HHS Public Access

Author manuscript

J Psycholinguist Res. Author manuscript; available in PMC 2020 August 19.

Published in final edited form as:

J Psycholinguist Res. 2020 April ; 49(2): 335–350. doi:10.1007/s10936-020-09694-8.

The Factor Structure of Vocabulary: An Investigation of Breadth and Depth of Adults with Low Literacy Skills

An H. Tran[1], **Kathryn A. Tremblay**[1], **Katherine S. Binder**[1]

[1]Department of Psychology and Education, Mount Holyoke College, 50 College St, South Hadley, MA 01075, USA

Abstract

The purpose of this study was to examine the factor structure of vocabulary. We believe that not only is vocabulary multidimensional, but depth of vocabulary knowledge should also be assessed with multiple measures since it too, is composed of multiple aspects. Furthermore, to explore the predictive validity of the different aspects of vocabulary knowledge, we assessed the relationship between vocabulary breadth, vocabulary depth, and reading comprehension in adults with low literacy skills. The participants were 103 adults. They completed 12 tasks that have been used in past studies to measure vocabulary breadth, depth, and reading comprehension. We had several important findings. First, we confirmed that all of the assessments were highly reliable for adults with low literacy skills. Second, the results of the factor analysis indicated two distinct vocabulary factors. Finally, both breadth and depth contribute independently to explaining variance in reading comprehension. Implications for vocabulary measurement are suggested.

Keywords

Vocabulary depth; Vocabulary breadth; Reading comprehension; Adult literacy; Factor analysis

Introduction

Adult literacy is an important issue around the globe. Results from the Program for the International Assessment of Adult Competencies (PIAAC) indicated that for all participating countries, workers with low levels of literacy proficiency had incomes among the lowest 20% in their country (Grotlüschen et al. 2016), and other studies have shown that low literacy skills also correlate with adverse health conditions (Boyes et al. 2016). Statistics in the United States suggest that one in six adults has weak literacy skills, which is equivalent to about 52.4 million adults (OECD 2013). This is significantly more than in 2003, when about 30 million adults in the United States were below a basic literacy level (National Center for Education Statistics [NCES] 2003). These statistics show that the situation is getting worse, at least in the US. Such studies suggest reading deficits might leave a pessimistic mark on future generations, in addition to creating a negative impact on employment and limiting access to health care and social services. Thus, the effectiveness of adult education programs is crucial, but more research is needed to understand the cognitive

Katherine S. Binder, kbinder@mtholyoke.edu.

profiles of these adult learners. While there are many component skills that comprise literacy, we evaluated which vocabulary factors contribute to adults' reading comprehension (RC), and if the assessments that we used are reliable and valid in measuring adults' vocabulary skills. We selected vocabulary abilities since these skills are crucially important to both oral and written comprehension and to writing skills as well.

Vocabulary Skills

Knowing a word is not an all or none proposition. A person may have seen a word, but not know what it means or how to use it. There are some words for which a person knows what that word means in a limited contextual situation, while there are other words for which a person has a nuanced understanding of the various meanings and contexts in which the word can be known or used. Despite these varying levels of familiarity, historically, researchers conceptualized vocabulary as a unidimensional construct, and as such, used a single measure of vocabulary knowledge to represent it in analyses. More recently, however, several researchers have recognized that vocabulary is multidimensional and is comprised of at least two different dimensions—breadth and depth (Li and Kirby 2015; Ouellette 2006; Proctor et al. 2012; Qian 1999; Read 1989; Stæhr 2009; Wesche and Paribakht 1996). These researchers have argued that vocabulary knowledge is not simply how many words you know, which represents breadth, but also depends on how well you understand the meanings of those words, which represents depth.

This account of the multidimensionality of vocabulary is grounded in various theoretical accounts of lexical organization. For example, Levelt et al. (1999) argued that a word is represented in the lexicon with a stored phonological representation and its accompanying semantic information. In this account, the number of stored phonological entries represents vocabulary breadth, while the extent of the stored semantic information represents vocabulary depth. A more contemporary account is the Lexical Quality Hypothesis (Perfetti and Hart 2001) in which it is proposed that words are represented as a bonded set of orthographic, phonological, and semantic constituents. Full specification of meaning is context dependent in ways that orthographic and phonological specification is not. Nonetheless, over repeated encounters with a word in different semantic contexts, a common core representation of the semantic constituent should emerge and be refined over time. In this model, the orthographic and phonological constituents can be thought of as the breadth dimension, while the semantic constituent conveys depth of vocabulary knowledge.

While there is a growing consensus that vocabulary is comprised of at least two dimensions, depth itself is comprised of different aspects. Some conceptions of depth refer to the quality of the semantic information contained in the lexicon for a given word. For example, in some studies, a participant is asked to reflect on how well they know a given target word. Do they recognize it? Can they figure out the meaning of the word if it is found in context? Might they be able to provide a detailed definition of the word when the target is not found in context? Another component of semantics is an appreciation of polysemy (Carlo et al. 2004). Understanding that words can have more than one meaning or sense is important. While there are many words that have distinct, multiple meanings (e.g., “diamond,” which could mean a jewel or a baseball field), there are more words that merely have different senses, and one needs to understand that a word's meaning may change slightly depending on the context (e.g., "home" could mean a place, or it could mean a sense of belonging). Finally, many researchers argue that how well words are organized in the lexicon contributes to the semantic component of depth (Henriksen 1999; Meara and Wolter 2004; Read 1998; Stæhr 2009). Henriksen (1999) argued that truly knowing a word requires not only knowing the meaning, but also knowing how that word is connected to other words in a semantic network. Words that have extensive sets of connections are thought to be known more deeply (Meara and Wolter 2004).

Other conceptualizations of depth are more expansive and outline other components that contribute to the depth of word knowledge. According to Proctor et al. (2012), in addition to semantic relations, vocabulary depth covers components such as morphological awareness and syntactic awareness. Morphological awareness consists of the recognition that words are made up of roots and affixes. For example, we may derive other words by adding prefixes and suffixes to the word "place" such as: placement, displace, misplace, placed, etc. When encountering a complex word, readers with proficient morphological skills may break the word into smaller parts to decipher its meaning, thereby, increasing their vocabulary knowledge. Knowing how to understand and/or generate different morphological forms from the same root helps to grow vocabulary. The syntactic aspect of vocabulary depth covers knowledge about the possible syntactic structures into which the word fits. For "place," depending on the context, it could be a noun or a verb. In addition, when changing the morphological structure of the word, it can also play a variety of grammatical roles (e.g., "misplaced" can be a verb or an adjective). To assess this dimension, researchers have administered a task in which a participant is shown a picture and given a target word. The participant needs to use the target in a single sentence and their score depends on whether the sentence structure is accurate (Proctor et al. 2012).

Interestingly, while researchers have described vocabulary depth in this multi-component manner, very few studies have used more than one measure to assess it, so we currently lack a full understanding of vocabulary depth, and how this construct relates to other literacy outcomes (Li and Kirby 2015). When researchers have used a single measure, they most often have assessed the semantic component of depth by asking participants to provide definitions of words (Cavalli et al. 2016; Dickinson et al. 2019; Hadley et al. 2015; McGregor et al. 2013; Ordóñez et al. 2002; Proctor et al. 2011; Swart et al. 2017; van Goch et al. 2019) or they asked participants to identify synonyms or evaluate associates of a given target word (Corrigan 2007; Enayat and Babaii 2018; Qian and Schedl 2004; Silverman et al. 2015; Stæhr 2009). A few studies have incorporated two measures of depth in their protocol (Cain and Oakhill 2014; Ouellette and Shaw 2014; Stothers 2018). However, these studies have only included definitional and synonym/associate tasks. The exception is that Stothers (2018) also used a task to measure polysemy in which participants were asked to provide as many definitions as possible for a list of homographs. There are a few studies for which the researchers used three or more tasks to measure depth (Lawrence et al. 2019; Li and Kirby 2015; Montecillo Leider et al. 2013; Ouellette 2006; Proctor et al. 2012). Importantly, Li and Kirby (2015) included tasks that assessed morphological knowledge by asking participants to identify roots in multi-morphemic words. Additionally, Richard (2011) used the Vocabulary Depth Task, which is a task that attempts to measure all three elements (morphology, semantics, and syntax). Participants are presented with sets of six sentences, each with a blank. The task is to determine one target word that can fit in all six blanks. With this test, participants need to use their morphological and syntactic awareness to adjust the target word so it suits each of the sentences. They also have to understand the word meaning in different contexts, which taps into the semantic aspect of vocabulary depth. Clearly, to gain a better understanding of the complexity of depth, we need to evaluate a full battery of tasks that may assess this construct. This is one of the goals of the current study.

Relationships Among Vocabulary Abilities and RC

While the main goal of this study was to gain a better appreciation of the structure of vocabulary abilities, we also wanted to explore how these dimensions of vocabulary were related to RC. Researchers have demonstrated that vocabulary shares a strong relationship with RC in various populations, such as in primary and secondary school children (Ouellette and Beers 2010; van Gelderen et al. 2007; Verhoeven and van Leeuwe 2008). Tannenbaum et al. (2006) also claimed that in most studies, the correlations between vocabulary knowledge and RC were positive in the range of .30–.80. This past research on reading acquisition has established a firm foundation indicating that vocabulary breadth contributes significantly to RC. Many researchers have also examined the relationship between vocabulary depth and breadth across various tasks. Ouellette (2006) investigated the effects of breadth and depth on different literacy skills in a sample of fourth grade children. Specifically, the researcher administered a battery of tests to determine participants’ levels of nonverbal intelligence, decoding, visual word recognition, RC, oral vocabulary, word definitions, and synonyms. Results showed that vocabulary breadth had significant influence on decoding skills. Having a larger vocabulary size increases a person’s familiarity with phonemic units, which results in decoding efficiency. Furthermore, results also indicated that both breadth and depth explained variance in word recognition because knowing more words and understanding each word will help with retrieval efficiency. Finally, the study demonstrated that only depth was a unique predictor of RC as a thorough understanding of a word’s meaning and its role in a context help make sense of the text.

Binder et al. (2016) examined the relationships among vocabulary breadth, vocabulary depth, RC, and reading rate by administering the Nelson Denny Reading Test of Vocabulary, Comprehension, and Reading Rate to skilled college readers. They used the Vocabulary Depth Task and the Word Families task to determine vocabulary depth. Specifically, in the Word Families task, participants were given a root word and were asked to produce as many words as possible by adding prefixes and suffixes. The results demonstrated that firstly, both vocabulary breadth and depth had a significant correlation with RC and reading rate. Secondly, both breadth and depth contributed to explaining variance in RC with contributions of 33% and 6%, respectively. However, only vocabulary breadth explained unique variance in reading rate. These findings are consistent with Ouellette’s (2006) results that different dimensions of vocabulary contribute differently to various reading skills. On the one hand, possessing a large vocabulary size helps to decode words faster, which increases reading rate. On the other hand, understanding a word thoroughly aids in choosing the appropriate meaning to link words together and establish a more coherent perception of the text, which improves RC overall.

Current Study

In the current study, we attempted to broaden our understanding of the multi-dimensional construct of vocabulary knowledge and its relationship to RC in a sample of adults with low literacy skills. We focused on deconstructing the components of vocabulary knowledge and assessed how these components are related to RC in adults. Specifically, we administered a total of 12 tasks that are known to measure vocabulary breadth, depth, and RC. Since various studies use different measures to assess breadth and depth, we first wanted to assess which tasks measured which aspects of vocabulary knowledge. First, we conducted analyses to see whether these vocabulary measures of breadth and depth were reliable and valid in examining the vocabulary level of adults with low literacy skills. We hypothesized that these measures were reliable and valid; thus, vocabulary knowledge would comprise at least two factors: breadth and depth. We further investigated to see which measures loaded the highest on the depth factor. Since morphological awareness had shown to contribute to vocabulary, we hypothesized that it might be the component that loaded the highest on the depth factor. Second, we examined whether the vocabulary depth factor explained variance in the RC of adults with low literacy after controlling for vocabulary breadth. We hypothesized that both components would make significant independent contributions to explain RC.

Method

Participants

We recruited a total of 117 adult participants who were currently enrolled in different levels of ABE programs in Massachusetts, USA. However, we eliminated the data associated with 14 participants who did not return for our second testing session. Therefore, our sample included a total of 103 adult participants ranging from 18 to 71 years old (M = 35.26, SD = 14.21), consisting of 78 females and 25 males. The sample included 47 White participants, 16 Black/African American participants, five multiracial participants, four Asian participants, three participants who identified as American Indian/Alaska Native, and 28 participants who identified with another race or did not know their race. Participants were provided with $20 for compensation. We collected reading achievement data using the Woodcock Johnson Broad Reading Scale (Woodcock et al. 2001). For the Letter-Word Identification task, participants read letters and words of increasing difficulty from a flashcard. For Word Attack, participants pronounced a series of non-words of increasing difficulty. For Reading Fluency, participants read three pages of sentences and identified whether the statements were true or false by circling “Y” or “N.” Testing was discontinued after three minutes. For Passage Comprehension, participants read a series of sentences of increasing difficulty, each with one missing word, and had to supply the missing word. From these tasks, we constructed a grade equivalent score. On average, the adults performed at a 5th grade reading level (M = 5.18, SD = 2.48), which is much lower than would be expected for a typically developing adult.

Materials

We administered 12 tasks to measure participants’ vocabulary depth, vocabulary breadth, and RC.

Word Definitions—We used the Word Definitions task taken from the Test of Word Knowledge (Wiig and Secord 1992). In this task, the experimenter presented a word in both written and spoken forms and asked participants to explain its meaning. If the meaning was ambiguous, the participant would be asked to tell more about the word. There were a total of 32 words. Each target word would be scored from zero to two, with points given for the number of semantic features, and a definition must also include the semantic category of the word. The task was discontinued if the participant made five consecutive errors or non-responses. Wiig and Secord (1992) reported reliability ranging from .85 to .96 across age groups.

Polysemy—We used the materials presented by Carlo et al. (2004) to examine participants' knowledge of polysemous words. The experimenter read a total of six lexically ambiguous words (a word with two or more possible meanings) one at a time with 30 s between each word. During these 30 s, participants had to say as many sentences as possible to convey different meanings of the words. Points were calculated based on the frequency of the meaning demonstrated in the sentences. Responses with the dominant meaning (most frequent meaning) were worth one point. Responses with the secondary meaning (less frequent meaning) were awarded two points. Responses with either a tertiary or quaternary meaning (least frequent meaning) were given three points. The points were cumulative. Finally, if the participant utilized a specific meaning more than once, the repeated meaning would not be awarded any points. The frequency order of meanings for each word was based on the results of a norming task conducted at the researchers' institution. Carlo et al. (2004) reported reliability of .64 for this task.

Semantic Category Fluency—We conducted the Semantic Category Fluency task used by Tannenbaum et al. (2006). In this task, participants were presented with eight categories (e.g. farm animals, fruits, things people drink, etc.) and instructed to list as many items in that category as possible within 10 s. Participants' scores included the total number of correct items, summed across the eight categories. Tannenbaum et al. reported good reliability of .87 in their sample.

Synonyms—The Synonyms task from the Test of Word Knowledge (Wiig and Secord 1992) was administered. Participants were shown a target word and asked to choose the synonym of that word from three or four given options. All words were presented in both written and spoken forms. There were a total of 42 target words, and testing was suspended after five consecutive errors or non-responses from the participant. Wiig and Secord report reliability ranging from .84 to .93 across age groups.

Morphological Awareness—Morphological awareness was examined by three tests. The first one, the Derivational Morphemes task, determined participants' knowledge of morphological structure through assessing their understanding of derivational morphemes (Carlisle 2000). The experimenter read the target word, followed by a sentence in which one word was replaced with the word "blank." The participant then had to derive a new word from the target word to fill in the blank.

The second morphological awareness test was the Suffix Choice task, which involved manipulating morphemes using pseudowords (Mahony 1994; Singson et al. 2000; Tyler and Nagy 1989, 1990). The experimenter gave participants a printed form of the test consisting of 14 items. Each item included a sentence with a blank and four answer choices. For example, for the sentence "Our teacher taught us how to _____ long words," the participant had to choose the correct answer among four pseudowords: *jittling, jittles, jittled*, and *jittle*. The experimenter read the sentence and answer choices aloud to the participant to eliminate the dependence on participants' decoding ability. For both tasks, testing was terminated after six incorrect answers. Tighe and Binder (2015) reported the reliabilities of .94 and .86 for these two tasks, respectively.

The last morphological awareness test was the Word Families task. The experimenter presented participants with 10 root words, one at a time, and requested that they say as many derivatives of each target word as possible in 30 s. Each correct word participants produced was worth one point, and the total score was summed across the 10 root words.

VDT—We administered the Vocabulary Depth Task (VDT) designed by Richard (2011). Participants were given 30 sets of six sentences with a blank in each sentence. Participants had to find a target word that could fit into all six of the sentences within each set. For each correct target word identified, the participant scored a point. An example of this task is as follows: Target Word: Lose; A. I hear Linda is going to [] her job.; B. I do not want to [] this game.; C. To [] your family so young is sad.; D. I want to [] ten pounds, so I am exercising more.; E. Do not [] time in attacking your enemy.; F. Relax! Do not [] it.

PPVT—The Peabody Picture Vocabulary Test-Third Edition (PPVT) is a receptive task (Dunn and Dunn 1997). In this task, the experimenter read a word while showing the participant a set of four pictures. The participant pointed to the picture that conveyed the meaning of that word. Testing was discontinued when the participant made eight or more mistakes within a set. There were 12 vocabulary words in each set. Dunn and Dunn report a split-half reliability of .94.

PVT—The Picture Vocabulary Test (PVT) is an expressive task taken from the Woodcock Johnson III Achievement Test (Woodcock et al. 2001). Participants had to name the picture that the experimenter showed them. The level of difficulty increased as the participant progressed through the task, and testing was continued until the participant made six consecutive errors. The reported reliability for this task is .81.

RC Assessments

We administered two tests to assess RC. The first one was the Passage Comprehension task from the Woodcock Johnson III Achievement Test (Woodcock et al. 2001). This test indicates participants' ability to utilize contextual clues to fit a missing word into a sentence. The sentence was presented to the participant in written form while the experimenter read it aloud to them, and participants responded orally. The experimenter continued the task until the participant made six consecutive errors. Woodcock et al. report a reliability of .88 for this task.

The second test was the Test of Sentence Reading Efficiency and Comprehension (TOSREC), which was reported to have good reliability (.93) and construct validity (.87–.89) (Wagner et al. 2010). The task included ninety-eight items and two example items. With each item, the participant read and determined if a statement was true or false. Within 2 min, the participant tried to read and answer as many items as possible.

Procedure

The experimenter started by interviewing participants to collect demographic data. Then, all of the assessments were conducted over the course of two sessions which lasted between 40 and 60 min each with breaks as needed. The process took place over 2 days in a room at participants' instructional sites.

Results

The 12 variables represent scores from 10 tests of vocabulary and two tests of RC. Descriptive statistics are presented in Table 1.

Cronbach's Reliability

Cronbach's reliability estimates were calculated for all 12 measures. All 12 tasks were found to be highly reliable, with alpha values ranging from $\alpha = .764$ to $\alpha = .987$ (see Table 2). The Polysemy test was reported to have the lowest reliability ($\alpha = .764$) amongst all the tests. However, all alpha values were in the moderate to high range of reliability, which indicated a sufficient level of reliability for assessing vocabulary knowledge. Similarly, Passage Comprehension and Reading Fluency (TOSREC) also had very high reliability of over .913.

Correlation Analysis

Correlation analysis was used to examine the relationship among 10 vocabulary tests before we conducted the factor analysis. The results indicated all of the tasks significantly and positively correlated with one another (see Table 3). The highest correlation was the relation between the PPVT and Picture Vocabulary Test (PVT) (r =.905, $p < .001$), while the lowest was the correlation between the Suffix Choice task and Semantic Fluency task (r =.481, $p < .001$). With all of the assessments significantly positively correlated, results from this correlation analysis suggested that factor analysis could produce distinct and reliable factors. However, the PPVT task was highly correlated (rs > .80) with Word Definitions, Synonyms, and the PVT, which indicated a problem of multicollinearity for our factor analysis. Therefore, we excluded the PPVT in our factor analysis model.

Factor Analysis for Vocabulary Assessments

We examined the factorability of nine vocabulary tests. Several well-recognized criteria for the factorability were used. Firstly, all the tests correlated significantly and positively with one another with a range from $r = .517$ to $r = .798$, indicating reasonable factorability. Secondly, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was .940 (above the recommendation of .60), which implied that our data were suitable for factor analysis. The Bartlett's test of sphericity was significant ($\chi^2 = 845.634$, $p < .001$), which implied that there were patterned relationships between the variables. The diagonals of the anti-image correlation matrix were all over .933, which supported the inclusion of each task in the factor analysis. Moreover, the communalities were all above .753, further confirming that each test shared some common variance with other tests. Given these overall indicators, factor analysis was conducted with all nine variables.

We examined several indicators to determine how many factors to extract. There was one eigenvalue that exceeded 1, (6.522), but examination of the scree plot indicated that the plot did not level until after the second factor. When we ran the factor analysis that would produce one factor and examined the residual correlation matrix, we found that there was a large percentage (47%) of non-redundant residuals with absolute values greater than .05. This suggested that additional factors should be extracted. When we specified two factors, the percentage of non-redundant residuals fell to 25%, which seemed more reasonable. Given the scree plot and the examination of the non-redundant residuals, we decided that it would be appropriate to extract two factors.

Data were subjected to factor analysis using principle component analysis and direct oblimin rotation method. We used principle component analysis because our primary purpose was to identify common factors underlying the tests and compute composite scores for the components. Direct oblimin rotation was used to allow for correlation as we suspected these factors were breadth and depth, which positively correlated with each other. By specifying two factors, we explained 81.27% of the total variance. Specifically, the initial eigenvalues showed that the first factor explained 72.47% of the variance, while the second factor added another 8.80% of the variance.

Results from the pattern matrix indicated that there were two factors constructing vocabulary, which were vocabulary depth and vocabulary breadth (factor 1 and factor 2, respectively). See Table 4. The Target Word Task (VDT), Derivational Morphology task, Word Families, and Suffix Choice all had loadings above .694 for the depth factor. Word Definitions, Polysemy, Semantic Fluency, and the PVT all had loadings above .637 for the breadth factor. Examination of the pattern matrix shows that the Synonym task had a loading on the depth factor of .511 and a cross-loading of .492 on the breadth factor. Thus, in a subsequent analysis we eliminated this task.

We conducted a principle component factor analysis of the remaining eight vocabulary tests, with direct oblimin rotations, with the two factors explaining 81.171% of the variance. All assessments had primary loadings over .64. The factor loading matrix for this final solution is presented in the two right-most columns of Table 4. Four assessments, including Suffix Choice, Target Word Task, Word Families, and Derivational Morphemes loaded on vocabulary depth (factor 1) with loading over .69. Word Definitions, Polysemy, Semantic Fluency, and PVT loaded on vocabulary breadth (factor 2) with loading over .698.

Regression Analysis

Composite scores were created for each of the two factors using the multiple regression approach. We also created a composite score with the two measures that we had for RC (WJ Passage Comprehension and TOSREC Reading Fluency). A hierarchical multiple regression analysis was conducted to examine whether vocabulary breadth and vocabulary depth significantly explained participants' RC. A two stage hierarchical multiple regression was run with RC Factor as the dependent variable. Vocabulary Breadth Factor was entered at stage one of the regression to control for the influence of vocabulary breadth knowledge. Vocabulary Depth Factor was entered at stage two of the regression to identify if vocabulary depth contributed additional variance to explain RC beyond the effect of vocabulary breadth. The hierarchical multiple regression revealed that at stage one, Vocabulary Breadth contributed significantly to the regression model, $F(1, 101) = 208.635$, $p < .001$ with an R^2 of .674. Introducing the Vocabulary Depth Factor variable explained an additional 11.8% of variation in RC Factor and this change in R^2 was significant, $F(1, 100) = 56.607$, $p < .001$. Together, the two independent variables accounted for 79.2% of the variance in RC Factor. The β coefficients for both of the predictors were significant (vocabulary breadth: $\beta = .509$, $t = 8.261$, $p < .001$; vocabulary depth: $\beta = .464$, $t = 7.524$, $p < .001$).

Discussion

The purpose of this study was to examine the factor structure of vocabulary and to assess the relationship between vocabulary breadth, vocabulary depth, and RC in adults with low literacy skills. We had several important findings. First, we confirmed that all of the assessments were highly reliable for adults. Second, the results of the factor analysis indicated two distinct vocabulary factors. Finally, both breadth and depth contributed independently to explaining variance in RC.

Vocabulary Knowledge and RC Assessments Validity

After the data were subjected to factor analysis, we were able to find two underlying factors among ten vocabulary tests: vocabulary breadth and vocabulary depth, which is consistent with findings from past literature about the construct of vocabulary knowledge (Binder et al. 2016; Nurweni and Read 1999; Ouellette 2006; Qian 1999). Our results indicated that Suffix Choice, Target Word Task, Word Families, and Derivational Morphemes loaded on vocabulary depth. For the breadth factor, in addition to the Picture Vocabulary Test, vocabulary breadth also comprised Semantic Category Fluency, Polysemy, and Word Definitions. Contrary to our initial prediction, the Synonyms task was shown to assess not only vocabulary depth but also vocabulary breadth. One surprising finding was that the Semantic Category Fluency, Polysemy, and Word Definition tasks, which are usually thought to measure vocabulary depth, loaded highly on vocabulary breadth (Tannenbaum et al. 2006). One possible explanation could be due to the testing and scoring procedures. For Semantic Category Fluency, participants are asked to name as many words as possible in a given category within 10 s. Since our categories are familiar to participants, they only need to understand the basic meaning of the words to list them under a category. Thus, while the task is designed to assess participants' depth of knowledge of each category, it might actually ask participants how many words they know under a category, which might tap on breadth rather than vocabulary depth. In addition, a participant's score is based on the total amount of correct words listed without taking into account the difficulty levels of the words. A participant who lists more common words might not understand the category as well as one that produces lower frequency words. For Word Definitions, participants' scores were determined based on a scale from 0 to 2. Since the range of the scale was small, it might not be able to differentiate the depth aspect of word understanding among participants. Therefore, researchers could consider adjusting the tests to address these shortcomings if they would like to use them to assess depth.

The results showed that Suffix Choice, Target Word Task, Word Families, and Derivational Morphemes loaded on vocabulary depth. This supported the idea that vocabulary depth consists of different components including morphological awareness, semantic relations, and syntactic awareness (Proctor et al. 2012). Our findings revealed that all three morphological awareness tasks highly loaded on vocabulary depth, which further emphasizes the important contribution of the measurement of morphological awareness to the depth of vocabulary as presented in past literature (Fracasso et al. 2016; Keiffer and Lesaux 2012; To et al. 2016). Fracasso et al. (2016) found that morphological awareness was a unique predictor of spelling ability, vocabulary, and listening comprehension in adult learners. By understanding the morphological structure of a word, readers could manipulate affixes to derive different word meanings. Our finding also supports the idea that morphological awareness indirectly contributes to RC through vocabulary (Fracasso et al. 2016; Keiffer and Lesaux 2012; Nagy et al. 2006). When encountering a complex word, readers with proficient morphological skills may break it into smaller parts to decipher its meaning, thereby enriching their vocabulary and contributing to RC.

Relationship Among Vocabulary Breadth, Vocabulary Depth, and RC

Several studies have investigated the relationship among vocabulary breadth, depth, and RC. Akbarian and Alavi (2013) demonstrated that vocabulary breadth contributed significantly to participants' IELTS and TOEFL (i.e., tests that assess English language proficiency) RC scores in their sample. Our findings are consistent with those studies, and they are also in line with other studies that indicated that vocabulary breadth significantly explains variance in RC (Carlisle 2000; Hall et al. 2014; Nelson and Stage 2007; Ouellette 2006). After vocabulary depth was added to the regression model, the results demonstrated that vocabulary depth measures explain another 11.8% of variance in the RC of adults with low literacy after controlling for vocabulary breadth. These findings support our second hypothesis that both components make significant independent contributions to explain RC. They are also consistent with other studies and add to past literature by extending the significant contribution of both vocabulary breadth and depth to RC in the adult beginning reader population (Binder et al. 2016; Nation and Snowling 2004; Qian 1999). While knowing more words (vocabulary breadth) helps readers decode and process words more efficiently, understanding more about each individual word (vocabulary depth) assists readers in making sense of the whole text, thereby aiding RC (Binder et al. 2016).

Limitations

We were only able to analyze data from 103 adults in our factor analysis for 10 variables. Although some researchers suggested the subjects-to-variables ratio of above five is sufficient, others recommended a ratio of 10:1 (as cited in Garson 2008). However, research has indicated that these "general rules" concerning minimal sample sizes are not valid or useful (MacCallum et al. 1999; Preacher and MacCallum 2002). There are other factors related to the analysis that seem to matter more. "As long as communalities are high, the number of expected factors is relatively small, and model error is low (a condition which often goes hand-in-hand with high communalities), researchers and reviewers should not be overly concerned about small sample sizes" (Preacher and MacCallum 2002, p. 160). The communalities in our study were all quite high. All of these conditions were met in the current project, so we have confidence in the results.

Implications and Future Studies

As has become apparent over the past many years, vocabulary can no longer be thought of as a unidimensional construct. Rather, there is now a substantial body of evidence that persuasively demonstrates that vocabulary has at least two dimensions: breadth and depth (Li and Kirby 2015; Ouellette 2006; Proctor et al. 2012; Qian 1999; Read 1989; Stæhr 2009; Wesche and Paribakht 1996). Moreover, the depth component needs to be appreciated for its complexity. As Proctor et al. (2012) argued, depth consists of semantic, morphological, and syntactic components. Thus, for us to gain a better understanding of how depth is related to RC, vocabulary acquisition, and writing quality, we need to make sure we are measuring all of these aspects. This study identified several tasks that can be used in future studies to more accurately assess the relationship between these constructs and other, higher-level literacy tasks.

While we conducted a factor analysis to empirically establish a two-factor solution to vocabulary, future studies could explore the multidimensionality of the depth factor. That is, as noted above, Proctor et al. (2012) have argued that depth is comprised of semantic, morphological, and syntactic components. Perhaps a better measurement model would be to include several different measures of each of these components, which may capture nuances as to how each of these components may be related to other behaviors, similar to how depth and breadth have independent contributions to different processes such as word recognition, reading speed, and RC.

Funding

The project described was supported by Grant Number R15HD067755 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development awarded to the last author. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health & Human Development or the National Institutes of Health.

References

Akbarian I, & Alavi SM (2013). Comparing the contribution of vocabulary breadth to IELTS and TOEFL reading subtests. Porta Linguarum, 20, 135–151.

Binder KS, Gilbert Cote N, Lee C, Bessette E, & Vu H (2016). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. Journal of Research in Reading, 40, 333–343. [PubMed: 28717257]

Boyes ME, Leitao S, Claessen M, Badcock NA, & Nayton M (2016). Why are reading difficulties associated with mental health problems? Dyslexia, 22, 263–266. [PubMed: 27465209]

Cain K, & Oakhill J (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? L'année psychologique/Topics in Cognitive Psychology, 114, 647–662.

Carlisle JF (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. Reading and Writing, 12, 169–190.

Carlo MS, August D, McLaughlin B, Snow CE, Dressler C, Lippman DN, et al. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. Reading Research Quarterly, 39, 188–215.

Cavalli E, Casalis S, El Ahmadi A, Zira M, Poracchia-George F, & Colé P (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. Research in Developmental Disabilities, 51–52, 89–102.

Corrigan R (2007). An experimental analysis of the affective dimensions of deep vocabulary knowledge used in inferring the meaning of words in context. Applied Linguistics, 28, 211–240.

Dickinson DK, Nesbitt KT, Collins MF, Hadley EB, Newman K, Rivera BL, et al. (2019). Teaching for breadth and depth of vocabulary knowledge: Learning from explicit and implicit instruction and the storybook texts. Early Childhood Research Quarterly, 47, 341–356.

Dunn LM, & Dunn LM (1997). Peabody picture vocabulary test—third edition [assessment instrument]. Circle Pines, MN: American Guidance Service.

Enayat MJ, & Babaii E (2018). Reliable predictors of reduced redundancy test performance: The interaction between lexical bonds and test takers' depth and breadth of vocabulary knowledge. Language Testing, 35, 121–144.

Fracasso LE, Bangs K, & Binder KS (2016). The contributions of phonological and morphological awareness to literacy skills in the adult basic education population. Journal of Learning Disabilities, 49, 140–151. [PubMed: 24935886]

Garson DG (2008). Factor analysis: Statnotes. Encorewiki. Retrieved January 9, 2017, from https://www.encorewiki.org/download/attachments/25657/Factor+Analysis_+Statnotes+from+North+Carolina+State+University.pdf

Grotlüschen A, Mallows D, Reder S, & Sabatini J (2016). Adults with low proficiency in literacy or numeracy. OECD Education Working Papers, 131, pp. 1–151.

Hadley EB, Dickinson DK, Hirsh-Pasek K, Golinkoff R, & Nesbitt KT (2015). Examining the acquisition of vocabulary knowledge depth among preschool students. Reading Research Quarterly, 51, 181–198.

Hall R, Greenberg D, Gore JL, & Pae HK (2014). The relationship between expressive vocabulary knowledge and reading skills for adult struggling readers. Journal of Research in Reading, 37, 87–100. [PubMed: 24778459]

Henriksen B (1999). Three dimensions of vocabulary development. Studies in Second Language Acquisition, 21, 303–317.

Keiffer MJ, & Lesaux NK (2012). Direct and indirect roles of morphological awareness in the English reading comprehension of native English, Spanish, Filipino, and Vietnamese speakers. Language Learning, 62, 1170–1204.

Lawrence JF, Hagen AM, Hwang JK, Lin G, & Lervåg A (2019). Academic vocabulary and reading comprehension: Exploring the relationships across measures of vocabulary knowledge. Reading and Writing, 32, 285–306.

Levelt WJM, Roelofs A, & Meyer AS (1999). A theory of lexical access in speech production. Behavioral and Brain Sciences, 22, 1–38. [PubMed: 11301520]

Li M, & Kirby JR (2015). The effects of vocabulary breadth and depth on English reading. Applied Linguistics, 36, 611–634.

MacCallum RC, Widaman KF, Zhang S, & Hong S (1999). Sample size in factor analysis. Psychological Methods, 4, 84–99.

Mahony DL (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. Reading and Writing, 6, 19–44.

McGregor KK, Oleson J, Bahnsen A, & Duff D (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. International Journal of Language & Communication Disorders, 48, 307–319. [PubMed: 23650887]

Meara P, & Wolter B (2004). V-links: Beyond vocabulary depth In Albrechtsen D, Haastrup K, & Henriksen B (Eds.), Writing and vocabulary in foreign language acquisition (pp. 85–96). Copenhagen: Museum Tusculanum Press.

Montecillo Leider C, Proctor CP, Silverman RD, & Harring JR (2013). Examining the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. Reading and Writing, 26, 1459–1485.

Nagy W, Berninger V, & Abbott R (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. Journal of Educational Psychology, 98, 134–147.

Nation K, & Snowling MJ (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. Journal of Research in Reading, 27, 342–356.

National Center for Education Statistics. (2003). National assessment of adult literacy. Retrieved September 17, 2016, from https://nces.ed.gov/naal/kf_demographics.asp

Nelson JR, & Stage SA (2007). Fostering the development of vocabulary knowledge and reading comprehension through contextually-based multiple meaning vocabulary instruction. Education and Treatment of Children, 30, 1–22.

Nurweni A, & Read J (1999). The English vocabulary knowledge of Indonesian university students. English for Specific Purposes, 18, 161–175.

Ordóñez CL, Carlo MS, Snow CE, & McLaughlin B (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? Journal of Educational Psychology, 94, 719–728.

Organization for Economic Cooperation and Development [OECD]. (2013). Time for the U.S. to reskill?: What the survey of adult skills says. Retrieved September 17, 2016, from 10.1787/9789264204904-en

Ouellette G (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. Journal of Educational Psychology, 98, 554–566.

Ouellette G, & Beers A (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. Reading and Writing, 23, 189–208.

Ouellette G, & Shaw E (2014). Oral vocabulary and reading comprehension: An intricate affair. L'année psychologique/Topics in Cognitive Psychology, 114, 623–645.

Perfetti CA, & Hart L (2001). The lexical basis of comprehension skill In Gorfein DS (Ed.), Decade of behavior. On the consequences of meaning selection: Perspectives on resolving lexical ambiguity (pp. 67–86). Washington, DC, USA: American Psychological Association.

Preacher KJ, & MacCallum RC (2002). Exploratory factor analysis in behavior genetics research: Factor recovery with small sample sizes. Behavior Genetics, 32, 153–161. [PubMed: 12036113]

Proctor CP, Dalton B, Uccelli P, Biancarosa G, Mo E, Snow C, et al. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. Reading and Writing, 24, 517–544.

Proctor CP, Silverman RD, Harring JR, & Montecillo C (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish-English bilingual children in elementary school. Reading and Writing, 25, 1635–1664.

Qian DD (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. The Canadian Modern Language Review, 56, 282–307.

Qian DD, & Schedl M (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. Language Testing, 21, 28–52.

Read J (1989). Towards a deeper assessment of vocabulary knowledge Paper presented at the 8th annual meeting of the international association of applied linguistics (Sydney, NSW, Australia, August 16–21, 1987). ED 301048. Washington, DC: ERIC Clearing House on Languages and Linguistics.

Read J (1998). Validating a test to measure depth of vocabulary knowledge In Kunnen A (Ed.), Validation in language assessment (pp. 41–60). Hillsdale, NJ: Erlbaum.

Richard JPJ (2011). Does size matter? The relationship between vocabulary breadth and depth. Sophia International Review, 33, 107–120.

Silverman RD, Proctor CP, Harring JR, Hartranft AM, Doyle B, & Zelinke SB (2015). Language skills and reading comprehension in English monolingual and Spanish–English bilingual children in grades 2–5. Reading and Writing, 28, 1381–1405.

Singson M, Mahony D, & Mann V (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. Reading and Writing, 12, 219–252.

Stæhr LS (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. Studies in Second Language Acquisition, 31, 577–607.

Stothers ME (2018). Breadth over depth in the semantic representations of adults with nonverbal learning disabilities. Language and Cognition, 10, 56–82.

Swart NM, Muijselaar MML, Steenbeek-Planting EG, Droop M, de Jong PF, & Verhoeven L (2017). Differential lexical predictors of reading comprehension in fourth graders. Reading and Writing, 30, 489–507. [PubMed: 28255199]

Tannenbaum KR, Torgesen JK, & Wagner RK (2006). Relationships between word knowledge and reading comprehension in third grade children. Scientific Studies of Reading, 10, 281–298.

Tighe EL, & Binder KS (2015). An investigation of morphological awareness and processing in adults with low literacy. Applied Psycholinguistics, 36, 245–273. [PubMed: 25926711]

To NL, Tighe EL, & Binder KS (2016). Investigating morphological awareness and the processing of transparent and opaque words in adults with low literacy skills and in skilled readers. Journal of Research in Reading, 39, 171–188. [PubMed: 27158173]

Tyler A, & Nagy WE (1989). The acquisition of English derivational morphology. Journal of Memory and Language, 28, 649–667.

Tyler A, & Nagy WE (1990). Use of derivational morphology during reading. Cognition, 36, 17–34. [PubMed: 2383968]

van Gelderen A, Schoonen R, Stoel RD, de Glopper K, & Hulstijn J (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. Journal of Educational Psychology, 99, 477–491.

van Goch MM, Verhoeven L, & McQueen JM (2019). Success in learning similar-sounding words predicts vocabulary depth above and beyond vocabulary breadth. Journal of Child Language, 46, 184–197. [PubMed: 30236168]

Verhoeven L, & van Leeuwe J (2008). Prediction of the development of reading comprehension: A longitudinal study. Applied Cognitive Psychology, 22, 407–423.

Wagner RK, Torgesen JK, Rashotte CA, & Pearson NA (2010). Test of silent reading efficiency and comprehension. [Assessment instrument]. Austin, TX: Pro-ED Inc.

Wesche M, & Paribakht TS (1996). Assessing second language vocabulary knowledge: Depth versus breadth. The Canadian Modern Language Review, 53, 13–40.

Wiig EH, & Secord WA (1992). Test of word knowledge. [Assessment instrument]. San Antonio, TX: The Psych Corporation.

Woodcock RW, McGrew KS, & Mather N (2001). The Woodcock–Johnson test of achievement—Third Edition. [Assessment instrument]. Itasca, IL: Riverside Publishing.

Table 1

Descriptive statistics for 12 assessments

	N	Minimum	Maximum	Mean	SD
Passage comprehension	103	12	40	29.11	6.553
Reading fluency (TOSREC)	103	3	76	32.01	13.939
Word definitions	103	0	49	24.86	12.677
Synonyms	103	1	41	26.52	10.410
Polysemy	103	0	34	16.46	7.802
Semantic category fluency	101	9	54	35.94	8.997
Target word task (VDT)	103	0	21	6.18	5.001
Derivational morpheme	103	0	33	13.11	11.200
Word families	103	0	38	16.72	7.594
Suffix choice	103	0	14	7.17	4.408
Picture vocabulary test (PVT)	103	11	41	29.02	7.252
PPVT	103	52	185	134.87	37.459

Table 2

Cronbach's alpha of 12 measures

	α	No. of items
Derivational morpheme	.972	33
Suffix choice	.891	14
Target word task (VDT)	.873	30
Synonyms	.950	42
Word definitions	.922	32
Word families	.885	10
Polysemy	.764	6
Semantic category fluency	.790	8
Picture vocabulary test (PVT)	.923	54
PPVT	.987	204
Passage comprehension	.913	47
Reading fluency (TOSREC)	.968	98

Table 3

Correlation between 10 vocabulary measures

	Word definitions	Synonyms	Polysemy	Semantic fluency	Target word	Word families	Suffix choice	Derive morph	PVT score	PPVT
Word definitions										
Synonyms	.797**									
Polysemy	.742**	.740**								
Semantic fluency	.710**	.684**	.652**							
Target word VDT	.691**	.736**	.646**	.517**						
Word families	.680**	.791**	.713**	.586**	.712**					
Suffix choice	.620**	.709**	.584**	.481**	.753**	.739**				
Derive morph	.680**	.759**	.688**	.519**	.782**	.701**	.713**			
PVT score	.798**	.796**	.761**	.770**	.631**	.685**	.575**	.631**		
PPVT	.827**	.864**	.795**	.770**	.651**	.711**	.618**	.651**	.905**	

** < .001

Table 4

Pattern matrix loading

	1	2	1	2
Word definitions	.264	.699	.267	.698
Synonyms	.511	.492	–	–
Polysemy	.299	.637	.303	.641
Semantic fluency	− .160	1.014	− .152	1.011
Target word task (VDT)	.896	.017	.894	.025
Word families	.694	.242	.692	.244
Suffix choice	.976	− .109	.972	− .101
Derivational morpheme	.831	.088	.828	.093
PVT	.104	.857	.108	.854