



Published in final edited form as:

Read Writ. 2015 September ; 28(7): 959–988. doi:10.1007/s11145-015-9557-0.

Examining the Underlying Dimensions of Morphological Awareness and Vocabulary Knowledge

Mercedes Spencer, Andrea Muse, Richard K. Wagner, Barbara Foorman, Yaacov Petscher, Christopher Schatschneider, Elizabeth L. Tighe, and M. Denise Bishop

Florida State University and the Florida Center for Reading Research

Abstract

We report results from two studies on the underlying dimensions of morphological awareness and vocabulary knowledge in elementary-aged children. In Study 1, 99 fourth-grade students were given multiple measures of morphological awareness and vocabulary. A single factor accounted for individual differences in all morphology and vocabulary assessments. Study 2 extended these results by giving 90 eighth-grade students expanded measures of vocabulary and morphology that assessed (a) definitional knowledge, (b) usage, (c) relational knowledge, and (d) knowledge of morphological variants, with each potential aspect of knowledge assessed using an identical set of 23 words to control for differential knowledge of specific vocabulary items. Results indicated that a single-factor model that encompassed morphological and vocabulary knowledge provided the best fit to the data. Finally, explanatory item response modeling was used to investigate sources of variance in the vocabulary and morphological awareness tasks we administered. Implications for assessment and instruction are discussed.

Keywords

vocabulary knowledge; morphological awareness; structural equation modeling; item response modeling

Morphological awareness is defined as the ability to recognize and manipulate morphemes (Carlisle, 1995). Morphemes provide information about a word's meaning, spelling, and pronunciation (Carlisle, 2003). Because orthographic representation in English is morphophonemic, morphological information provides important cues for word pronunciation as well as information regarding semantic relationships (Chomsky & Halle, 1968). Given this, it is no surprise that morphological awareness has been identified as an important predictor of spelling, text processing speed, word reading, and reading comprehension proficiency (Carlisle, 1995, 2000, 2003; Carlisle & Nomanbhoy, 1993; Carlisle & Stone, 2005; Deacon & Kirby, 2004; Elbro & Arnbak, 1996; Kieffer & Lesaux, 2008; Kirby, Parrilla, Wade-Woolley, & Deacon, 2009; Kirby et al., 2012; Kuo & Anderson, 2006; Mann & Singson, 2003; Nagy, Berninger, & Abbott, 2006; Nunes, Bryant,

& Bindman, 2006; Roth, Lai, White, & Kirby, 2006; Singson, Mahoney, & Mann, 2000; Tong et al., 2011; Treiman & Cassar, 1996).

Morphological awareness has a fairly predictable developmental trajectory (e.g., Kuo & Anderson, 2006). Children's use of inflectional morphemes appears to develop prior to their use of derivational morphemes (Adams, 1990; Carlisle, 2003; Kuo & Anderson, 2006); most young children demonstrate an ability to only successfully derive words that are phonetically (e.g., *hand/handful*) and semantically transparent (e.g., *happy/unhappy*) (Carlisle, 1995). Productive words (i.e., verbs) are often easier to derive as well (Carlisle, 1995). Although kindergarteners and first graders demonstrate a mastery of inflectional morphological awareness, their knowledge remains relatively implicit, and this may explain why younger children often have problems successfully completing morphological awareness tasks (Carlisle, 1995).

Across studies, a number of different morphological awareness measures have been used. However, relatively little attention has been paid to whether possible differences in methods may explain differences in the predictive contributions of these skills to other literacy-based skills (see Deacon, Parilla, & Kirby, 2008 and Kirby et al., 2012 for reviews of morphological awareness measures). Several dimensions of interest include whether tasks require individuals to generate an oral response as opposed to a multiple-choice format and whether task administration is oral or written. For instance, tasks that have a written administration format may be tapping an individual's reading in addition to morphological awareness because such tasks require reading to complete the items. Similarly, individual differences in writing or spelling may affect performance on tasks that require written responses. Thus, such differences among various morphological awareness tasks may affect the predictive relations between morphological awareness and other literacy-based skills, such as is the case for other literacy skills (e.g., Jenkins, Johnson, & Hileman, 2004).

A recent investigation by Apel, Diehm, and Apel (2013) demonstrated that students' performance on different morphological awareness assessments were differentially predictive of their reading skills. The authors assessed morphological awareness in 156 kindergarten, first-, and second graders using multiple experimenter-created measures that tapped morphological production, identification, and judgment. Tasks measured students' inflectional and derivational knowledge and included both real words and nonwords. For second graders, two tasks measuring morphological production and judgment were significantly predictive of performance on tasks that assessed sight word reading, pseudo word decoding, and silent reading and comprehension; however, a task that measured morphological identification (affix identification task) was not significantly predictive of performance for any of the same reading measures. These differences may have emerged because these tasks are tapping different theoretical constructs but these findings may also be due to the fact that the affix identification task was a written task whereas the other two tasks were oral response tasks (Apel et al., 2013).

Relations between Morphological Awareness and Vocabulary

More words are learned incidentally when they are heard or read than are taught directly, and morphology and context are the two sources of information available for learning new words (Carlisle, 2007; Nagy & Scott, 2000). Nagy and Anderson (1984) estimated that 60% of the unfamiliar words a reader encounters in text have meanings that can be predicted on the basis of their component morphemes. Anglin (1993) estimated that between first and fifth grade, students learned an average of 1,100 root words per year but an average of 3,500 derived words per year.

Morphological awareness and vocabulary knowledge are related (Anglin, 1993; McBride-Chang, Wagner, Muse, Chow, & Shu, 2005; Nagy & Anderson, 1984; Wagner, Muse, & Tannenbaum, 2007); however the nature of their relation may change over time (Carlisle, 1995; Singson et al., 2000; Nagy, Berninger, & Abbot, 2006). For example, Nagy, Berninger, Abbott, Vaughn, and Vermeulen (2003) gave multiple measures of morphological awareness, orthographic knowledge, reading, and writing, and also single measures of oral vocabulary and phonological awareness to second and fourth graders. Morphological awareness was more strongly related to vocabulary than to any other construct measured in both second ($r = .60$) and fourth ($r = .80$) grades. Note that morphological awareness and vocabulary become more related across grades (for similar findings, see also Kirby et al., 2012 and McBride-Chang et al., 2005). This finding may result from the fact that the English language is morphophonemic, and that as children's knowledge of morphological structure becomes more sophisticated, they are better able to use morphological information in a way that aids their acquisition of new vocabulary words (Nunes, Bryant, & Bindman, 2006).

Despite the strong relations between morphological awareness and vocabulary knowledge detailed above, the two skills may become less related and more distinct over time. For example, Deacon and colleagues (2014) measured vocabulary and morphological awareness in 100 English-speaking children from third to fourth grade. Children's vocabulary knowledge (as indexed by a modified version of the Peabody Picture Vocabulary Test) was more strongly correlated with a word analogy task in third grade than in fourth grade ($r = .47$ and $.37$, respectively) (see also Deacon, Benere, & Pasquarella, 2013; Kieffer & Lesaux, 2012b). Although quality classroom instruction that focuses specifically on building morphological awareness has been shown to affect the development of both morphological awareness and vocabulary knowledge (Carlisle, 2010; Lesaux, Kieffer, Faller, & Kelley, 2010), differences in literacy-based experiences may be one explanation for why morphological awareness and vocabulary become more separable as children get older.

Variability associated with morphological awareness instruction as well as differing experiences related to print exposure may affect the relation between these skills because morphological awareness and reading comprehension are reciprocally related (Deacon et al., 2014; Kruk & Bergman, 2013). For instance, Nunes et al. (2006) examined relations between spelling ability and morphological awareness in 363 children from years (i.e., grade levels) 2, 3, and 4. Children were given a spelling task and two measures of morphological awareness that included word and sentence analogy tasks; the two morphological awareness

tasks were administered again a year later. Results indicated that performance on the spelling task was a significant longitudinal predictor of performance on tasks measuring morphological awareness. Thus, experience with printed text may be further exemplified by differences in writing experience because formal writing instruction tends to occur primarily during the later grades. This would also support previous findings that vocabulary and morphological awareness become more distinct as children get older (Deacon et al., 2013, 2014; Kieffer & Lesaux, 2012b).

Overall, past research has demonstrated that morphological awareness and vocabulary are related constructs and that morphological awareness explains unique variance in vocabulary (Anglin, 1993; Carlisle, 1995; Kieffer & Box, 2013; McBride-Chang et al., 2005; Nagy & Anderson, 1984; Nagy et al., 2003; Singson et al., 2000; Wagner et al., 2007); however, these skills tend to be measured by a single indicator, meaning that, in most instances, the reported correlations are likely to underestimate the true correlation between morphological awareness and vocabulary because of measurement error. One way to solve this issue is examine these relations using latent variable analyses, and several studies have explicitly investigated the factor structure of vocabulary and morphological awareness. For instance, Tannenbaum, Torgesen, and Wagner (2006) examined the structure of vocabulary knowledge in 203 third graders. Six standardized vocabulary measures were administered, and confirmatory factor analyses and structural equation modeling were used to test alternative models of vocabulary knowledge and also relations between vocabulary and reading comprehension. The results supported a two-factor model of vocabulary knowledge consisting of a breadth factor and a combined depth/fluency factor. Although breadth was found to explain more unique variance in reading comprehension scores than did the combined depth/fluency factor, the factors were found to be highly related ($r = .87$).

More recently, Kieffer and Lesaux (2012a, 2012b) carried out two investigations examining relations between morphological awareness and vocabulary. In one study, Kieffer & Lesaux, (2012a) examined these literacy-related skills in 90 Spanish-speaking language minority students. Data were collected longitudinally, following participants from fourth through seventh grade. Combinations of standardized and researcher-created measures were used, and data collection occurred at four time points. Growth curve modeling showed that growth in vocabulary was correlated with growth in morphological knowledge. In their second investigation, Kieffer and Lesaux (2012b) examined relations between morphological awareness and vocabulary knowledge in 584 sixth graders. Participants included English language learners and their English-speaking peers. The authors assessed vocabulary and morphological awareness at two time points, in October and again three months later in December. Tasks tapped relational and semantic knowledge, words with multiple meanings, the ability to use contextual information to derive word meanings, and morphological decomposition and derivation tasks that included real and nonwords, respectively. Multigroup confirmatory factor analysis indicated that knowledge about words was best represented as three factors – breadth, contextual sensitivity, and morphological awareness – for both language groups.

In this article we report the results of two studies. The focus of the first study was twofold. The first aim was to examine the underlying dimensions of commonly used morphological

awareness tasks. The second aim of the first study was to investigate relations between morphological awareness and vocabulary. The focus of the second study was to extend the findings of the first study in three primary ways. The first was to examine the relations between morphological awareness and vocabulary in older children. The second was to extend the measures of vocabulary from definitional measures of breadth to include measures of depth of vocabulary knowledge (Pearson, Hiebert, & Kamil, 2007). Different aspects of vocabulary knowledge may be somewhat distinct (Oulette, 2006; Wise, Sevcik, Morris, Lovett, & Wolf, 2007) or related to differences in task difficulty among various measures. For instance, children's receptive vocabulary knowledge tends to be more developed than their expressive vocabulary (Lehr, Osborn, & Hiebert, 2004; Kamil & Hiebert, 2005). Third and finally, we employed explanatory item response modeling (EIRM; De Boeck & Wilson, 2004) to determine how much of the variance in performance was attributable to students, words, and tasks.

Study 1

Study 1 had two main goals: (a) to test for the possibility of method effects among various types of morphological awareness tasks and (b) to examine relations between morphological awareness and vocabulary knowledge. Regarding the first goal of testing alternative models of the underlying dimensions of morphological awareness, nine morphological tasks were administered that varied in format. Several tasks had an oral response format whereas others required the selection of one correct answer from four possible distracters (i.e., multiple-choice format). Additionally, the tasks differed in terms of administration (i.e., oral versus written format). Confirmatory factor analysis was used to compare the fit of a unidimensional model (i.e., all nine tasks measured the same underlying factor regardless of task administration) to two models that specified two latent factors that corresponded to the possible formats described above: oral versus multiple-choice response and also oral versus written administration.

Two vocabulary tasks were given that allowed us to examine relations between morphological awareness and vocabulary. Specifically, we compared a model that specified vocabulary to be a distinct but related factor to the morphology factors identified in the goal one analyses to a model that specified that the vocabulary tasks measured the same underlying dimension as one or more of the morphology factors.

Method

Participants—The sample consisted of 99 English-speaking fourth graders ranging from 9 to 12 years of age. They were obtained from elementary schools in a moderately sized Southeastern city within the United States. According to school-wide estimates, 29.8% of students were eligible for free and reduced-price lunch. Participants identified themselves as Caucasian (70%), African American (24%), Asian American (4%), Middle Eastern (1%) and Hispanic (1%). Fifty-five percent were female.

Measures and Procedure—Nine experimenter-created morphological awareness tasks were administered. Five tasks involved the use of derivational morphology (three derivational suffixes choice tasks that involved real words, nonwords, and improbable

suffixes; comes from task; morphological derivation task), three tasks involved the use of compounding (bee grass task; morpheme identification task; morphological construction task), and one task involved the use of inflectional morphology (morphological decomposition task). Additionally, five tasks involved bound morphemes (three derivational suffixes choice tasks that involved real words, nonwords, and improbable suffixes; morphological decomposition and derivation tasks) and four tasks involved free morphemes (bee grass task; comes from task; morpheme identification task; morphological construction task).

Regarding task format, three tasks required an oral response (morphological decomposition, derivation, and construction) and six tasks had a multiple-choice or yes/no format (three derivational suffixes choice tasks that involved real words, nonwords, and improbable suffixes; morpheme identification; comes from task; bee grass task). Additionally, four tasks were administered orally (morphological decomposition, derivation, and construction; morpheme identification) and the remaining five tasks had written administration (three derivational suffixes choice tasks that involved real words, nonwords, and improbable suffixes; comes from task; bee grass task). Each task is described in greater detail below.

Test of morphological structure—decomposition (Carlisle, 2000): This was a 30-item orally administered free response task that involved providing a morphologically derived stem and then a sentence context requiring the participant to provide the morphological base form. For example, “The word is *driver*. The sentence is: Children are too young to ____.” The base and derived forms were equivalent in word frequency. The task contained equal numbers of morphological relations that were transparent (i.e., the sound of the base form is intact in the derived form) and involved a shift (i.e., the phonological representation shifts from base to derived form). Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Test of morphological structure—derivation (Carlisle, 2000): This was a 30-item free response task that was identical to the first task except that it involved providing a base stem and then a sentence context that required the participant to provide the morphological derived form. For instance, “The word is *farm*. The sentence is: My uncle is a ____.” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

The next three morphology tasks all involved derivational suffix items. These tasks were developed at the University of Washington (1999), using items that were based on prior research by Mahony (1994), Singson and others (2000), Nagy, Diakidoy, and Anderson (1993), and Tyler and Nagy (1989, 1990).

Derivational suffix choice—real words: This was a 25-item multiple-choice assessment that required the participant to choose from among four derivationally related suffix options that signaled parts of speech. For example, “Did you hear the _____? (a) announce, (b)

announcing, (c) announced, (d) announcement.” All items were presented in writing for the child to read silently while the experimenter read them aloud. Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Derivational suffix choice—nonwords: This was a 14-item multiple-choice assessment that was comparable to the previous task except that the response alternatives were nonwords. An example item is: “Our teacher taught us how to ____ long words. (a) jittling, (b) jittles, (c) jittle, (d) jittled.” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Derivational suffix choice—improbable suffixes: This was a five-item multiple-choice assessment in which the participant was given an improbable stem and then had to circle one of four sentences that correctly used the stem. For instance, the item is “Dogless. (a) The dogless can run fast. (b) He was in the dogless. (c) When he got a new puppy, he was no longer dogless. (d) He did not try to dogless.” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Bee grass task: This task was based on the research of Elbro and Arnbak (1996) and Fowler and Liberman (1995) and contained 14 multiple-choice items. The written format assessment required children to decide which of two options was a better answer to a riddle. An example of an item is: “Which is a better name for a bee that lives in the grass? A grass bee or a bee grass?” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Comes from task: This was 12-item multiple-choice task based on tasks used by Berko (1958), Carlisle (1995), Derwing (1976), Mahony (1994), and Mahony et al. (2000) that was administered verbally and in writing and required the child to decide if the second word was derived from the first word. For example, “quickly” comes from “quick” but “mother” doesn’t come from “moth.” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Morpheme identification (McBride-Chang et al., 2005): This was a 13-item orally administered free response task that assessed the ability to distinguish different meanings across homophones. For each item, two different pictures were presented simultaneously to the child and each of the pictures was spoken by the experimenter. The child was then given a word or phrase containing the target morpheme and was asked to choose the picture that best corresponded to the meaning of that morpheme. For example, the child was shown

pictures of a steak and a stake, and asked, “Which contains the meaning of the “steak” in “steakhouse”? Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Morphological construction (McBride-Chang et al., 2005): This was a 20-item orally administered free response task that assessed the ability to combine morphemes to construct a new meaning. Scenarios consisting of 2- to 4-sentences were presented and children were asked to construct a word that fit the scenario. Fourteen items involved morpheme compounding (e.g., “Early in the morning, we can see the sun coming up. This is called a sunrise. At night, we might also see the moon coming up. What could we call this?” The correct response was *moonrise*.), while six other items involved syntactic knowledge (e.g., “This is a musical instrument called a hux. Now we have three of them. There are three ____.” The correct response was *huxes*.). Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly.

Two standardized vocabulary tasks were administered.

Vocabulary subtest from the Stanford-Binet, 4th edition (Thorndike, Hagen, & Sattler, 1986): For this expressive vocabulary task, the word was presented to the child in writing while the tester said the word out loud. The child was then asked to define the word orally.

Peabody Picture Vocabulary Test (Dunn & Dunn, 1997): For this receptive vocabulary task, the child was presented with four pictures while the tester said a word aloud. The child was then required to point to the picture that best represented the spoken word.

Procedure—Trained research assistants administered all tasks individually over two or three testing sessions using a fixed order of task administration designed to distribute indicators for the same construct across multiple testing sessions, thereby reducing or eliminating time sampling error from the resultant latent variables. Data collection began in late fall and continued through early spring.

Results

Preliminary analyses—Outliers were identified using the median plus or minus two interquartile ranges (IQR) criterion. Twenty-one outlying data points (less than two percent of the data) were identified and were substituted with a value equal to the median plus or minus two IQR, depending on whether they were high or low. All skewness and kurtosis values were within the acceptable range, and no bivariate outliers or nonlinearity were found.

Descriptive analyses are provided in Table 1. Because many of the tasks were experimental, Cronbach’s Alpha reliabilities were calculated. Reliability values $.70$ were deemed adequate (Nunnally & Bernstein, 1994), and reliabilities for all tasks were $.75$ with four exceptions (see Table 1). Although scores obtained using measures with lower reliabilities

tend to include a substantial amount of measurement error, we left these tasks in the analyses because confirmatory factor analysis extracts the available common variance from low-reliability tasks. We also replicated our analyses after dropping these three tasks and the pattern of results remained unchanged.

Underlying dimensions of morphological awareness—Alternative models of the underlying dimensions of morphological awareness were compared. The first model was a base model that specified a single underlying dimension of morphological awareness. This model consisted of a single factor with all nine morphological awareness measures as indicators. Two additional models were specified. Both models included two distinct but potentially related underlying dimensions based on task features: multiple-choice versus oral responses and also oral versus written administration. This was done by specifying two correlated factors with the morphological tasks that required multiple-choice responses as indicators of the first factor and tasks that were categorized as requiring oral responses as indicators of the second factor. Similarly, we also investigated the possibility that a two-factor model may result from administration format. For this model, we specified two correlated factors with the morphological tasks that were administered orally representing one factor and written tasks representing the second factor.

The unidimensional model provided an outstanding fit to the data (see Table 2). This model was nested in the two-factor model, enabling the use of chi-square difference testing. Chi-square difference test results indicated that the constraints imposed by the single underlying dimension model did not result in a significantly poorer fit than either of the two-factor models representing method variance; the oral versus written administration [χ^2 difference = 2.98(1), $p = .08$, *ns*] and multiple-choice versus oral response format [χ^2 difference = -0.35(1), $p = .55$, *ns*] did not provide a better fit to the data than did the unidimensional model.

Relations between morphological awareness and vocabulary—We next examined relations between morphological awareness and vocabulary by specifying two additional models: (a) A model specifying morphological awareness and vocabulary to be separate but potentially related constructs, and (b) a model specifying that the morphological awareness and vocabulary tasks were measuring the same underlying construct. The motivation for these models were previous reports of strong correlations between measures of morphological awareness and vocabulary in past studies (e.g., Nagy et al, 2003) and also reports that morphological awareness and vocabulary are potentially distinct skills (e.g., Kieffer & Lesaux, 2012a, 2012b; Ramirez, Walton, & Roberts, 2014).

Model fit results are presented in Table 2. Because the unidimensional model was nested in the two-factor model, chi-square difference testing was used to compare the model fits. The results indicated that the additional constraints imposed by the unidimensional model did not significantly reduce model fit [χ^2 difference = 3.76(1), $p = .052$] thereby supporting the unidimensional model (see Figure 1).

Summary

There were two key results of Study 1. First, of the morphological awareness tasks that were administered, all were measures of a single underlying dimension of morphological awareness regardless of response type or administration format. Given the outstanding fit of the unidimensional model to the data, it is not surprising that an alternative model that specified distinctions between morphological awareness based on task features did not provide a significantly better fit to the data. Second, the morphological awareness tasks and vocabulary tasks that were given were measures of the same underlying ability. A model specifying morphological awareness and vocabulary as two distinct but correlated abilities did not provide a significantly better fit to the data than did a model specifying a single, combined morphological awareness and vocabulary factor.

Study 2

Although the results of Study 1 were that the morphological awareness and vocabulary tasks were measures of the same underlying ability, a limitation of Study 1 was that the measures of vocabulary were limited to assessing definitional knowledge. Definitional knowledge refers to an individual's ability to provide a definition when presented with a word (i.e., expressive vocabulary) or pick out the correct vocabulary word when presented with a definition (i.e., receptive vocabulary). However, being able to provide definitions may not capture the full extent of vocabulary knowledge because such tasks do not go beyond assessing the number of words an individual recognizes or can provide a definition for. Both receptive and expressive vocabulary tasks fail to address how much an individual may know about a particular word (i.e., depth).

Morphological awareness may be related to aspects of vocabulary knowledge beyond definitional knowledge (e.g., relational knowledge) (Kieffer & Lesaux, 2012b), and the fact that only definitional knowledge was assessed in Study 1 may have influenced relations between morphological awareness and vocabulary as presented in Study 1. Thus, Study 2 had three main goals: (a) To extend the study of the dimensions of morphological awareness and vocabulary with a richer set of vocabulary and morphological awareness measures that assessed not only breadth of vocabulary knowledge but also depth; (b) to address the potential artifact that differential knowledge of the words used on different morphological knowledge and vocabulary measures affects the correlations between measures of morphology and vocabulary; and (c) to examine sources of variance in measures of morphological awareness and vocabulary knowledge.

In each of the previous studies examining relations between vocabulary and morphological awareness, the factor structure was characterized by either highly correlated factors (Kieffer & Lesaux, 2012a, 2012b; Tannenbaum et al., 2006) or a single unidimensional factor (Study 1). An issue that was addressed in the present study was whether differential knowledge of individual vocabulary words was partly responsible for the high degree of correlation among vocabulary and morphological awareness. For a fairly common word such as *fast*, it may be easy to provide a definition, use the word in a sentence, identify an antonym or synonym, and subsequently, generate a morphologically related version (e.g., *faster*). However, for relatively uncommon words like *legate*, an individual is probably less likely to be able to do

these same things. Because vocabulary performance is dependent on how well an individual knows a given word, this phenomenon could possibly inflate the true correlations among morphological awareness and vocabulary and result in a spurious unidimensional model, and therefore, should be investigated.

To address this issue, the present study used experimenter-created vocabulary and morphological awareness assessments that contained the same 23 words throughout in an effort account for differential word knowledge on the morphological awareness and vocabulary knowledge tasks. Three theoretical models were examined: (a) a four-factor model in which vocabulary knowledge was represented by four distinct, yet related, factors of definitional knowledge, usage, relational knowledge, and morphological awareness; (b) a two-factor model in which vocabulary knowledge was represented by two distinct, yet related factors of vocabulary knowledge and morphological awareness, and (c) a unidimensional model in which all indicators loaded onto a single factor.

To explore sources of variance in performance on vocabulary and morphology tasks, we additionally fit an explanatory item response their model (EIRM) to the data in order to: (a) estimate the variances in responses which were due to individuals, question type, and word type; and (b) compare the log-odds of correctly responding to the item types presented. This was made possible by the fact that the morphological awareness and vocabulary assessment contained the same 23 words throughout, which resulted in words being crossed by questions. Higher variance associated with question or word as compared to participant variance would indicate that variability in performance was largely due to what an individual is required to do (i.e., question type) or to whether words were more or less familiar. Higher variance associated with participant compared to word or question would indicate that variability in performance was largely due to individual differences in vocabulary or morphology.

Method

Participants—The sample consisted of 90 English-speaking eighth graders who were involved in a larger reading comprehension study. They were obtained from a moderately-sized Southeastern school. According to school-wide estimates, 23.9% of students were eligible for free and reduced-price lunch. The sample identified themselves as Caucasian (43%), African American (32%), Hispanic (13%), Asian (3%), and mixed ethnicity (2%). Fifty-four percent were females.

Measures and Procedure

Vocabulary tasks: This assessment contained three parts for each of the 23 words (see Appendix A). The first part asked questions related to general definitional knowledge and assessed participants' breadth of vocabulary knowledge (e.g., "Tell me what the word ___ means; Tell me another definition of the word ___; What is another meaning of the word ___?"). The other two sections – word usage and word relatedness – assessed what is commonly conceptualized as the depth of vocabulary. For the word usage ability sections, participants were prompted with, "Give me a sentence using the word ___; Give me another sentence that uses a different meaning of the word ___." For the word relatedness items,

participants were prompted with, “Give me words that mean the same as ____; Give me words that mean the opposite of ____.” Dichotomous scoring was used, with a correct answer earning a score of 1 and an incorrect answer earning a score of 0. No partial credit was given. Participants’ final scores for each question type (e.g., being able to generate one definition) were calculated as the sum of the total number of items answered correctly across words.

Morphological awareness tasks: This measure had two parts that assessed students’ morphological awareness for the same 23 vocabulary words. For the morphological completion section, students were provided with a sentence and a present tense verb and were required to modify the word to accommodate the sentence. For example, for the word *light*, the sentence is “Her mom _____ the birthday candles” (the correct answer is either *lit* or *lights*). Across items, morpheme manipulation was balanced; 11 items required the use of inflectional morphology (9 regular and 4 irregular¹) and 12 items required to use of derivational morphology.

The morphological generation section required participants to give other words that could be obtained by modifying a target word. For instance, “Give me other words that you can get by changing the word *light*.” Examples of correct responses included *delight*, *lightly*, and *lighter*. Dichotomous scoring was used for the morphological completion section; no partial credit was given. Participants’ final scores were calculated as the sum of the total number of items answered correctly. For the morphological generation section, participants were awarded 1 point per correctly generated morphological variant regardless of whether the generated word was an inflected or derived word. Participants’ final scores were calculated by summing the total number of correctly generated morphological variants. Throughout both tasks, the morphemes presented were free morphemes.

The entire assessment was conducted orally. If participants were unable to provide an answer, testers cued them further (e.g., “Give me any definition you can think of.”). If participants repeated previous answers, testers would ask for an answer that was not identical to previous responses. If no answers were elicited, testers went on to the next item. Responses were recorded exactly as they were given.

Procedure—Trained testers administered all tasks individually over two sessions and utilized a fixed order of task administration. Administration occurred during one week of the spring academic year.

Results

Preliminary analyses—Outliers were identified using the median plus or minus two IQR criterion. Six univariate outliers were found but were not attributable to data entry errors or other malfunction and were not particularly extreme, so they remained in the analyses. There were no multivariate outliers. Partial data were included for one case that had data missing at random. Normality and linearity estimates were within acceptable ranges.

¹Based on the prompt, two items, *run* and *light*, could involve the use of either regular or irregular inflections (i.e., *runs* or *ran* and *lights* or *lit*).

Descriptive statistics are provided in Table 3. Data were also checked for floor and ceiling effects. The item asking for a third definition of the vocabulary words was dropped from further analyses because of a pronounced floor effect, but there was adequate variation among all remaining tasks. Skewness and kurtosis values were all within acceptable ranges.

To analyze within-word variance, a correlation matrix was created for each word. Because each of the tasks included the same 23 words throughout, we were able to create a dataset for each word that had columns representing the eight vocabulary questions and rows representing each participant. In this way, we were able to account for differential knowledge of words because each of the 23 datasets included task performance that was specific to each of the 23 words that were used. Thus, for the creation of any one dataset, vocabulary and morphology items would have looked something like this: “Tell me what the word *catch* means; Tell me another definition of the word *catch*; Give me a sentence using the word *catch*; Give me another sentence that uses a different meaning of the word *catch*; Give me words that mean the same as *catch*; Justin *caught* a cold yesterday; Give me other words that you can get by changing the word *catch*”. In other words, we should theoretically be able to state that variability in participant performance resulted from individual- and task-level variance, not the different vocabulary words, because all vocabulary questions within each dataset included the same word throughout.

We then averaged these 23 correlation matrices to create a single correlation matrix in which each of the 23 words contributed equally thereby accounting for differential word knowledge across items. We used this averaged correlation matrix for analysis. We used mean correlations rather than mean covariances because correlations equalize the contribution of each word to the summary statistic. The composite correlation matrix is presented in Table 4. In general, the entries were lower than the correlations reported in Table 3, reflecting the fact that word-level variance was contributing to the original correlations.

Regarding the 23 individual datasets (one dataset per word), less than 10% of the total correlations among the eight tasks were in the moderate range (i.e., greater than or equal to .30); only four correlations exceeded .50, with no single correlation being greater than .57. Regarding specific relational patterns, being able to provide a second definition most often predicted the likelihood that an individual would be able to generate a second sentence ($r = .31 - .57$, all $ps < .01$), which is not wholly unexpected. In several instances, relational knowledge – being able to generate a synonym or an antonym – was related to performance on the morphological generation task ($r = .30 - .46$, all $ps < .01$) as well as performance on the two definition questions ($r = .30 - .52$, all $ps < .01$). Regarding the possibility of word characteristics influencing specific relational patterns, the words *part*, *stable*, and *consume* tended to demonstrate moderate correlations between the definitional and relational items; the word *catch* had moderate correlations between the definitional and usage items. Only six words had more than three correlations exceeding .30 among all questions.²

²These words were part, stable, store, resign, suspend, and consume.

Underlying Dimensions of Vocabulary Knowledge

Confirmatory factor analyses: Confirmatory factor analyses of the averaged within-word correlation matrix were used to test three alternative models of vocabulary knowledge: (a) a one-factor model was specified by having observed indicators loading onto a single vocabulary/morphological awareness factor; (b) a two-factor model, was specified by having definitional knowledge, usage, and relational knowledge loading on a vocabulary factor, and the morphological questions loading on a morphological awareness factor; and (c) four-factor model was specified with definitional knowledge, usage, relational knowledge, and morphological awareness identified as four distinct yet possibly correlated factors. For these models, definitional knowledge was represented by the two definition questions in the vocabulary assessment, usage was represented by the two sentence generation questions on the measure, relational knowledge was represented by synonym and antonym generation questions, and morphological awareness was represented by the morphological sentence completion and morphological generation questions.

The unidimensional model provided an outstanding fit to the data (see Table 5). The relatively lower loadings for the indicators on the factors are consistent with the magnitudes of the within-word correlations. Chi-square difference testing revealed that none of the less constrained models provided a significantly better fit to the data than the unidimensional model, with nonsignificant chi-square values for the comparison of the one- and two-factor models and [χ^2 difference = 0.21(1), $p = .65$, *ns*] and for the comparison of the one- and four-factor models [χ^2 difference = 7.63(6), $p = .27$, *ns*], respectively. Additionally, we ran a bi-factor model, with all tasks loading on both a general factor of vocabulary knowledge and specific factors representing the four task types of definitional knowledge, usage, relational knowledge, and morphological awareness, which also provided a good fit. However, we favored the unidimensional model given its near perfect fit and because the bi-factor model results might not generalize as well as the unidimensional model due to the fact that the bi-factor model required an additional eight parameters to fit the data. These results further supported a unidimensional model that includes morphological awareness and vocabulary as a single factor (see Figure 2).

Explanatory Item Response Modeling—Our initial hypothesis was that variance could be attributed to individuals, question type, and word type; thus, the initial model fit to the data was most closely aligned to the latent regression linear logistic test model (LR-LTTM). Where our model departs from the traditional LR-LTTM specification is that not only were responses simultaneously nested within individuals and items, but also within words because we had the same 23 words throughout the assessment. EIRM model estimates for individuals, question type, and words can be found in Table 6. For Model 1, the mean estimates for variance explained by individuals, question types, and words were .26 [confidence interval (CI) = .20–.38], .86 (CI = .37–3.62), and .23 (CI = .13–.47), respectively. Most notable is the high degree of variability associated with question type.

Due to the substantial variability found for question type, subsequent EIRM investigated the effects of different tasks (i.e., questions) on performance (see Table 7). The model building process began with specifying an unconditional model (Model 1) to estimate the random and

fixed effects associated with item responses when no predictors were included. Log-odds for average item responses were converted to predicted probabilities, or probability of answering the item correctly, in order to facilitate interpretation of the coefficients. Following Model 1, dummy-coded features of the items (i.e., whether the item required individuals to provide a definition, sentence, synonym or antonym, complete a morphological sentence, or generate morphologically related words) were used as predictors of item responses. Pseudo- R^2 statistics were then computed to estimate how much of the variance in Model 1 was explained by the selected covariates (i.e., the dummy coded features) included in Model 2.

The different types of tasks (i.e., questions) resulted in substantial variability in student performance. The threshold parameter estimate (i.e., difficulty) for generating a single definition of a word was .95 and had a predicted probability of .72. This indicates that with all other covariates included within the model, an individual had a 72% chance of correctly providing one definition. However, asking a student to generate a second definition was considerably more difficult. The estimate for generating a second definition of a word was -1.35 with a probability of .40, which suggests that individuals had only a 40% chance of correctly providing a second definition of a word. Individuals were 32% less likely to successfully complete this second task than the first one. Asking an individual to provide a sentence that correctly used a target word was the easiest task. This question had an estimate of 1.45 and predicted probability of .92, suggesting that individuals had 92% of successfully providing one sentence using the word. The estimate for generating a second sentence was $-.55$, and its associated probability was .60. Generating a synonym and antonym were deemed to be equally challenging. Estimates of $-.90$ and corresponding predicted probabilities of .51 were found for both question parameters. Although each of these tasks assessed some component of vocabulary knowledge, there was a significant difference in task performance between the easiest and most difficult questions: Students were over 50% less likely to be able to generate a second definition of word than to generate one sentence correctly using that word.

For the morphological portion of the assessment, completing a sentence by changing a present tense verb to a correct morphologically appropriate word was more difficult than freely generating morphologically related variants of a target word. Estimates were $-.27$ and .75 and associated probabilities were .66 and .85 for the morphological sentence completion task and morphological generation task, respectively. These results suggest that individuals were nearly 20% more likely to be able to correctly generate morphologically related words than to complete a sentence that required a specific morphological variant.

Discussion

Two studies were carried out to investigate relations between morphological awareness and vocabulary knowledge. The results of the first study were that a unidimensional model explained individual differences in performance on the administered morphological awareness tasks and measures of definitional vocabulary. The results of the second study extended the results of the first study in three primary ways. First, a unidimensional model continued to account for individual differences when an expanded set of vocabulary

measures and a new measure of morphological awareness were obtained. These results are particularly interesting given that the vocabulary tasks measured both breadth and depth of word knowledge. Second, when between-word variance was eliminated by analyzing task variance within-word, a unidimensional model continued to account for individual differences in morphological awareness and vocabulary knowledge. Third, EIRM revealed that individual differences in vocabulary only accounted for a proportion of the variance in task performance on vocabulary assessments; words on the assessment and types of questions asked also impacted performance, with the types of questions asked having the most profound influence on performance.

Our interpretation of these results is that morphological awareness is an integral part of vocabulary knowledge and may even be considered an additional facet of an individual's depth of knowledge. When vocabulary is acquired normally from context, what one knows about a word affects one's ability to define the word, use it in context, identify related words, and identify morphological variants. Furthermore, different morphological tasks appear to vary in their difficulty even though they all potentially tap morphological awareness, which may be explained by Schreuder and Baayen's (1995) model of morphological processing. Being able to generate morphologically related variants of words is easier than determining how to complete a sentence with a morphologically appropriate version of a word because the processing demands are greater (e.g., additional consideration of grammatical information) for the latter task.

However, because our data were correlational in nature and obtained at a single time point, it is not possible for us to test alternative causal models that explain the development of morphological awareness and vocabulary knowledge. For example, if it was the case that individual differences in morphological awareness were the sole determinant of individual differences in vocabulary, one might expect both kinds of measures to load on a single factor. The same result would be expected if performance on morphological awareness tasks was primarily determined by vocabulary knowledge. Furthermore, our results are dependent on the specific morphological awareness and vocabulary tasks that were used in both studies. In Study 1, five of the nine morphological tasks involved real rather than nonwords, and the morphological assessment in Study 2 involved only real words. In English, word knowledge predicts individuals' performance on morphological tasks involving real words (Mitchell & Brady, 2014). This is perhaps not surprising given that morphological awareness can be conceptualized as another facet of depth of word knowledge (Proctor, Silverman, Haring, & Montecillo, 2012). In line with this argument, certain features about the specific morphological awareness tasks used in the present studies may explain why a unidimensional model emerged as the preferred model, in contrast to other investigations of these same constructs (e.g., Kieffer & Lesaux, 2012b).

Kieffer and Lesaux (2012b) gave multiple tasks that represented potential dimensions of vocabulary knowledge including synonyms, semantic associations, multiple meanings, use of context clues, morphological decomposition with real words, and morphological derivation with nonwords. They reported that vocabulary was comprised of three highly-related yet distinct dimensions of breadth, contextual sensitivity, and morphological awareness. From a model comparison point of view, the results could not have been more

different. In both of the present studies, a single factor underlay individual differences in performance on the vocabulary and morphology tasks that were administered. Although Kieffer and Lesaux (2012b) settled on a three factor model, they reported that both four- and five-factor models provided significantly better fits to the data.

What might have accounted for differences in model comparison results between the Keiffer and Lesaux (2012b) and present studies? Although we can only speculate, three key differences between the studies might have accounted for the differences in results. First, 11 of the 13 tasks that Keiffer and Lesaux gave were group-administered written tasks that required reading the items and possible responses, and the two orally presented morphological decomposition tasks required written responses. Consequently, individual differences in reading and writing skill may have influenced task performance. In contrast, the tasks used in the present studies included (Study 1) or used exclusively (Study 2) oral presentation, thereby minimizing or eliminating spurious effects of individual differences in reading on task performance. Second, Keiffer and Lesaux addressed the nesting of students in classrooms and schools by using multiple regression to partial out the effects of classroom and doing all analyses on residual scores. They suggest that if classrooms differ in which aspects of vocabulary they teach, failing to account for classroom effects could result in stronger relations among tasks and ultimately factors; that would not be the case if instructional effects either did not occur or were taken into account in the analyses. This was not done in the present study because classroom identity was not available for Study 1, and Study 2 was carried out in only several classrooms within a single school. Keiffer and Lesaux may be correct about the need to handle classroom effects; however, if classrooms and schools differed in average levels of vocabulary for their students, analyzing residuals would have eliminated genuine variance among students that would have lessened correlations among tasks and potentially generated more factors. What is required to resolve this uncertainty is a new study that includes enough classrooms and schools so that multi-level analysis can be used to model variance at both the student and classroom levels. An example of doing this for literacy and oral language is provided by Mehta, Foorman, Branum-Martin, and Taylor (2005), who modeled the data at both the student and classroom levels. At the student level, they found literacy (i.e., reading and writing) to be unidimensional and correlated with but distinct from oral language. At the classroom level, literacy and language were indistinguishable, except that writing rather than reading or language was affected by teacher effects. Third and finally, the construct of contextual sensitivity, which was one of three factors in Keiffer and Lesaux's retained solution, was not represented in the present study.

Although the results of Keiffer and Lesaux (2012b) and those of the present investigation differ from a model comparison perspective, from a parameter-estimation perspective, the results are more similar. For example, in Study 1, the correlation between the vocabulary and morphological awareness factors was .91. Factor correlations reported by Keiffer and Lesaux ranged from .71 to .85.

Other previous studies have also shown that when definitional vocabulary knowledge is controlled, a unique portion of the variance in reading comprehension abilities can be accounted for by morphological awareness (Carlisle, 2007; Deacon & Kirby, 2004; Kirby et

al., 2012; Kieffer & Lesaux, 2008; Nagy et al., 2006), suggesting that morphological awareness provides a unique contribution to predicting reading comprehension over and above definitional vocabulary knowledge and is therefore a discrete skill. At first glance, our results appear to be inconsistent with previously reported findings that conclude that measures of morphological awareness and vocabulary are separable skills as indicated by their independent contributions to other literacy skills, such as reading comprehension. However, several investigations of the relations between these skills included statistical analyses that analyzed observed rather than latent variables (e.g., regression) (e.g., Kieffer & Lesaux, 2008; Kirby et al., 2012), which differs from the present study. One outcome of analyzing observed rather than latent variables is that the correlation between observed variables is almost never perfect and is therefore less likely to indicate unidimensionality between factors.

The present findings also suggested that the features of the items, specifically the varying difficulty levels associated with the different question types, could explain the variability in student responses. Generating one definition and one sentence of a target word as well as generating morphological variants of a word were deemed to be the easiest types of tasks whereas generating a second definition and coming up with synonyms and antonyms were identified as being the most difficult. Therefore, it is important to be attentive to these item features when selecting and creating vocabulary and morphology assessments, as different task types likely introduce additional sources of variance.

Nonetheless, it is important to acknowledge several limitations of our studies. Firstly, our data were not collected longitudinally, meaning that our results only provide insight to skills at the two developmental levels at which they were assessed rather than provide an opportunity to model how these skills change over time. Secondly, due to the absence of manipulation within our studies, causal inferences cannot be made. Our data only provide correlational evidence for the relationship between morphological awareness and vocabulary knowledge. Thirdly, our results might be different if alternative measures of vocabulary or morphological awareness were used (e.g., *Word Analogy*, Kirby et al., 2012) because there are various methods employed to assess morphological awareness that differ across several dimensions (see Deacon, Parilla, & Kirby, 2008, for a review). Morphology measures can be presented in oral, written, or combined oral and written form; tasks can assess judgment, production, or decomposition abilities. However, similar results have been found across studies that employ measures with different task characteristics, reducing the likelihood that the relation is due to method effects (Carlisle, 2003), as demonstrated in Study 1. Further, our findings are limited to English, as participants included in both studies were native English speakers. It is certainly possible that the present results may not hold for non-English languages, such as Chinese and Japanese, both of which contain many homophones; or Finnish and German languages, which rely heavily on compounding and for which the length of compounded words is nearly unlimited. Fourthly, the sample sizes of both studies may have contributed to the present findings. The unidimensional models in Study 1 and 2 were identified as preferred models based on nonsignificant chi-square differences between these models and other comparable models (e.g., a two-factor model of vocabulary knowledge and morphological awareness; Kieffer & Lesaux, 2012b). It is plausible that with

a larger variety of tasks, multiple dimensions might be supported. Additionally, the sample sizes of the two studies constrained the modeling techniques that could be employed. In the future, a larger sample size and longitudinal assessment could allow for more complex modeling and future investigation of developmental relations between these constructs.

Practical Implications

Our results provide evidence that when vocabulary is learned from context under typical conditions, individuals acquire knowledge about morphology, definitions, usage, and relational knowledge (i.e., antonyms and synonyms), as opposed to merely learning word definitions in isolation. Each of these kinds of knowledge should be considered essential facets of vocabulary knowledge. There are implications of these findings for assessment and intervention.

Beginning with assessment, if tasks that require definition, usage, identifying morphological variants, and relational knowledge all appear to be tapping the same underlying latent ability, why is it not enough to just assess definitional knowledge as is commonly done in vocabulary assessments? Task performance on any single task is determined not only by the underlying skill being measured, but also by task specific strategy and method variance. Including items that assess each of these kinds of knowledge will result in a more robust examination of vocabulary knowledge. Additionally, just because normal acquisition of vocabulary from context by a random sample of students is characterized by learning about morphology, definitions, usage, and relational knowledge, this may not be the case for some groups (i.e., English language learners) or some individuals. A better representation of items that assess each of these facets of knowledge is likely to yield a more robust and complete assessment of vocabulary knowledge.

We would like to emphasize that the present findings do not suggest that assessing only one aspect of vocabulary knowledge, such as definitional knowledge, is all that is necessary to fully gauge an individual's word knowledge. This idea runs contrary to the present findings, as vocabulary knowledge includes an individual's knowledge of word meanings in addition to relational knowledge, the appropriate usage of words, and morphological awareness. The present findings underscore the need for the use of comprehensive vocabulary assessments because vocabulary knowledge is represented multiple underlying facets. Thus, the assessment of multiple component skills, as opposed to only one or two, is crucial.

Turning to intervention, it is important to know whether a particular intervention affects each of the kinds of knowledge that appears to be learned when vocabulary is acquired naturally from context. Also, our results are consistent with the idea that vocabulary instruction and intervention is more likely to be successful if it goes beyond memorizing definitions of new words. Across several decades, Beck, McKeown, and colleagues have argued that there are distinct advantages of incorporating robust vocabulary instruction as a means of improving children's acquisition of new words (Beck & McKeown, 1983, 2004, 2007; Beck, McKeown, & McCaslin, 1983; Beck, McKeown, & Omanson, 1987; McKeown, Beck, Omanson, & Perfetti, 1983; McKeown, Beck, Omanson, & Pople, 1985). Such instructional practices encourage the inclusion of activities that use new vocabulary in multiple contexts that extend beyond the classroom (Beck, McKeown, & Kucan, 2013) and

give children multiple opportunities to acquire new information about words (Rupley, Logan, & Nichols, 1998).

Similarly, classroom instruction should additionally target students' morphological knowledge, as several studies have provided evidence that morphological awareness instruction leads to noticeable gains in vocabulary knowledge (for reviews, see Goodwin & Ahn, 2013 and Bowers, Kirby, & Deacon, 2010). The fact that morphological awareness represents another facet of vocabulary knowledge would indicate that quality morphological instruction and intervention could affect vocabulary development directly and that quality vocabulary instruction should also include a morphological component.

Along the same vein, if good instructional practices are described as those that seek to promote children's knowledge about words beyond their definitions, an assessment with items that measure only definitional knowledge neglects to measure additional aspects of a child's knowledge about words, such as how they are also used and related as well as their morphological variants. Thus, it is important to recognize that vocabulary knowledge consists of more than an individual's ability to define a word, and this multi-faceted nature of vocabulary knowledge should be reflected in our assessments and instructional practices.

Acknowledgments

This research was supported by Grant Number R305F100005 and Grant Number R305F100027 from the Institute for Education Sciences, Grant Number P50 HD52120 from the National Institute of Child Health and Human Development, and a Predoctoral Interdisciplinary Training Grant Number R305B090021 from the Institute for Education Sciences.

References

- Adams, MJ. *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press; 1990.
- Anglin JM. Knowing versus learning words. *Monographs of the Society for Research in Child Development*. 1993; 58(10):176–186.
- Apel K, Diehm E, Apel L. Using multiple measures of morphological awareness to assess its relation to reading. *Topics in Language Disorders*. 2013; 33(1):42–56.
- Beck IL, McKeown MG. Learning words well: A program to enhance vocabulary and comprehension. *The Reading Teacher*. 1983:622–625.
- Beck, IL.; McKeown, MG. Direct and rich vocabulary instruction. In: Baumann, JF.; Kameenui, EJ., editors. *Vocabulary instruction: Research to practice*. New York, NY: Guilford; 2004. p. 13-27.
- Beck IL, McKeown MG. Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*. 2007; 107(3):251–271.
- Beck, IL.; McKeown, MG.; Kucan, L. *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford Press; 2013.
- Beck IL, McKeown MG, McCaslin ES. Vocabulary development: All contexts are not created equal. *The Elementary School Journal*. 1983:177–181.
- Beck, IL.; McKeown, MG.; Omanson, RC. The effects and uses of diverse vocabulary instructional techniques. In: McKeown, MG.; Curtis, ME., editors. *The nature of vocabulary acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1987. p. 147-163.
- Berko J. The child's learning of English morphology. *Word*. 1958; 14:150–177.
- Berninger VW, Abbott RD, Nagy W, Carlisle J. Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguist Research*. 2010; 39(2):141–163.

- Bowers PN, Kirby JR, Deacon SH. The effects of morphological instruction on literacy skills a systematic review of the literature. *Review of Educational Research*. 2010; 80(2):144–179.
- Carlisle JF. Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*. 1988; 9(03):247–266.
- Carlisle, JF. Morphological awareness and early reading achievement. In: Feldman, L., editor. *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum; 1995. p. 189-209.
- Carlisle JF. Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Journal of Reading and Writing*. 2000; 12(3):169–190.
- Carlisle JF. Morphology matters in learning to read: A commentary. *Reading Psychology*. 2003; 24:373–404.
- Carlisle, JF. Fostering morphological processing, vocabulary development, and reading comprehension. In: Wagner, RK.; Muse, AE.; Tannenbaum, KR., editors. *Vocabulary acquisition: Implications for reading comprehension*. NY: Guilford Press; 2007. p. 78-103.
- Carlisle JF. Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*. 2010; 45(4):464–487.
- Carlisle JF, Nomanbhoy DM. Phonological and morphological awareness in first graders. *Applied Psycholinguistics*. 1993; 14(2):177–195.
- Carlisle JF, Stone C. Exploring the role of morphemes in word reading. *Reading Research Quarterly*. 2005; 40(4):428–449.
- Chomsky, N.; Halle, M. *The sound pattern of English*. Cambridge, MA: MIT Press; 1968.
- De Boeck, P.; Wilson, M., editors. *Explanatory item response models: A generalized linear and nonlinear approach*. New York, NY: Springer; 2004.
- Deacon SH, Benere J, Pasquarella A. Reciprocal relationship: Children's morphological awareness and their reading accuracy across grades 2 to 3. *Developmental Psychology*. 2013; 49(6):1113. [PubMed: 22845830]
- Deacon SH, Kieffer MJ, Laroche A. The Relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal models. *Scientific Studies of Reading*. 2014:1–20. ahead-of-print.
- Deacon SH, Kirby J. Morphological awareness: Just more phonological? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*. 2004; 25(1):223–238.
- Deacon, SH.; Parrila, R.; Kirby, JR. A review of evidence on morphological processing in dyslexics and poor readers. In: Reid, G.; Fawcett, A.; Manis, F.; Siegel, L., editors. *The SAGE Handbook of Dyslexia*. London: Sage Publications; 2008. p. 212-237.
- Derwing B. Morpheme recognition and the learning of rules for derivational morphology. *The Canadian Journal of Linguistics*. 1976; 21:38–66.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test. 3*. Circle Pines: MN: American Guidance Service; 1997.
- Elbro C, Arnbak E. The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia*. 1996; 46:209–240. [PubMed: 24234273]
- Fowler, AE.; Liberman, IY. The role of phonology and orthography in morphological awareness. In: Feldman, LB., editor. *Morphological aspects of language processing*. Hillsdale, NJ: Erlbaum; 1995.
- Goodwin AP, Ahn S. A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*. 2013; 17(4):257–285.
- Harcourt Measurement. *Technical Manual: Stanford Achievement Test Series. 10*. San Antonio, TX: Harcourt Publishing; 2003.
- Jenkins JR, Johnson E, Hileman J. When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading*. 2004; 8(2):125–151.
- Kamil, ML.; Hiebert, EH. Teaching and learning vocabulary: Perspectives and persistent issues. In: Hiebert, EH.; Kamil, ML., editors. *Teaching and learning vocabulary: Bringing research to practice*. Mahwah, NJ: Erlbaum; 2005. p. 1-23.

- Kieffer MJ, Box CD. Derivational morphological awareness, academic vocabulary, and reading comprehension in linguistically diverse sixth graders. *Learning and Individual Differences*. 2013
- Kieffer MJ, Lesaux NK. The role of derivational morphological awareness in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing: An Interdisciplinary Journal*. 2008; 21:783–804.
- Kieffer MJ, Lesaux NK. Development of morphological awareness and vocabulary knowledge in Spanish-speaking language minority learners: A parallel process latent growth curve model. *Applied Psycholinguistics*. 2012a; 33(1):23– 54.
- Kieffer MJ, Lesaux NK. Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing*. 2012b; 25(2):347–373.
- Kirby JR, Deacon SH, Bowers PN, Izenberg L, Wade-Woolley L, Parilla R. Children’s morphological awareness and reading ability. *Reading & Writing*. 2012; 25:389– 410.10.1007/s11145-010-9276-5
- Kuo LJ, Anderson RC. Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*. 2006; 41(3):161–180.
- Lehr F, Osborn J, Hiebert EH. Research-based practices in early reading Series: A focus on vocabulary. Pacific Resources for Education and Learning PREL. 2004
- Lesaux NK, Kieffer MJ, Faller SE, Kelley JG. The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*. 2010; 45(2):196–228.
- Kruk RS, Bergman K. The reciprocal relations between morphological processes and reading. *Journal of Experimental Child Psychology*. 2013; 114(1):10–34. [PubMed: 23123144]
- Mahony D. Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing: An Interdisciplinary Journal*. 1994; 6:19–44.
- Mahony D, Singson M, Mann V. Reading ability and sensitivity to morphological relations. *Reading and Writing: An Interdisciplinary Journal*. 2000; 12:191–218.
- Mann, V.; Singson, M. Reading complex words. Springer; US: 2003. Linking morphological knowledge to English decoding ability: Large effects of little suffixes. In; p. 1-25.
- Mcbride-Chang C, Wagner RK, Muse A, Chow BW, Shu HUA. The role of morphological awareness in children’s vocabulary acquisition in English. *Applied Psycholinguistics*. 2005; 26(3):415.
- McKeown MG, Beck IL, Omanson RC, Pople MT. Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*. 1985:522–535.
- McKeown MG, Beck IL, Omanson RC, Perfetti CA. The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Literacy Research*. 1983; 15(1):3–18.
- Mehta PD, Foorman BR, Branum-Martin L, Taylor WP. Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in grades 1 to 4. *Scientific Studies of Reading*. 2005; 9(2):85–116.
- Mitchell AM, Brady SA. Assessing affix knowledge using both pseudoword and real-word measures. *Topics in Language Disorders*. 2014; 34(3):210–227.
- Muse A. The nature of morphological knowledge. (Unpublished doctoral dissertation). Retrieved from PsycINFO. Dissertations Abstracts International: Section 2800 Developmental Psychology. 2005; 66(11):6314.
- Nagy WE, Anderson RC. How many words are there in printed school English? *Reading Research Quarterly*. 1984; 19:304–330.
- Nagy W, Berninger V, Abbott R. Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle school students. *Journal of Educational Psychology*. 2006; 98(1): 134–147.
- Nagy W, Berninger V, Abbott R, Vaughn K, Vermeulen K. Relationship of morphology and other language skills to literacy skills in at-risk second grade readers and at-risk fourth grade writers. *Journal of Educational Psychology*. 2003; 95(4):730–742.
- Nagy W, Diakidoy I, Anderson R. The acquisition of morphology: Learning the contribution of suffixes to the meanings of derivatives. *Journal of Reading Behavior*. 1993; 25:155–170.

- Nagy WE, Scott JA. Word schemas: What do people know about words they don't know? *Cognition & Instruction*. 1990; 7:105–127.
- Nunes T, Bryant P, Bindman M. The effects of learning to spell on children's awareness of morphology. *Reading and Writing*. 2006; 19(7):767–787.
- Ouellette GP. What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*. 2006; 98(3):554.
- Pearson PD, Hiebert EH, Kamil ML. Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*. 2007; 42(2):282–296.
- Proctor CP, Silverman RD, Harring JR, Montecillo C. The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing*. 2012; 25(7):1635–1664.
- Ramirez G, Walton P, Roberts W. Morphological awareness and vocabulary development among kindergarteners with different ability levels. *Journal of Learning Disabilities*. 2014; 47(1):54–64. [PubMed: 24306459]
- Roth, L.; Lai, S.; White, B.; Kirby, JR. Orthographic and morphological processing as predictors of reading achievement. annual meeting of the Society for the Scientific Study of Reading; Vancouver, BC. 2006.
- Rupley WH, Logan JW, Nichols WD. Vocabulary instruction in a balanced reading program. *The Reading Teacher*. 1998:336–346.
- Schreuder, R.; Baayan, RH. Modeling morphological processing. In: Feldman, L., editor. *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum; 1995. p. 131-154.
- Singson M, Mahony D, Mann V. The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*. 2000; 12:219–252.
- Tannenbaum KR, Torgesen JK, Wagner RK. Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*. 2006; 10(4):381–398.
- Thorndike, RL.; Hagen, E.; Sattler, J. *Stanford-Binet Intelligence Scales*. 4. Itasca, IL: Riverside; 1986.
- Tong X, Deacon SH, Kirby JR, Cain K, Parilla R. Morphological awareness: A key to understanding poor reading comprehension in English. *Journal of Educational Psychology*. 2011; 103(3):523–534.
- Treiman R, Cassar M. Effects of morphology on children's spelling of final consonant clusters. *Journal of Experimental Child Psychology*. 1996; 63(1):141–170. [PubMed: 8812040]
- Tyler A, Nagy WE. The acquisition of English derivational morphology. *Journal of Memory and Language*. 1989; 28(1):649–667.
- Tyler A, Nagy W. Use of English derivational morphology during reading. *Cognition*. 1990; 36:17–34. [PubMed: 2383968]
- University of Washington Morphological Awareness Battery. Unpublished experimental test battery. Seattle, WA: 1999.
- Wise JC, Sevcik RA, Morris RD, Lovett MW, Wolf M. The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research*. 2007; 50(4):1093–1109.

Appendix A. List of Words Used in the Vocabulary and Morphological Assessment

1. Run (*Example*)
2. Cover
3. Clear

4. Light
5. Catch
6. Part
7. Clean
8. Drop
9. Store
10. Rough
11. Stable
12. Tangle
13. Treat
14. Stand
15. Wake
16. Note
17. Still
18. Found
19. Store
20. Resign
21. Suspend
22. Honor
23. Support
24. Consume

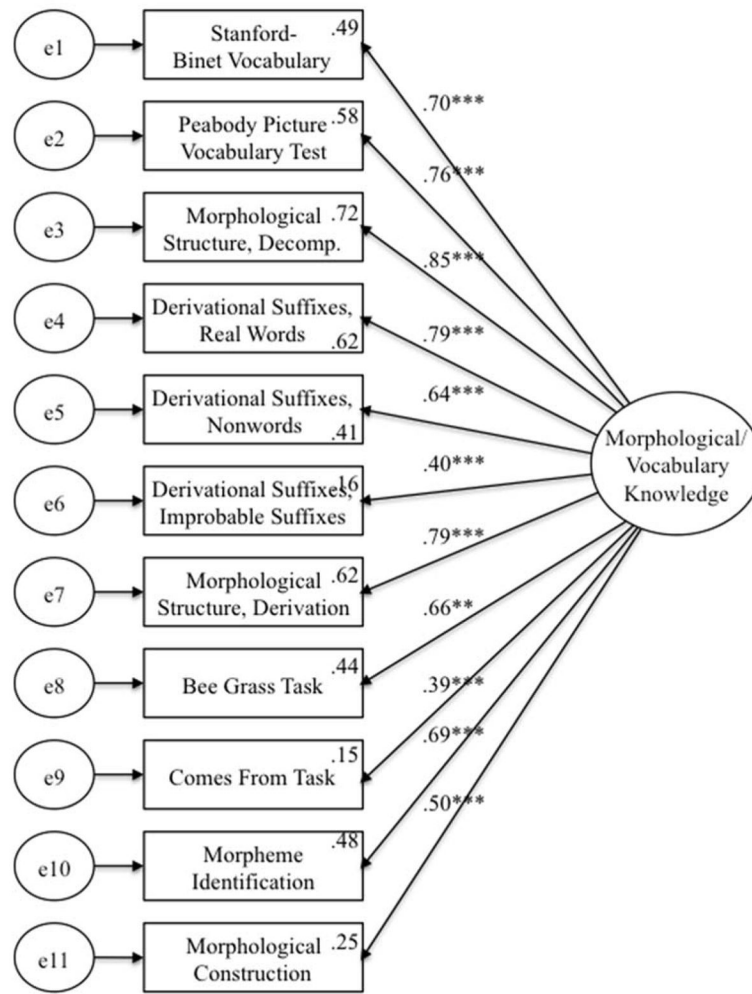


Figure 1. A unidimensional model of morphological and vocabulary knowledge found in Study 1. Decomp. = Decomposition. *** $p < .001$. ** $p < .01$.

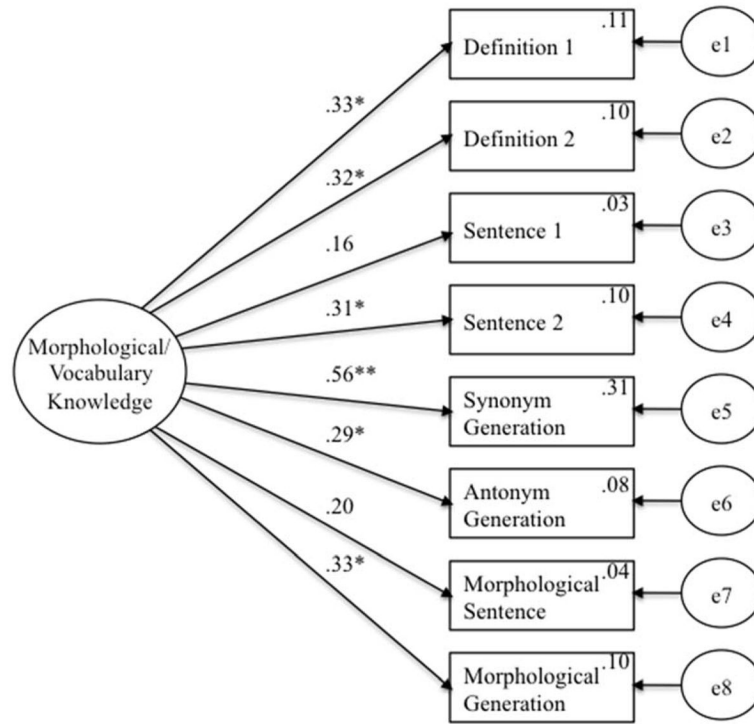


Figure 2. A unidimensional model of morphological and vocabulary knowledge found in Study 2. ** $p < .01$. * $p < .05$.

Study 1 Correlations, Means, Standard Deviations, and Reliability Coefficients for the Vocabulary and Morphology Tasks (N = 99)

Table 1

Task	1	2	3	4	5	6	7	8	9	10	11
1. Morphological Structure Decomp.	--										
2. Derivational Suffix, Real Words	.67**	--									
3. Derivational Suffix, Improbable	.32**	.27**	--								
4. Derivational Suffix, Nonwords	.55**	.69**	.30**	--							
5. Morphological Structure, Derivation	.64**	.55**	.27**	.53*	--						
6. Bee Grass Task	.52**	.52**	.28**	.39**	.50**	--					
7. Comes From Task	.32**	.39*	-.00	.32**	.28**	.31**	--				
8. Morpheme Identification	.61**	.52**	.34**	.42**	.54**	.50**	.22*	--			
9. Morphological Construction	.44**	.40**	.31**	.32**	.30**	.36**	.25*	.31**	--		
10. Peabody Picture Vocabulary Test	.65**	.50**	.30**	.37**	.62**	.63**	.31**	.54**	.44**	--	
11. Stanford-Binet Vocabulary	.64**	.51**	.33**	.32*	.62**	.38**	.20	.46**	.31**	.62	--
Means	25.60	27.62	21.87	3.46	8.90	19.72	10.23	9.66	10.20	12.43	136.66
SD	2.95	2.98	2.92	1.00	3.25	4.48	2.59	2.75	2.47	0.79	50.16
Possible Range	0-46	0-30	0-25	0-5	0-14	0-30	0-14	0-12	0-13	0-20	0-817
α	.88	.75	.77	.33	.76	.80	.68	.84	.54	.44	.95

Note. Decomp. = Decomposition. SD = Standard deviation.

** $p < .01$.

* $p < .05$.

Study 1 Fit Indices of One- and Two-Factor Models of Morphological Awareness and Vocabulary Knowledge

Table 2

Fit Indices	Oral vs. Written		Multiple-Choice vs. Oral		Morphology and Vocabulary	
	Two-Factor	One-Factor	Two-Factor	One-Factor	Two-Factor	One-Factor
χ^2	25.47	28.45	28.80	71.53	75.29	
df	26	27	26	43	44	
<i>p</i>	.49	.39	.53	<.01	<.01	
RMSEA	<.001	.02	<.001	.08	.09	
<i>p</i>	.78	.70	.80	.07	.05	
TLI	1.00	.99	1.00	.92	.88	
CFI	1.00	1.00	1.00	.94	.93	

Note. χ^2 = Chi-square. df = Degrees of freedom. RMSEA = Root Mean Squared Error of Approximation. TLI = Tucker-Lewis Index. CFI = Comparative Fit Index.

Study 2 Correlations, Means, and Standard Deviations for the Vocabulary and Morphology Tasks (N = 90)

Table 3

Question	1	2	3	4	5	6	7	8
1. Definition 1	--							
2. Definition 2	.46**	--						
3. Sentence 1	.42**	.33**	--					
4. Sentence 2	.34**	.60**	.25*	--				
5. Synonym Generation	.64**	.46**	.41**	.45**	--			
6. Antonym Generation	.56**	.41**	.44**	.40**	.71**	--		
7. Morphological Sentence	.50**	.37**	.47**	.47**	.52**	.52**	--	
8. Morphological Word Generation	.49**	.31**	.28**	.30**	.62**	.67**	.39**	--
Means	16.18	9.05	21.43	12.97	15.56	13.81	16.87	56.26
Standard Deviations	3.17	3.37	1.28	3.81	6.94	4.18	3.42	20.35
Possible Range	0 – 23	0 – 23	0 – 23	0 – 23	0 – 23	0 – 23	0 – 23	0 – 117 ^a

Note.

** $p < .01$.

* $p < .05$.

^a Observed maximum of generated morphological variants.

Table 4

Study 2 Correlation Matrix of Within-Word Variance

	1	2	3	4	5	6	7	8
1. Definition 1	--							
2. Definition 2	-.03	--						
3. Sentence 1	.14	.04	--					
4. Sentence 2	.09	.32	-.04	--				
5. Synonym Generation	.23	.19	.08	.12	--			
6. Antonym Generation	.07	.08	.04	.06	.17	--		
7. Morphological Sentence	.11	.03	.10	.07	.08	.05	--	
8. Morphological Word Generation	.11	.06	.05	.08	.19	.14	.11	--

Table 5

Study 2 Fit Indices of the One-, Two-, and Four-Factor Models of Morphological Awareness and Vocabulary Knowledge

Fit Indices	Four-Factor	Two-Factor	One-Factor
χ^2	4.14	11.56	11.77
df	14	19	20
<i>p</i>	.99	.90	.92
RMSEA	<.001	<.001	<.001
<i>p</i>	.98	.97	.98
TLI	1.00	1.00	1.00
CFI	1.00	1.00	1.00

Note. χ^2 = Chi-square. df = Degrees of freedom. RMSEA = Root Mean Squared Error of Approximation. TLI = Tucker-Lewis Index. CFI = Comparative Fit Index. RC = Reading Comprehension.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 6
 Explanatory Item Response Models Estimating Response Variance due to Individuals, Question Type, and Words

Model	Subject	Estimate	SE	95% Confidence Interval		
				Lower Bound	Upper Bound	Upper Bound
Model 1	Individual	0.26	0.04	0.20	0.38	0.38
	Question	0.86	0.46	0.37	3.62	3.62
	Word	0.23	0.07	0.13	0.47	0.47
Model 2	Individual	0.26	0.04	0.20	0.38	0.38
	Question	0.00	---	---	---	---
	Word	0.23	0.07	0.13	0.47	0.47

Note. SE = Standard error.

Explanatory Item Response Models Estimating the Log-Odds of Correct Responses due to Question Type with Different Question Types as Predictors

Table 7

Model	Effect	Threshold	SE	df	t-value	p-value	pp
Model 1	Intercept	0.73	0.35	7	2.09	0.07	0.67
Model 2	Intercept	0.95	0.12	22	7.68	<.001	0.72
	Definition 2	-1.35	0.07	13431	-19.43	<.001	0.40
	Sentence 1	1.45	0.11	13431	12.70	<.001	0.92
	Sentence 2	-0.55	0.07	13431	-8.02	<.001	0.60
	Synonym	-0.90	0.07	13431	-13.24	<.001	0.51
	Antonym	-0.90	0.07	13431	-12.21	<.001	0.51
	Morphological Sentence	-0.27	0.07	13431	-3.78	<.001	0.66
	Morphological Generation	0.75	0.11	13431	6.71	<.001	0.85

Note. Definition 1 serves as the intercept in the present models. SE = Standard error. df = Degrees of freedom. pp = Predicted probability.