareness and vocabulary are correlated in young students and therefore phonological awareness was used to compare the strength of the associations between the short and full versions of the PPVT-5 in the current study.

Vocabulary Reaction Time and the Relationship to Reading Comprehension

An aspect of vocabulary that is also examined in this study is the time taken to identify the correct picture representation for a known word; that is, one measure of the speed of accessing vocabulary knowledge. Poulsen and Elbro (2013) conducted one of the few studies to examine the relationship between a somewhat similar measure, speed of lexical access, and reading comprehension. They defined and investigated two different components of lexical access (phonological and semantic) in Grade 5 Danish students, although only semantic access is relevant to this study and discussed here. Semantic access was assessed by measuring isolated picture naming speed using a computer-based task (Poulsen & Elbro, 2013). In this task, pictures were presented on the computer screen one at a time, and each picture depicted a target word of three or four syllables. The amount of time elapsed between the presentation of the picture and the onset of the appropriate response was recorded as the response time (Poulsen & Elbro, 2013). Reading comprehension was measured using a group-administered task that consisted of two expository texts followed by 19 cloze questions for the first and 15 cloze questions for the second (Poulsen & Elbro, 2013). The authors found that picture naming speed was significantly correlated with reading comprehension ($r = 0.46$). Additionally, results from the hierarchical

regression analyses showed that picture naming speed accounted for unique variance in reading comprehension after accounting for variance due to vocabulary depth and letter naming speed (8%). In another hierarchical regression analysis controlling for word reading fluency and vocabulary depth, picture naming speed still accounted for 5% of the variance in reading comprehension (Poulsen & Elbro, 2013).

As previously mentioned, in their study investigating the relationship between different dimensions of vocabulary and reading comprehension, Tannenbaum et al. (2006) defined and measured speed of lexical access differently from Poulsen and Elbro (2013). Instead of using the term speed of lexical access, Tannenbaum et al. used the term vocabulary fluency, and they measured it in two different ways: first by giving participants a specified amount of time to use as many target words as possible in sentences, and then by giving participants a specified amount of time to name as many items as possible from eight categories (Tannenbaum et al., 2006). Unfortunately, Tannenbaum et al. (2006) combined their measures of vocabulary depth and fluency before conducting their analyses, and so the contribution of vocabulary fluency to reading comprehension could not be determined. However, it should be recalled that they found their measures of depth and fluency loaded on the same construct.

Although the two studies defined and measured their terms differently, it can be argued that response time/picture naming speed in the Poulsen and Elbro study and vocabulary fluency in the Tannenbaum et al. study were both measuring a similar construct, speed of lexical access; the speed at which an individual can access words from their lexicon (Poulsen & Elbro, 2013; Tannenbaum et al., 2006). The current study explores a similar construct which can be thought of as the time taken to access the meaning of a given word. This study examines whether the speed

of accessing the meaning of words plays a role in predicting unique variance in students' reading comprehension.

Goals of the Current Study

Based on previous research, it is clear that there exists a significant, positive correlation between vocabulary and reading comprehension (e.g., Binder et al., 2017; Nation & Snowling, 2004; Proctor et al., 2006). This has been shown across a wide range of studies examining participants of different ages (but see Fricke et al., 2016 for a study where this is not the case) and first languages and using a variety of measures to assess vocabulary and reading comprehension. One measure of vocabulary that has been extensively used in both research and practice is the Peabody Picture Vocabulary Test. Some initial validity work has been done on shortened forms of the PPVT-III and PPVT-IV. As previously mentioned, two studies reported that scores on the full and shortened item sets administered at a younger age correlated similarly with scores on the shortened item sets administered at older ages (Deacon et al., 2013; Deacon et al., 2014). This indicates that the shortened forms are sensitive to individual differences in children's receptive vocabulary. However, whether the strength of these shortened versions in predicting reading comprehension is as strong as the full version has yet to be examined. These shortened forms of the PPVT-III and PPVT-IV have been used in a handful of studies to date, with these studies focusing predominately on examining the unique relationship between morphological awareness and reading comprehension, using vocabulary as one of the control variables (Deacon et al., 2013; Deacon et al., 2014). Similarly, a recent study found that vocabulary predicted gains in later reading comprehension for second and third grade students (Metsala et al., 2021). However, vocabulary was not a predictor with morphological awareness in the equation. Likewise, a shortened form of the PPVT-IV did not contribute unique variance

to later reading comprehension in first grade students (Sparks & Metsala, 2021). A shortcoming of all these studies may be that using shortened versions rather than the full form of the PPVT weakened the strength of the association found between vocabulary and reading comprehension. This is an open empirical question.

Another goal of this study was to examine the validity of shortened versions of the PPVT-5 in a sample including younger students, those in kindergarten and first grade. Since reading comprehension can be challenging to measure in younger students, the predictive relationship for different versions of the PPVT was examined with phonological awareness. Similar to the relationship between vocabulary and reading comprehension, there is a substantial body of research supporting the positive association between vocabulary and phonological awareness (e.g., Carroll & Snowling, 2001; Lonigan et al., 2000; Metsala, 2011; Sénéchal & LeFevre, 2002).

A final goal of this study concerns the speed of accessing word meanings and its relationship to reading comprehension. I am not aware of prior research examining the relationship between reaction time for correct items on the PPVT and reading comprehension. As reviewed previously, one study found that speed of lexical access accounted for unique variance in reading comprehension, even after controlling for vocabulary (Poulsen & Elbro, 2013).

Given the overall goals of this study, my first research question is whether two shortened forms of the PPVT-5 are weaker in their association with reading comprehension compared to the full version in a sample of first and second grade students. My second research question addresses the validity of shortened forms of the PPVT-5 in a sample including younger students and asks whether these are more weakly associated with phonological awareness than the full version, in a sample of kindergarten and first grade students. Across both these samples, PPVT-5

shortened forms were calculated by including: i) every third item, and ii) every fourth item. This approach is consistent with previous studies that used shortened versions of the PPVT in which every third item (Metsala et al., 2021; Sparks & Metsala, 2021) or every fourth item (Deacon et al., 2013; Deacon et al., 2014) were administered. The final research question is whether first and second grade students' reaction times to identify correct items on the PPVT-5 explain additional variance in reading comprehension after controlling for the number of correct answers, or vocabulary breadth.

I hypothesize that the full version of the PPVT-5 will correlate more strongly with reading comprehension than the shortened versions of the PPVT-5 in the first and second grade students. Similarly, I hypothesize that the full version will correlate more strongly with phonological awareness than the shortened versions in the younger sample. Confirmation of these hypotheses would have implications for future research. I also predict that individual differences in students' mean reaction time to items on the PPVT-5 will explain additional variance in reading comprehension after controlling for the effects of vocabulary breadth. This would be consistent with and add to previous research concerning how the timing of lexical access influences making meaning from text (Poulsen & Elbro, 2013).

Methods

Participants and Procedure

Sample 1

Sample 1 consists of 27 kindergarten students (52% female, mean age 66.11 months) and 18 Grade 1 students (55% female, mean age 79.50 months) who participated in a larger vocabulary learning study. These students were from three different schools in a rural region of Nova Scotia. All participants learned English as their first language, as reported by their

guardians. Trained research assistants conducted all testing in a quiet room at the participants' schools.

Sample 2

Sample 2 consists of 32 Grade 1 students (50% female, mean age 76.88 months) and 33 Grade 2 students (64% female, mean age 86.16 months) who participated in a larger vocabulary instruction study. Ninety-six per cent of the participants spoke English as their first language, and the remaining children learned English between 2 and 5 years of age, as reported by their guardians. This study included students from eight classrooms in one suburban school in Nova Scotia. Trained research assistants conducted all testing in a quiet room at the participants' school.

Measures

Sample 1

Vocabulary Breadth. The Peabody Picture Vocabulary Test – Fifth Edition (PPVT-5; Dunn & Dunn, 2018) was used to measure receptive vocabulary breadth. The PPVT-5 is a standardized, individually administered test of receptive vocabulary in which the examinee is required to select the picture (from an array of four) that they think best illustrates the meaning of the stimulus word presented orally by the examiner (Dunn & Dunn, 2018). The test deviated from standard administration as it was administered using a tablet. All other standardized procedures were followed. The start item was determined by the age of each participant, and testing was discontinued following six consecutive incorrect responses.

PPVT-5 Short Form 3. In this shortened version of the PPVT-5, a score was calculated for each participant that summed every third item.

PPVT-5 Short Form 4. In this shortened version of the PPVT-5, a score was calculated for each participant that summed every fourth item.

Elision. The Elision subtest of the Comprehensive Test of Phonological Processing – Second Edition (CTOPP-2; Wagner et al., 2013) was used to measure phonological awareness. For this test, participants are required to remove phonological segments from spoken words to form other words (e.g., "Say cat without saying /c/"). Standard procedures were followed, with testing discontinued following three consecutive incorrect responses.

Sample 2

Vocabulary Breadth and Vocabulary Reaction Time. The Peabody Picture Vocabulary Test – Fifth Edition (PPVT-5; Dunn & Dunn, 2018) was also used to measure receptive vocabulary breadth (see Sample 1). As per Sample 1, the test deviated from standard administration as it was administered using a tablet and participants' reaction time for each item was recorded. All other standardized procedures were followed. The start item was determined by the age of each participant, and testing was discontinued following six consecutive incorrect responses.

PPVT-5 Short Form 3. In this shortened version of the PPVT-5, a score was calculated for each participant that summed every third item.

PPVT-5 Short Form 4. In this shortened version of the PPVT-5, a score was calculated for each participant that summed every fourth item.

PPVT-5 Reaction Time. A program on each tablet measured reaction time. The clock started at the end point of the presentation of the spoken word by a computerized voice. The clock stopped with the participants' tapping on one of the four items on the screen. Participants' mean reaction time across all correct items was calculated and represented each participant's

measure of reaction time. Reaction times for only correct items were used because it represents the construct of interest, the time taken to access the meaning of a known word; this would not be the case for reaction times for incorrect items which are presumably unknown words for the participant. Extreme reaction times, defined as values above two standard deviations from an individual participant's mean, were replaced with this cutoff value (two standard deviations above the mean). This method of handling outliers is consistent with the practices used in past, similar research (e.g., Poulsen & Elbro, 2013).

Reading Comprehension. The Passage Comprehension subtest of the Woodcock Reading Mastery Tests – Third Edition (WRMT-III; Woodcock, 2011) was used to measure reading comprehension. The Passage Comprehension subtest requires participants to read short passages silently and supply the missing word in each. All standardized procedures were followed, including start points based on the grade of the participants, and testing was discontinued following four consecutive incorrect responses.

Word Reading. The Word Identification subtest of the Woodcock Reading Mastery Tests – Third Edition (WRMT-III; Woodcock, 2011) was used to measure single word reading. This task is comprised of 46 words presented in order of increasing difficulty. Participants were required to read the list of words aloud until they made four consecutive errors.

Results

Descriptive Statistics

Sample 1

Initial analysis of z scores for each variable in Sample 1 revealed that 3.33% of the data were outliers and these were replaced with the cutoff value of two standard deviations above or below the mean (Jajodia, 2017). Distributions for the three ways of scoring the vocabulary

measure and for the elision task were all normal (George & Mallery, 2010). Table 1 provides the means and standard deviations for the raw scores for all variables. When available, this information is also provided for standardized scores in order to help characterize the sample's performance on these measures. As can be seen in Table 1, the mean standard score for the vocabulary measure fell within the average range.

Sample 2

From the original 32 Grade 1 students in Sample 2, two participants were excluded from analyses because they both had only one correct response on the PPVT-5 and thus scores on each version of the PPVT-5 could not be calculated. From the 33 Grade 2 students, one participant was excluded because they did not complete the reading comprehension measure. Analyses of z scores for each variable, other than vocabulary reaction time, revealed one value as an outlier and this value was replaced with the cutoff value of two standard deviations above the mean. For the reaction time data, 215 values (5.21%) were identified as outliers and were replaced with the cutoff value of two standard deviations above the mean. This method of handling outliers is consistent with common practices (Jajodia, 2017). Table 1 provides descriptive statistics, including means and standard deviations for the raw scores used in analyses for all variables; standard scores calculated for each measure used in the study are also reported. As can be seen in Table 1, the mean standard scores for the vocabulary, reading comprehension, and word reading measures fell within the average range.

Table 1

Means and Standard Deviations for Study Measures for Grades Primary, 1, and 2 Students

	Sample 1 (N = 45)		Sample 2 (N = 62)		Grade 1 (Sample 2) (n = 30)		Grade 2 (Sample 2) (n = 32)	
Measure	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Vocabulary (full scale; RS)	110.14	15.85	124.02	20.97	116.70	16.62	130.88	22.51
Vocabulary (full scale; SS)	91.00	------	95.00	------	93.00	------	97.00	------
Reading Comp. (RS)	------	------	8.74	4.98	6.03	4.52	11.28	3.98
Reading Comp. (SS)	------	------	101.33	14.21	------	------	------	------
Vocabulary Reaction Time	------	------	2.82	.40	2.69	.28	2.94	.46
Phonological Awareness (RS)	11.02	4.89	------	------	------	------	------	------
Word Reading (RS)	------	------	14.22	7.37	9.63	6.44	18.53	5.36
Word Reading (SS)	------	------	104.09	14.96	------	------	------	------

Note: RS = raw score. SS = standard score.

Examining the Strength of Associations with Full and Shortened Versions of the PPVT-5

Sample 1

Zero-order correlations among measure raw scores for Sample 1 are presented in Table 2. Preliminary analyses were conducted to ensure no violations of the assumptions of normality, linearity, and homoscedasticity (Laerd Statistics, 2018). The Shapiro-Wilk test of normality revealed that each variable was approximately normally distributed. Visual inspection of scatterplots confirmed a linear relationship between each predictor variable and the response variable. Homoscedasticity was determined by visual inspection of scatterplots, which showed constant variance of the data points. Taken together, the data met assumptions for the planned correlational analyses (Laerd Statistics, 2018).

I first examined whether the short form of the PPVT-5 based on every third item was more strongly associated with the full version of the PPVT-5 than the short form based on every fourth item. The correlation between PPVT-5 Short Form 3 and the full version was $r = .963$, $p < .001$. The correlation between PPVT-5 Short Form 4 and the full version was $r = .904$, $p < .001$. To determine if the correlations between the shortened forms and the full version of the PPVT-5 were statistically different from each other, a *z* value and a *p* value were calculated (IBM, 2020). Since the research questions and hypotheses related to each comparison of correlations in this study are unidirectional, all tests (and associated *p*-values) comparing correlations are one-tailed. For this comparison, $z = 2.87$ and $p = .002$. The *z* value for the comparison was greater than 1.96 and the *p* value was less than .05, indicating that the difference between the correlations was statistically significant. This means that an administration in which every third item is included shared more variance with the full scale than an administration in which every fourth item is included.

I next examined whether scores on the shortened forms of the PPVT-5 were more weakly associated with the elision task than the full version of the PPVT-5. Conducting these comparisons between correlations allows us to further examine the validity of shortened forms of the PPVT-5. The correlations between the different forms of the PPVT-5 and elision were $r = .389, .408, .398$, *p's* $< .01$ for PPVT-5 full scale, PPVT-5 Short Form 3, and PPVT-5 Short Form 4, respectively. For the comparison between PPVT-5 Short Form 3 and the full version of the PPVT-5, $z = .50$ and $p = .31$, indicating that the strength of the correlations were not shown to differ. For the comparison between PPVT-5 Short Form 4 and the full version of the PPVT-5, $z = .16$ and $p = .44$, again indicating that the difference between the correlations was not statistically

significant. These results demonstrate that the shortened forms and the full version of the PPVT-5 are similarly correlated with elision.

Table 2

Zero-order Correlations for Sample 1

Measures	1.	2.	3.	4.
Full Sample				
1. Vocabulary (full scale)	-			
2. Vocabulary (every 3rd item)	.96**	-		
3. Vocabulary (every 4th item)	.90**	.86**	-	
4. Phonological Awareness	.39**	.41**	.40**	-

Note. $^{**}p < .01$

Sample 2

To address whether the shortened forms of the PPVT-5 were less strongly associated with reading comprehension than the full version, I compared the correlations between scores on the shortened forms of the PPVT-5 and reading comprehension to the correlation between the full version of the PPVT-5 and reading comprehension. This was done first across the entire sample and then separately by grade (see Table 3 for correlations across the entire sample and Table 4 for correlations for each grade separately).

Preliminary analyses were conducted to ensure no violations of the assumptions of normality, linearity, and homoscedasticity (Laerd Statistics, 2018). The Shapiro-Wilk test of normality revealed that each variable was approximately normally distributed. Visual inspection of scatterplots confirmed a linear relationship between each predictor variable and the response variable. Homoscedasticity was determined by visual inspection of scatterplots, which showed constant variance of the data points. Taken together, the data met assumptions for the planned correlation analyses (Laerd Statistics, 2018).

The correlations between the different forms of the PPVT-5 and reading comprehension were r = .530, .492, .481, p*'s* < .001 for PPVT-5 full scale, PPVT-5 Short Form 3, and PPVT-5 Short Form 4, respectively. For the comparison between PPVT-5 Short Form 3 and the full version of the PPVT-5, z = -1.27 and p = .10, indicating that the difference between the correlations was not statistically significant. For the comparison between PPVT-5 Short Form 4 and the full version of the PPVT-5, z = -1.41 and p = .08, indicating that the difference between the correlations was not statistically significant.

The Simple View of Reading (Hoover & Gough, 1990) suggests that the relationship between oral language and reading comprehension increases as children's reading develops, and

so similar comparisons were also made for each grade level separately (see Table 4 for correlations). For the Grade 1 students, the comparison between PPVT-5 Short Form 3 and the full version of the PPVT-5, z = -.84 and p = .20, and for the comparison between PPVT-5 Short Form 4 and the full version of the PPVT-5, z = .13 and p = .45. These comparisons indicate that the shortened forms and the full version of the PPVT-5 were not found to differ in the strength of their association with reading comprehension in Grade 1 students. Thus, for first grade students, both shortened versions of the PPVT-5 appear to be as strongly predictive of reading comprehension as the full version of the PPVT-5.

For the comparison between PPVT-5 Short Form 3 and the full version of the PPVT-5 for the Grade 2 students, z = -.83 and p = .21. For the comparison between PPVT-5 Short Form 4 and the full version of the PPVT-5 for the Grade 2 students, z = -1.88 and p = .03. So, PPVT-5 Short Form 3 was not found to be more weakly associated with reading comprehension than the full version of the PPVT-5. On the other hand, PPVT-5 Short Form 4 is not as strongly related to reading comprehension as the full version of the PPVT-5 in this sample of second grade students.

Table 3

Zero-order Correlations for Sample 2

Measures	1.	2.	3.	4.	5.	6.
Full Sample						
1. Vocabulary (full scale)	-					
2. Vocabulary (every 3rd item)	.96**	-				
3. Vocabulary (every 4th item)	.95**	.93**	-			
4. Reading Comprehension	.53**	.49**	.48**	-		
5. Vocabulary Reaction Time	.09	.06	.04	.23	-	
6. Word Reading	.43**	.40**	.38**	.88**	.35**	-

Note. $^{**}p < .01$

Table 4

Comparing Correlations for Grade 1 (top diagonal) vs. Grade 2 (bottom diagonal) Students in Sample 2

Measures	1.	2.	3.	4.	5.	6.
1. Vocabulary (full scale)	-	.95**	.92**	.24	.07	.01
2. Vocabulary (every 3rd item)	.97**	-	.90**	.19	.01	-.01
3. Vocabulary (every 4th item)	.96**	.94**	-	.25	-.03	.07
4. Reading Comprehension	.62**	.59**	.54**	-	.23	.87**
5. Vocabulary Reaction Time	-.06	-.08	-.07	-.01	-	.29
6. Word Reading	.55**	.52**	.45**	.78**	.18	-

Note. $^{**}p \leq .01$.

Contribution of Vocabulary Reaction Time to Reading Comprehension

The next planned step was to examine the relative contribution of vocabulary reaction time to reading comprehension. As can be seen in Table 3, the zero-order correlation between vocabulary reaction time and reading comprehension was not statistically significant; the planned analyses were nonetheless carried out which were designed to address whether students' reaction times on the PPVT-5 explain unique variance in reading comprehension after controlling for vocabulary breadth (and then word reading as well). Preliminary explorations were conducted to ensure all assumptions for multiple regression analysis were met. Visual inspection of scatterplots confirmed a linear relationship between each predictor variable and the response variable. Collinearity statistics were within acceptable limits. Independence of observations was tested using the Durbin Watson statistic and all were within an acceptable range, confirming independence of the residuals (Statology, 2021). Homoscedasticity was determined by visual inspection of plots of standard residuals versus standard predicted values. Inspection of residual error plots demonstrated that distribution of these scores were approximately normal. Inspection of normal P-P plots showed that all measures were normally distributed. Taken together, the data met assumptions for the planned multiple regression analyses (Statology, 2021).

Two hierarchical regression analyses were conducted. In the first hierarchical regression, grade was entered as the first step, vocabulary breadth was entered as the second step, and vocabulary reaction time was entered as the third step. As seen in Table 5 (Regression 1), at the second step, vocabulary breadth accounted for an additional and significant 14% of the variance in reading comprehension. As expected from the correlations, vocabulary reaction time, entered at the third step, did not account for additional variance in reading comprehension.

In the second hierarchical regression, word reading was also controlled as it is an important contributor to reading comprehension. As the second step, word reading accounted for an additional 50% of the variance in reading comprehension (see Table 5, Regression 2). Vocabulary breadth was entered as Step 3 and continued to account for significant, unique variance in reading comprehension (3%). As expected from the correlations, vocabulary reaction time did not account for unique variance as the final step in the regression.

Table 5

Hierarchical Regression Analyses Predicting Reading Comprehension (n = 62)

Step	Outcome	Predictor	ΔR^2	β	Final β
Regression 1	Reading Comprehension				
1		Grade	.28**	.53**	.37**
2		Vocabulary (full scale)	.14**	.40**	.40**
3		Vocabulary Reaction Time	.01	.08	.08
Regression 2	Reading Comprehension				
1		Grade	.28**	.53**	-.02
2		Word Reading	.50**	.89**	.84**
3		Vocabulary (full scale)	.03**	.19**	.18**
4		Vocabulary Reaction Time	.01	-.07	-.07

Note. $^{**}p \leq .01$.

Discussion

There were two primary goals of this study. The first was to explore whether shortened versions of the PPVT-5 predict reading related skills (including reading comprehension and phonological awareness) as well as the full version of the PPVT-5, a commonly used measure of receptive vocabulary breadth. Several recent research studies have used short forms of the PPVT to save time in lengthy research batteries and have provided some initial support for the validity of these shortened versions (Deacon et al., 2013; Deacon et al., 2014; Metsala et al., 2021; Sparks & Metsala, 2021); however, there still remain questions regarding the validity of shortened versions of the Peabody Picture Vocabulary Test. Furthermore, very little research has examined the relationship between the speed of accessing vocabulary information and any unique contribution to reading comprehension. No research, to our knowledge, has examined this specifically using reaction time data for items on the PPVT and reading comprehension, which is the second overall goal of this research study.

In order to examine the validity of shortened versions of the PPVT-5 in a sample of kindergarten and Grade 1 students (Sample 1), I looked at the predictive relationship with phonological awareness, since reading comprehension can be difficult to measure in younger students. As an initial step, results first showed that PPVT-5 Short Form 3 was more strongly correlated with the full scale than PPVT-5 Short Form 4. This finding indicates that PPVT-5 test scores based on every third item explain more variance in the full-scale version than scores based on every fourth item. That being said, both shortened versions were very highly correlated with the full version, accounting for the majority of the variance in the full version (i.e., 92.7% and 81.7% of the variance in the full form for every third item version and every fourth item version, respectively).

When the correlations between different versions of the PPVT-5 and a measure of phonological awareness were compared in the same sample of students, no significant differences were found. This finding can help lend support to the validity of using shortened versions of the PPVT-5 in studies examining the relationship between vocabulary and reading related processes or measures. Additionally, the magnitude of the correlations found in the current study for each form of the vocabulary test appear consistent with the correlations reported in other studies. For example, Ouellette and Beers (2010) found a correlation of .435 between scores on the full version of the PPVT-4 and scores on a phonological awareness measure in Grade 1 students. Deacon et al. (2014) found a correlation of .44 between scores on a shortened version of the PPVT-III in which every fourth item was administered and scores on a phonological awareness measure in Grade 3 students. These values are comparable to the correlations found in the current study across all three forms of the PPVT-5 (ranging from .389 to .408). Overall, then, it appears that shortened forms of the PPVT-5 in which every third item or every fourth item is administered may adequately capture individual differences in young children's vocabulary breadth and be used as an acceptable proxy for the full-scale Peabody Picture Vocabulary Test.

A primary research question was whether the shortened vocabulary test versions were weaker in their associations with reading comprehension than the full form. Across the entire sample of first and second grade students (Sample 2), no differences were found in the strength of these associations. Although no previous research, to my knowledge, has directly compared shortened versions of the PPVT with the full version in terms of their relationships to reading comprehension, the magnitudes of the correlations in this study appear similar to correlations found in other studies. For example, Ouellette and Beers (2010) found a correlation of .495

between the PPVT-4 (full version) and scores on a reading comprehension measure in Grade 1 students. Deacon et al. (2014) found a correlation of .45 between scores on a shortened version of the PPVT-III in which every fourth item was administered and scores on a reading comprehension measure in Grade 3 students. These values seem comparable to the correlations between vocabulary and reading comprehension found for Sample 2. Although the hypothesis was not supported across the entire sample, the absolute magnitudes of the correlations followed the pattern that was expected (530, .492, .481 for correlations with reading comprehension for the full version, every third item, and every fourth item, respectively).

There are reasons to suspect developmental differences in the strength of the relationship between vocabulary and reading comprehension. The Simple View of Reading, for example, proposes that the relationship between vocabulary and reading comprehension becomes stronger as children get older (Gough & Tunmer, 1986; Hoover & Gough, 1990). Indeed, there was a different pattern of results for the association between vocabulary and reading comprehension for first versus second grade students. For the Grade 1 students, the correlations between the different versions of the PPVT-5 and reading comprehension ranged from .19 to .25 and were not statistically significant. For the Grade 2 students, the correlations ranged from .54 to .62 and were statistically significant. The weak correlations found for the Grade 1 students are in contrast to Ouellette and Beers' (2010) findings for Grade 1 students, although it should be noted that after controlling for other reading-related skills in their regression analyses, vocabulary breadth did not explain unique variance in the first-grade students' reading comprehension. Given both theoretical reasons and empirical findings concerning differences in the strength of the relationship between vocabulary and reading comprehension, it was planned to compare the magnitude of these correlations separately for first and second grade students.

For second grade students, there was some indication that not all forms of the vocabulary test were equal. The association between PPVT-5 Short Form 4 and reading comprehension was weaker than the association between the full version and reading comprehension. As predicted, scores based on every fourth item of the vocabulary test were not as strongly associated with reading comprehension as scores based on every item in this sample of second grade students. This raises potential issues in studies which have used a form of the test administering only every fourth item. In their studies, Deacon et al. (2013; 2014) justified the use of such a shortened version of the PPVT-III by noting that scores on the full and shortened item sets administered at a younger age correlated similarly with scores on the shortened item sets administered at older ages. However, they didn't call attention to the fact that the shortened form and the full version were only moderately correlated, indicating that there was variance in the full version that was not being accounted for by the shortened version. They also did not compare the correlations between the full and shortened versions of the PPVT and other variables, such as reading comprehension. Although the current study found high correlations between shortened and full versions of the PPVT-5, there was still a significant difference in how well scores based on every fourth item predicted reading comprehension. The consequences may be that for these past studies (Deacon et al., 2013; 2014), findings that morphological awareness made a unique contribution to reading comprehension may have been different had the full version of the vocabulary test been used. Metsala et al. (2021) and Sparks and Metsala (2021) used a version of the PPVT in which every third item was administered. Although differences were not found in the current study, further research with larger samples might explore whether this shortened version can take the place of the full test in examining unique relationships between variables. For example, it is possible that had they used the full version of the PPVT-IV in their studies, as

opposed to a shortened version with every third item, they may have found that vocabulary did account for unique variance in reading comprehension even after controlling for syntactic and/or morphological awareness. It should be noted that the amounts of unique variance accounted for by each individual variable tends to be small in these studies (e.g., 1-4%), and so small differences in the association of the vocabulary test with the dependent variable may well influence statistical findings.

The final goal of this study was to investigate the contribution of speed or time taken to access vocabulary knowledge to reading comprehension in Grade 1 and 2 students (Sample 2). Vocabulary reaction time did not explain significant unique variance in reading comprehension, before or after accounting for the variance explained by grade, word reading, and vocabulary breadth. This finding was not surprising given that the zero-order correlation between vocabulary reaction time and reading comprehension was not statistically significant, although it contrasted with my prediction that vocabulary reaction time would explain additional unique variance in reading comprehension. A very limited number of studies have examined the relationship between timed aspects of vocabulary and reading comprehension, and these lead to my prediction that there would be a unique association. There are a few potential explanations for the differences between the findings of the current study and the findings of the other studies. The way Poulsen and Elbro (2013) defined and measured reaction time was different than in the current study. The task used in their study was a measure of expressive vocabulary and required an oral response. This contrasts with the current study, which used a measure of receptive vocabulary with a nonverbal response. Additionally, the PPVT-5 employed a multiple-choice format, while the task used in the Poulsen and Elbro study did not. Both of these differences may have contributed to the different findings. For example, it could be that the task in the Poulsen

and Elbro study was a more specific measure of the time for students to access individual words, since it required them to generate a one-word response. The PPVT-5 requires examination of multiple pictures before making a response. Comparison of visual stimuli may thus affect reaction times, rather than a purer measure of accessing vocabulary knowledge; the PPVT-5 is also susceptible to the effects of guessing. The differences in the manner of measuring vocabulary and timing may have led to Poulsen and Elbro finding a stronger relationship between vocabulary related timing and reading comprehension.

Another important difference was that the participants in Poulsen and Elbro's (2013) study were in Grade 5 compared to the first and second grade students in the current study. It may be that reaction time for accessing vocabulary knowledge has a larger influence on children's reading comprehension with development, as they become more skilled readers. This would make sense when considered in the context of the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990). The Simple View of Reading suggests that the contribution of vocabulary knowledge to reading comprehension increases as children develop, and so it is possible that this pattern extends to other aspects of vocabulary as well, such as speed of accessing vocabulary knowledge. Additionally, Tannenbaum et al. (2006) found that vocabulary fluency was related to reading comprehension. In that study, vocabulary fluency was a combined measure of timing to access a word and vocabulary depth. From that study, with students in Grade 3, perhaps it was the depth aspect of the combined measure that accounted for the significant association with reading comprehension.

Implications and Limitations of the Current Study

This study has implications for how vocabulary is measured in research studies, and also for whether vocabulary reaction time will explain individual differences in reading

comprehension in young students. When examining whether scores on the shortened forms and the full version of the PPVT-5 were differentially correlated with scores on measures of reading comprehension and phonological awareness, there were many similarities in the strength of the associations between the different forms. However, for students in Grade 2, the full version of the PPVT-5 was more strongly correlated with reading comprehension than the version which counted only every fourth item. These findings suggest that researchers might think twice before substituting shortened versions of the PPVT-5 for the full version. This study indicates that this is particularly applicable to shortened versions in which every fourth item is administered.

The results of this study were also meant to address the question of the role of timing in accessing vocabulary knowledge in explaining reading comprehension. The results from the current study indicated that vocabulary reaction time in first and second grade students on the Peabody Picture Vocabulary Test did not explain individual differences in reading comprehension. This could mean that the time taken to access vocabulary knowledge does not play as large of a role in reading comprehension, as was predicted based on previous research. Alternatively, it is possible that the predictive relationship between reaction time for vocabulary knowledge and reading comprehension is significant for older students, but not in the early elementary grades. Very little research in this area has been completed to date, and the measures used in this study may not have captured the variance in accessing vocabulary knowledge. Further research will build toward a consensus in understanding whether individual differences in the time to access vocabulary knowledge influences reading comprehension.

As with any study, the current study has limitations that need to be considered. One limitation was the relatively small sample size. Small sample sizes may influence the generalizability of the study findings or may limit the power to find statistically significant

differences. Not having a reading comprehension measure for Sample 1 meant that the relations between the different versions of the vocabulary test and reading comprehension could not be examined for this younger sample. However, using a measure of phonological awareness seems a better way to examine shortened versions of the PPVT-5 in the sample of kindergarten and Grade 1 students, given that scores on the different versions of the PPVT-5 and scores on the reading comprehension measure were not found to be significantly correlated in the Grade 1 students in Sample 2. Vocabulary was robustly related to phonological awareness in this young sample, allowing comparison of the different versions of the vocabulary test. Another limitation of the current study was that the analyses were not conducted separately by grade for Sample 1, and so developmental differences across kindergarten to second grade could not be examined. A final limitation was the absence of full demographic data on the samples, including SES and home literacy environment. We know there is a strong relationship between these factors and vocabulary acquisition, and it would be beneficial to include demographic variables in future research.

In summary, this study contributes to a sparse body of existing research on the validity of shortened forms of the PPVT, and to that on the relationship between vocabulary reaction time and reading comprehension in young elementary school students. The findings from the current study call into question the use of shortened versions of the PPVT as an alternative way of measuring vocabulary breadth. The assumption that these shortened versions are as strongly predictive of reading and reading-related skills as the full version was not completely supported in this study. Furthermore, the results from the regression analyses suggest that for young children, time to access vocabulary knowledge may not make unique contributions to reading comprehension.

References

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*(3), 251-271. https://doi.org/10.1086/511706

Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology, 24*(3-4), 323-335. https://doi.org/10.1080/02702710390227297

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research, volume 2* (pp. 41-51). The Guilford Press.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498-520. https://doi.org/10.1037/0022-0663.93.3.498

Binder, K. S., Cote, N. G., Lee, C., Bessette, E., & Vu, H. (2017). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of Research in Reading, 40*(3), 333-343. https://doi.org/10.1111/1467-9817.12069

Bochner, S. (1978). Reliability of the Peabody Picture Vocabulary Test: A review of 32 selected research studies published between 1965 and 1974. *Psychology in the Schools*, *15*(1), 320–327. https://doi.org/10.1002/1520-6807(197807)15:3<320::AID-PITS2310150302>3.0.CO;2-B

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research, 6*(2), 145-173. https://doi.org/10.1191/1362168802lr103oa

Campbell, S. (2021). *Addressing reading comprehension in elementary-school readers: How do decoding and oral language skills predict performance on two measures, and are the same students identified as having a comprehension difficulty?* [Master's thesis, Mount Saint Vincent University]. Mount E-commons.

Carroll, J. M., & Snowling, M. J. (2001). The effects of global similarity between stimuli on children's judgment of rime and alliteration. *Applied Psycholinguistics*, *22*(3), 327-342. https://doi.org/10.1017/S0142716401003034

Cervatiuc, A. (2007). Assessing second language vocabulary knowledge. *International Forum of Teaching and Studies*, *3*(3), 40-47.

Chen, K. Y. (2011). The impact of EFL students' vocabulary breadth of knowledge on literal reading comprehension. *Asian EFL Journal*, *51*, 30-40.

Coxhead, A., Nation, P., & Sim, D. (2015). Measuring the vocabulary size of native speakers of English in New Zealand secondary schools. *New Zealand Journal of Educational Studies*, *50*(1), 121-135. https://doi.org/10.1007/s40841-015-0002-3

Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly*, *30*(2), 74-88. https://doi.org/10.2307/30035543

Deacon, S. H., Benere, J., & Pasquarella, A. (2013). Reciprocal relationship: Children's morphological awareness and their reading accuracy across grades 2 to 3. *Developmental Psychology*, *49*(6), 1113-1126. https://doi.org/10.1037/a0029474

Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal

models. *Scientific Studies of Reading*, *18*(6), 432-451. https://doi.org/10.1080/10888438.2014.926907

Duff, D., & Brydon, M. (2020). Estimates of individual differences in vocabulary size in English: How many words are needed to 'close the vocabulary gap'? *Journal of Research in Reading*, *43*(4), 454-481. https://doi.org/10.1111/1467-9817.12322

Dunn, L. M., & Dunn, D. M. (2018). *Peabody Picture Vocabulary Test* (5th ed.). Pearson Education.

Farvardin, M. T., & Koosha, M. (2011). The role of vocabulary knowledge in Iranian EFL students' reading comprehension performance: Breadth or depth. *Theory and Practice in Language Studies*, *1*(11), 1575-1580. https://doi.org/10.4304/TPLS.1.11.1575-1580

Fricke, S., Szczerbinski, M., Fox-Boyer, A., & Stackhouse, J. (2016). Preschool predictors of early literacy acquisition in German-speaking children. *Reading Research Quarterly*, *51*(1), 29-53. https://doi.org/10.1002/rrq.116

George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update* (10th ed.). Boston: Pearson.

Goriot, C., Van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2021). Using the Peabody Picture Vocabulary Test in L2 children and adolescents: Effects of L1. *International Journal of Bilingual Education and Bilingualism*, *24*(4), 546-568. https://doi.org/10.1080/13670050.2018.1494131

Gottardo, A., Mirza, A., Koh, P. W., Ferreira, A., & Javier, C. (2018). Unpacking listening comprehension: The role of vocabulary, morphological awareness, and syntactic knowledge in reading comprehension. *Reading and Writing*, *31*(8), 1741-1764. https://doi.org/10.1007/s11145-017-9736-2

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

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator*, *27*(1), 4-9.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368-1378. https://doi.org/10.1111/1467-8624.00612

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, *2*(2), 127-160. https://doi.org/10.1007/BF00401799

IBM. (2020, April 16). *Testing differences between dependent correlations.* https://www.ibm.com/support/pages/testing-differences-between-dependent-correlations

Jajodia, P. (2017, February). *Removing outliers using standard deviation in Python*. KDnuggets. https://www.kdnuggets.com/2017/02/removing-outliers-standard-deviation-python.html

Kang, Y., Kang, H. S., & Park, J. (2012). Is it vocabulary breadth or depth that better predict Korean EFL learners' reading comprehension? *English Teaching*, *67*(4), 149-171. https://doi.org/10.15858/ENGTEA.67.4.201212.149

Karami, H. (2012). The development and validation of a bilingual version of the Vocabulary Size Test. *RELC Journal*, *43*(1), 53-67. https://doi.org/10.1177/0033688212439359

Kim, B. (2016, May 20). *Hierarchical linear regression*. University of Virginia library: Research data services + sciences. https://data.library.virginia.edu/hierarchical-linear-regression/

Kremmel, B., & Schmitt, N. (2018). Vocabulary Levels Test. In J. I. Liontas, M. DelliCarpini, & N. J. Anderson (Eds.), *The TESOL encyclopedia of English language teaching* (pp. 1-7). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0499

Laerd Statistics. (2018). *Pearson's product-moment correlation using SPSS statistics.* https://statistics.laerd.com/spss-tutorials/pearsons-product-moment-correlation-using-spss-statistics.php

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, *36*(5), 611-634. https://doi.org/10.1093/applin/amu007

Lonigan, C. J. (2007). Vocabulary development and the development of phonological awareness skills in preschool children. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 15-31). New York: Guilford Press.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36*(5), 596-613. https://doi.org/10.1037/0012-1649.36.5.596

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300-335. https://doi.org/10.3102/0034654310377087

McDowell, K. D., Lonigan, C. J., & Goldstein, H. (2007). Relations among socioeconomic status, age, and predictors of phonological awareness. *Journal of Speech, Language, and Hearing Research, 50*(4), 1079-1092. https://doi.org/10.1044/1092-4388(2007/075)

McEwan, E. K. (2008). *Root words, roots and affixes*. Reading rockets. https://www.readingrockets.org/article/root-words-roots-and-affixes

Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, *4*(2), 142-154. https://doi.org/10.1177/026553228700400202

Meara, P., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English-speaking world, volume 4: Writing and vocabulary in foreign language acquisition* (pp. 85-96). Museum Tusculanum Press.

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*(1), 3–19. https://doi.org/10.1037/0022-0663.91.1.3

Metsala, J. L. (2011). Lexical reorganization and the emergence of phonological awareness. *Handbook of Early Literacy Research, 3*, 66-84.

Metsala, J. L., Sparks, E., David, M., Conrad, N., & Deacon, S. H. (2021). What is the best way to characterise the contributions of oral language to reading comprehension: Listening comprehension or individual oral language skills? *Journal of Research in Reading*, *44*(3), 675-694. https://doi.org/10.1111/1467-9817.12362

Moats, L. C. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do*. American Federation of Teachers, Washington, DC.

Morrison, F. J., Williams, M. A., & Massetti, G. M. (1998). The contributions of IQ and schooling to academic achievement. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, San Diego, CA.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342-356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Nation, P., & Coxhead, A. (2014). Vocabulary size research at Victoria University of Wellington, New Zealand. *Language Teaching*, *47*(3), 398-403. https://doi.org/10.1017/S0261444814000111

National Institute for Literacy. (2008). Developing early literacy: Report of the National Early Literacy Panel. Jessup, MD.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nouri, N., & Zerhouni, B. (2016). The relationship between vocabulary knowledge and reading comprehension among Moroccan EFL learners. *IOSR Journal of Humanities and Social Science*, *21*(10), 19-26. https://doi.org/10.9790/0837-2110051926

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554-566. https://doi.org/10.1037/0022-0663.98.3.554

Ouellette, G. P., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*(2), 189-208. https://doi.org/10.1007/s11145-008-9159-1

Poulsen, M., & Elbro, C. (2013). What's in a name depends on the type of name: The relationships between semantic and phonological access, reading fluency, and reading

comprehension. *Scientific Studies of Reading*, *17*(4), 303-314. https://doi.org/10.1080/10888438.2012.692743

Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, *98*(1), 159-169. https://doi.org/10.1037/0022-0663.98.1.159

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 209-227). John Benjamins Publishing Company.

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235-257. https://doi.org/10.1080/10888430701344306

Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, *73*(2), 445-460. https://doi.org/10.1111/1467-8624.00417

Sénéchal, M., Lefevre, J. A., Thomas, E. M., & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *33*(1), 96-116. https://doi.org/10.1598/RRQ.33.1.5

Sparks, E., & Metsala, J. L. (2021). *Morphological awareness predicts reading comprehension in first grade students* [Manuscript submitted for publication]. Department of Psychology and Neuroscience, Dalhousie University.

Statology. (2021, November 16). *The five assumptions of multiple linear regression.* https://www.statology.org/multiple-linear-regression-assumptions/

Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly, 11*(3), 271-282. https://doi.org/10.1080/15434303.2014.922977

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934-947. https://doi.org/10.1037/0012-1649.38.6.934

Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398. https://doi.org/10.1207/s1532799xssr1004_3

Wagner, R.K., Torgesen J., & Rashotte, C. (2013). *Comprehensive Test of Phonological Processing* (2nd ed.). Austin, TX: Pro-Ed.

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98*(1), 63-74. https://doi.org/10.1037/0022-0663.98.1.63

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *International Journal of Applied Linguistics, 168*(1), 33-69. https://doi.org/10.1075/itl.168.1.02web

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests* (3rd ed.). Bloomington, MN: Pearson Education.