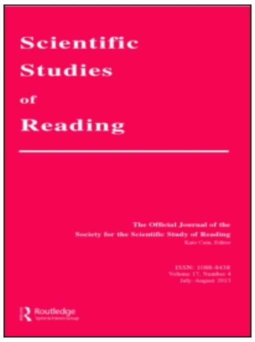

Scientific Studies of Reading

ISSN: 1088-8438 (Print) 1532-799X (Online) Journal homepage: http://www.tandfonline.com/loi/hssr20

Is the Whole Greater Than the Sum of Its Parts? Modeling the Contributions of Language Comprehension Skills to Reading Comprehension in the Upper Elementary Grades

Michael J. Kieffer, Yaacov Petscher, C. Patrick Proctor & Rebecca D. Silverman

SCIENTIFIC STUDIES OF READING
2016, VOL. 20, NO. 6, 436–454
http://dx.doi.org/10.1080/10888438.2016.1214591

Is the Whole Greater Than the Sum of Its Parts? Modeling the Contributions of Language Comprehension Skills to Reading Comprehension in the Upper Elementary Grades

Michael J. Kieffer[a], Yaacov Petscher[b], C. Patrick Proctor[c], and Rebecca D. Silverman[d]

[a]New York University; [b]Florida State University; [c]Boston College; [d]University of Maryland, College Park

ABSTRACT

Language comprehension is crucial to reading. However, theoretical models and recent research raise questions about what constitutes this multifaceted domain. We present two related studies examining the dimensionality of language comprehension and relations to reading comprehension in the upper elementary grades. Studies 1 (Grade 6; N = 148) and 2 (Grade 3–5; N = 311) contrasted factor models of language comprehension using item level indicators of morphological awareness and vocabulary (Studies 1 and 2) and syntactic awareness (Study 2). In both studies, a bifactor model—including general language comprehension and specific factors for each language component—best fit the data, and general language comprehension was the strongest predictor of reading comprehension. In Study 2, the morphology-specific factor also uniquely predicted reading comprehension above and beyond general language comprehension. Results suggest the value of modeling the common proficiency underlying performance on tasks designed to tap theoretically distinct language comprehension skills.

Successful reading comprehension requires not only accurate and fluent word recognition skills but also language comprehension skills (e.g., Gough & Tunmer, 1986; NICHD, 2000; Perfetti, Landi, & Oakhill, 2005; RAND Reading Study Group, 2002; Scarborough, 2001). Despite a general consensus on the important role of language comprehension in understanding written text, particularly in the upper elementary school grades, there is less agreement on what exactly constitutes this broad and potentially multidimensional domain. Theoretical models (e.g., those offered by Gough & Tunmer, 1986; Perfetti et al., 2005; Scarborough, 2001) list several component language skills and processes, implying that their contributions to reading comprehension are at least partially independent; however, these models also tend to emphasize the integration of these component skills during the reading process, suggesting that what is common across these components may be more important than what is unique to each.

Empirically, however, evidence generally suggests that reading comprehension is uniquely predicted by multiple components of language comprehension. The extant research base includes components of (a) broad and deep vocabulary knowledge, specifically knowledge of words' meanings, relations, and uses across contexts (e.g., Anderson & Freebody, 1981; Nation & Snowling, 2004; Ouellette, 2006); (b) morphological awareness (MA), or understanding how complex words are formed from smaller units of meaning (e.g., Carlisle, 2000; Deacon, Kieffer, & Laroche, 2014; Deacon & Kirby, 2004); and (c) syntactic awareness (SA), often operationalized as an understanding of how sentences are structured (e.g., Mokhtari & Thompson, 2006; Nagy, 2007). Recent research has shown that these linguistic skills are predictive of reading comprehension for bilingual as well as

CONTACT Michael J. Kieffer michael.kieffer@nyu.edu New York University, 239 Greene Street, 6th Floor, New York, NY 10012.

monolingual readers (e.g., Kieffer & Box, 2013; Kieffer & Lesaux, 2008; Proctor, Silverman, Harring, & Montecillo, 2012; Silverman et al., 2015; Swanson, Rosston, Gerber, & Solari, 2008).

However, the majority of this evidence (including all of the aforementioned studies) comes from studies that use multiple regression to estimate unique relations for observed measures. Unfortunately, this approach does not account for variance in performance that is due to random measurement error or construct-irrelevant task demands. As such, it is possible that these findings overestimate the unique contributions of individual language comprehension skills to reading comprehension. In addition, researchers often assume that their tasks are tapping separate dimensions based on face validity evidence, rather than testing this assumption empirically (i.e., through factor-analytic approaches). If this assumption does not hold, then inferences about the underlying constructs' relations to outcomes are called into question.

Latent variable modeling, by contrast, has the potential to help researchers distinguish signal from noise in language comprehension, and recently researchers using such modeling have called into question the uniqueness of these components and their contributions (Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015; Neugebauer, Kieffer, & Howard, 2015; Spencer et al., 2015; but see also Nagy, Berninger, & Abbott, 2006; Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003; Kieffer & Lesaux, 2012a, 2012b). Consequently, models of reading comprehension are being reevaluated and extended to test the extent to which components of language comprehension are independent dimensions of individual differences that are uniquely important. These new lines of inquiry stand to illuminate the degree to which these skills are in fact indicators of common underlying linguistic comprehension, or stand-alone, instructionally malleable linguistic abilities.

In this article, we report findings from two studies that investigated the dimensionality of language comprehension and the extent to which theoretically distinct components make unique and shared contributions to reading comprehension in the upper elementary grades. Both studies were conducted with linguistically and ethnically diverse student populations that are increasingly representative of U.S. public schools (Britz & Batalova, 2013). The first study focused on vocabulary and morphological awareness with students in Grade 6, whereas the second study investigated vocabulary, morphological awareness, and syntactic awareness with students in Grades 3–5. In each study, we used confirmatory factor analyses (CFA) with item-level data to contrast several theoretically viable measurement models (one-factor, correlated multifactor, and bifactor models) to represent the dimensionality of language comprehension, and we used structural equation modeling (SEM) to evaluate the unique contributions of the resulting factors to reading comprehension.

Components of language comprehension in theoretical models of reading comprehension

Reading comprehension is defined as the "process of simultaneously extracting and constructing meaning through interaction and engagement with written language" (RAND Reading Study Group, 2002, p. 11). As such, it is a process that draws heavily on a reader's abilities to make meaning when listening and speaking, but also requires different types and uses of language as demanded by the particularly complex nature of written text. Humans evolved to use language for a variety of basic social purposes (e.g., collecting enough food to stay alive; forming social relationships to reproduce), so acquiring listening and speaking skills for such purposes is a nearly universal human accomplishment (Pinker, 1999). By contrast, humans did not evolve to read (e.g., Wolf, 2007), so developing the language comprehension skills required to make meaning from sophisticated written texts is not universal, and indeed these skills demonstrate wide individual differences among children and adults. For instance, nearly all children develop the breadth of vocabulary knowledge required to refer to the objects and relationships in their daily lives, but only some develop the tens or hundreds of thousands of word meanings that appear in typical written texts. Similarly, all children develop enough implicit knowledge of morphology and syntax to understand their parents' instructions and communicate their needs, but many do not develop the metalinguistic and strategic knowledge of these systems required to actively parse written texts. Our focus, then, on language comprehension skills for reading comprehension requires that we go beyond everyday listening comprehension to incorporate both more sophisticated knowledge (e.g., rarer vocabulary that appears primarily in written text) and more metalinguistic abilities (e.g., morphological and syntactic awareness).

In describing this broad domain of language comprehension skills implicated in reading comprehension, theoretical models have drawn conceptual distinctions among the domains of vocabulary, morphology, and syntax. These follow in part from psycholinguistic descriptions that characterizes the mind as including a lexicon with entries for individual words (vocabulary) that are distinct from morphological rules, which are in turn distinct from syntactic rules (Pinker, 1999). However, theoretical models of reading comprehension demonstrate an unresolved tension between these conceptual distinctions and the models' emphases on the integration of students' resources for language comprehension during the reading comprehension process.

For instance, the Simple View of Reading (SVR) defines linguistic comprehension as "the process by which given lexical (i.e., word) information, sentences and discourses are interpreted" (Gough & Tunmer, 1986, p. 7). Although this definition separates out three levels of language, it suggests that these levels of language skills can be operationalized using a single global indicator of listening comprehension, according to seminal SVR research (Gough & Tunmer, 1986; Hoover & Gough, 1990) and subsequent studies (e.g., Adlof, Catts, & Little, 2006; Joshi & Aaron, 2000; Savage, 2006). Although the SVR framework has been powerful in describing variation observed among students, Kirby and Savage's (2008) review points out that it is not a full model and that each term in the SVR equation, including linguistic comprehension, captures a great deal of underlying complexity worthy of study.

Scarborough's (2001) model depicts vocabulary (e.g., breadth, precision, links) and language structures (e.g., syntax, semantics) as two strands within language comprehension that must be integrated together (and with other skills) to produce skilled reading. Here, the components of language comprehension are represented separately, but the emphasis is on their integration in skilled reading comprehension. Moreover, Scarborough indicated that language comprehension skills must become increasingly strategic over time, suggesting that purposeful and effortful use of one's awareness of language plays an increasingly important role in the active comprehension processes involved in skilled reading. Perfetti et al. (2005) listed the lexicon as a component of reading comprehension and further specified the lexicon as involving meaning, morphology, and syntax but were less clear on whether these constituents of the lexicon represent partially independent dimensions or whether they together compose a single dimension of individual differences in students' lexicons.

These models also make some predictions about the development of reading comprehension over time. In the SVR, over time, individual differences in listening comprehension explain increasing portions of variation in reading comprehension, whereas decoding explains decreasing portions, particularly as students transition from the primary grades to the upper elementary grades (e.g., Hoover & Gough, 1990; Language and Reading Research Consortium, 2015b). Scarborough's (2001) model indicates that the development of skilled reading involves increasing integration of componential strands, suggesting that the dimensionality of component language skills may decrease but that individual differences among readers may develop differently. Recent evidence from studies of individual differences indeed suggests that greater dimensionality among language skills emerges over development, particularly after the earliest grades (e.g., Language and Reading Research Consortium, 2015a; Tomblin & Zhang, 2006). Although a full investigation of these developmental questions is beyond the scope of the current study, these theories do support our focus on the upper elementary grades as a period in which language comprehension is likely to be partially independent of word reading skills, potentially multidimensional, and particularly important to reading comprehension.

Unique contributions of language comprehension skills

Across the models just reviewed, a question remains: Should we consider students' skills in theoretically distinct language domains as partially independent sources of variation in reading comprehension performance, or as indicators of a common, underlying, and integrated language proficiency that is required for reading comprehension? For the purposes of the current studies, we consider empirical work that has used multiple regression or path analyses that generally suggest that vocabulary, MA, and SA may each make unique contributions to reading comprehension.

Regarding MA, a growing body of work supports a unique contribution of this skill after controlling for vocabulary, including both cross-sectional studies (e.g., Carlisle, 2000; Kieffer, Biancarosa, & Mancilla-Martinez, 2013; Kieffer & Box, 2013; Kieffer & Lesaux, 2008; Ku & Anderson, 2003; Pasquarella, Chen, Lam, Luo, & Ramirez, 2012) and longitudinal studies (Deacon et al., 2014; Kirby et al., 2012; see also Deacon & Kirby, 2004, and Foorman, Petscher, & Bishop, 2012, for analyses that controlled for autoregressors, but not vocabulary). SA has been investigated less often, but the existing evidence points to a unique contribution after controlling for vocabulary (e.g., Bowey, 1986; Farnia & Geva, 2013; Geva & Farnia, 2012; Simard, Foucambert, & Labelle, 2014; Swanson et al., 2008). Some evidence further suggests that SA makes a unique contribution to reading comprehension after controlling for MA as well as vocabulary (e.g., Proctor et al., 2012; Silverman et al., 2015).

The emerging literature base using latent variable models attempts to overcome the known limitations of using independent total scores from individual assessments. Finding that vocabulary, SA, and MA uniquely relate to reading comprehension could point to a deeper theoretical understanding of individual differences in reading comprehension. At the same time, such findings could be due, in part, to an artifact of measurement. Many studies (including all of those cited in the previous paragraph) have relied on using summary scores (usually total scores) from observed measures of language comprehension skills. Without leveraging what the observed measures may have in common via latent variable models, results from these multiple regression models may result in spurious theoretical conclusions. If, in a multiple regression, the independent variables are highly correlated with each other, potential results from the predictive model may include the suppression of an effect for one or more of the independent variables or an underestimation of the unique effect for an independent variable.

For instance, consider the fifth-grade results from Kieffer and Lesaux's (2008) study of the unique contributions of MA to reading comprehension beyond vocabulary and word-reading skills. Their multiple regression model predicting passage comprehension included scores from three highly correlated word reading measures: Woodcock Letter-Word Identification, Woodcock Word Attack, and TOWRE Sight Word Efficiency. Although the coefficient for Sight Word Efficiency was positive and significant, as we would expect, the standardized coefficient on Woodcock Word Attack was negative (–.22; $p < .10$), suggesting a suppression effect and the (almost certainly) spurious conclusion that phonological decoding skills hinder reading comprehension. For the highly correlated MA and vocabulary measures also included, the coefficient for morphological awareness was positive and significant (.34; $p < .01$), but the coefficient for vocabulary was smaller and not significant (.18; $p > .05$); the latter result is likely an underestimate that suggests the (almost certainly) spurious conclusion that vocabulary, unlike morphological awareness, is not important to reading comprehension. Although Kieffer and Lesaux did not emphasize these spurious inferences in their conclusions, these examples nonetheless highlight the potential danger of neglecting the high correlations among multiple, related language skills when predicting reading comprehension.

In addition, because correlations between variables are attenuated by measurement error, what may theoretically exist as stronger true score correlations among vocabulary, SA, and MA may manifest as moderate observed scores correlations. Consequently, because multiple regression with observed scores assumes that predictors are measured without error, construct-irrelevant demands that are specific to measurement tasks (e.g., productive language demands in open-ended assessments or test-taking skills in multiple-choice assessments) or items (e.g., knowledge of the vocabulary in distractor items; working memory demands of longer item stems) may manifest as signal rather than noise.

Recent studies have used latent variable models to more comprehensively treat observed score data, finding some evidence to support the viability of a broader language construct. For instance, Spencer et al. (2015) used CFA to investigate whether MA and vocabulary constituted distinct dimensions of language proficiency in two data sets, each with multiple measures of vocabulary and MA. Contrary to theoretical predictions, they found that a unitary factor that encompassed both MA and vocabulary knowledge fit both data sets better than a two-factor model that distinguished between these constructs. By contrast, Kieffer and Lesaux (2012b)'s study using a similar CFA approach found evidence for distinct MA and vocabulary factors; however, the authors also found that the two factors were very strongly related (latent correlation = .76), suggesting more overlap in these skills than is typically assumed, as pointed out by Spencer et al. (2015) in their discussion. Regarding syntactic skills, Tomblin and Zhang (2006)'s CFA study found evidence for grammar and vocabulary factors that were distinct but very strongly related (with latent correlations ranging from .78 to .94), so the authors concluded that much of the variance tapped these measures is attributable to a common underlying proficiency. It is important to note that none of these studies examined the contributions of language proficiency to reading comprehension. In a study using vocabulary and syntax (along with decoding fluency and reading comprehension) measures with students in Grades 4–10, Foorman et al. (2015) found that the best fitting measurement model was a bifactor model that included both a general oral language factor (capturing the common variance among syntax and vocabulary items) and specific factors for each of syntax and vocabulary.

In addition, a few studies moved beyond the question of factor structure and used SEM to estimate the unique contributions of individual language comprehension skills to reading comprehension. These studies have produced conflicting results for the contributions of MA and SA to reading comprehension, after controlling for measurement error and/or task-specific demands. Although three SEM studies have found evidence for the unique, direct contributions of MA beyond vocabulary (Kieffer & Lesaux, 2012b; Nagy et al., 2003; Nagy et al, 2006), two others found no evidence for such direct contributions, only indirect contributions that were fully mediated by vocabulary (Goodwin, Huggins, Carlo, August, & Calderon, 2013; Neugebauer et al., 2015). Regarding SA, Foorman et al. (2015) found that their syntax-specific factor did not uniquely contribute to reading comprehension, after accounting for the general oral language factor in each of Grades 5–10, though the SA-specific factor did make a unique contribution to reading comprehension in Grade 4.

In summary, although the evidence from multiple regression studies generally converges in supporting the unique contributions of individual language comprehension skills, the evidence from CFA and SEM studies is more equivocal. The latter studies suggest that latent variable approaches to modeling the unique and shared variance across language comprehension measures may yield different, but ultimately more appropriate, inferences about the contributions of these skills to reading comprehension. In particular, the evidence from Spencer et al. (2015) and Foorman et al. (2015) suggests that the underlying construct of general language comprehension—as indicated by the variance in proficiency shared across diverse measures—is important to model and may be ultimately so predictive of reading comprehension that the unique contributions of individual skills (MA in Spencer et al.; SA in Foorman et al.) become trivial in comparison.

Current studies

Given tensions within theoretical models and conflicting prior findings, we aimed to investigate the dimensionality of language comprehension as it relates to reading comprehension in the current studies. We focused on students in Grades 3–6, a developmental period in which language comprehension plays an increased role, relative to the primary grades. Specifically, we explored multiple approaches to modeling conceptually distinct language skills (vocabulary and MA in Study 1; and vocabulary, MA, and SA in Study 2) as measurably separable component skills and as indicators of general language comprehension (i.e., the overlapping abilities shared among these skills). We further examined the extent to which variance specific to each skill and variance in general language comprehension uniquely predicted reading comprehension.

We report on findings from two secondary analyses that investigated these questions. Both analyses drew on data from larger projects designed to investigate the multifaceted nature of oral language and its relations to reading comprehension. In Study 1, we investigated the dimensionality and unique contributions of MA and vocabulary knowledge, using data on students in Grade 6 drawn from Kieffer and Box (2013) and Kieffer (2014). In Study 2, we expand the focus to incorporate SA as well as multiple measures of MA and vocabulary knowledge using data on students in Grades 3 to 5 drawn from Proctor et al. (2012) and Silverman et al. (2015).

The ordering of the studies reflects the sequence of analyses conducted, not the design of the projects, which were conceptualized and carried out separately. Although the original projects were not designed to sequentially address these specific research questions, the common data analytic approach articulated here provides a robust empirical foundation upon which to extract the independent strengths of both data sets. Although the two studies modeled linguistic comprehension with different observed indicators of linguistic comprehension, we argue that the latent variable approach employed here allows for the broad exploration of whether a given whole of linguistic comprehension contributes beyond the effects of unique components. As such, together both studies provide more robust insights into the roles of language proficiency in reading comprehension than could be produced by either analysis alone. Both studies were conducted in linguistically and ethnically diverse school contexts similar to those in many urban centers, and parallel CFA and SEM analyses were conducted with both data sets.

Method

Study 1 sample

One hundred forty-eight 6th-grade students (66 girls, 82 boys) were recruited from two schools in metropolitan Phoenix, Arizona. A majority (63%) of participants were Latino, 11% were African American, 20% White, 1% Asian, 1% Pacific Islander, and 3% multiethnic. In the two schools, 81% and 65% of the overall student populations were reported to be "economically disadvantaged." Thirteen students (9%) were English language learners according to school records.

Study 1 measures

For all measures in both studies, see the supplemental online materials for extended descriptions and see Table 1 for reliability estimates.

Morphological awareness was assessed using an 18-item nonword suffix choice task developed based on previous research (Nagy et al., 2006; Singson, Mahoney, & Mann, 2000).

Vocabulary knowledge was assessed using the test of academic vocabulary, a researcher-designed measure in which students choose a synonym for a given word on 18 multiple-choice items, drawn from previous research (Lesaux & Kieffer, 2010).

Reading comprehension was assessed using the reading comprehension subtest from the Gates-MacGinitie Reading Test, Fourth Edition (sixth-grade version; MacGinitie, MacGinitie, Maria, & Dreyer, 2000).

Word reading fluency was assessed using the Test of Silent Word Reading Fluency (Mather, Hammill, Allen, & Roberts, 2004). This measure was used to control for the role of word reading skills in reading comprehension, consistent with all of the theoretical models just presented.

Study 2 sample

Three hundred eleven 3rd-, 4th-, and 5th-grade students (164 girls, 147 boys) were recruited from two districts in the mid-Atlantic (n = 193) and Northeast (n = 118) United States. Almost half (47%) of the sample was Latino, 30% was Black, 20% White, and 3% multiethnic. Overall, 71% of the sample qualified for free and reduced meals, and 26% of the students were English language learners according to school records.

Table 1. Means, standard deviations, and correlations among measures for Study 1 (*N* = 148) and Study 2 (*N* = 311).

	Study 1			Study 2								
Construct: Instrument (Scale)	1	2	3	4	5	6	7	8	9	10	11	12
Study 1	1.00											
1. Morphological Awareness: Non-Word Suffix Choice (raw score out of 18)												
2. Vocabulary: Test of Academic Vocabulary (raw score out of 18)	.53	1.00										
3. Word Reading: Silent Word Reading Fluency (standard score)	.47	.42	1.00									
4. Reading Comprehension: Gates-MacGinitie (standard score)	.54	.70	.42	1.00								
Study 2					1.00							
5. Morphological Awareness: Extract the Base (raw score out of 56)												
6. Morphological Awareness: Non-Word Suffix Choice (raw score out of 18)					.63	1.00						
7. Vocabulary: Woodcock Picture Vocabulary (raw score out of 49)					.63	.51	1.00					
8. Vocabulary: CELF Word Class (raw score out of 21)					.65	.45	.65	1.00				
9. Syntactic Awareness: CELF Formulated Sentences (raw score out of 55)					.67	.52	.72	.69	1.00			
10. Reading Comprehension: Woodcock Passage Comprehension (W-score)					.59	.48	.53	.50	.58	1.00		
11. Reading Comprehension: Gates-MacGinitie (raw score out of 46)					.56	.55	.44	.49	.50	.63	1.00	
12. Word Reading: Woodcock Letter-Word Identification (standard score)					.75	.59	.49	.53	.54	.66	.60	1.00
Reliability	.78[a]	.74 [a]	.90-.92 [b]	.92[c]; .83[d]	.98[e]	.78[a]	.88-.92[f]	.77[a]	.76–.82[f]	.81–.91[f]	.89–.93[c]	.96–.98[f]
M	10.88	12.12	98.76	94.37	40.03	9.50	33.23	10.3	38.64	490.98	25.62	100.64
SD	3.99	3.41	12.03	13.53	9.39	4.08	5.27	3.56	10.31	15.65	10.39	14.03

[a]Sample-specific Cronbach's alpha. [b]Publisher-reported Kuder-Richardson Formula reliability. [c]Publisher-reported test–retest reliability. [d]Publisher-reported alternate-forms reliability. [e]Publisher-reported Rasch reliability. [f]Publisher-reported Cronbach's alpha.

Study 2 measures

Morphological awareness was assessed using two measures: the Extract the Base test (August, Kenyon, Malabonga, Louguit, & Caglarcan, 2001), which was designed drawing on the paradigm used by Carlisle (2000), and the same nonword suffix choice task used in Study 1.

Syntactic awareness was assessed using the Formulated Sentences subtest from the Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig, & Secord, 2003).

Vocabulary knowledge was assessed using two measures: the Woodcock–Muñoz Language Survey Picture Vocabulary subtest (Woodcock, Muñoz-Sandoval, Reuf, & Alvarado, 2005) and the Word Classes 2 subtest from the CELF.

Reading comprehension was assessed using two measures: the Gates–MacGinitie, Levels 3, 4, and 5 (MacGinitie et al., 2000) and the Woodcock–Muñoz Passage Comprehension subtest (Woodcock et al., 2005).

Word reading was assessed using the Woodcock–Muñoz Letter-Word Identification subtest (Woodcock et al., 2005).

Data analysis

To address the research questions across both studies, a combination of CFA and SEM were used. Each study tested a series of unidimensional (Figure 1a), correlated factor (Figure 1b), and bifactor models (Figure 1c; Chen, Sousa, & West, 2005), the latter of which we used to explain item correlation due to both a general factor and specific factors. Specifically, a bifactor fits specific factors of MA and vocabulary (Study 1), and MA, SA, and vocabulary (Study 2) to its respective items, as well as a general factor (i.e., language comprehension in both Study 1 and Study 2) across all items.

The bifactor may be viewed as theoretically similar to a second-order factor CFA; however, several distinguishing features of the bifactor model are worth noting. First, the general factor in a bifactor model represents a source of observed score variance above that of the specific factors. In a second-order factor model, the higher order latent construct is indicated by the first-order factors and is conceptually distinct from them. Second, bifactor models allow one to test whether variance across all observed measures is due to a general factor, such as language comprehension, compared to the individual factors of MA-specific knowledge, SA-specific knowledge, and vocabulary-specific knowledge. A second-order factor is ill-equipped to make this comparison, as it captures variance due to differences among the first-order factors. Third, the bifactor model preserves inherent multidimensionality but still provides a single, global construct that underlies observations. Where bifactor models are especially advantageous compared to second-order models is in unpacking the appropriateness of subscales. That is, from a factor-analysis perspective, when the loadings for observed measures are high on a general factor and low on specific factors, it suggests that only a global score is needed and that the specific factors potentially reflect residual noise in the data. Conversely, when loadings are high on a combination of general and specific factors, it suggests that global and subscale scores could be used. Fourth, the measurement specification of bifactor models is such that measurement invariance can be tested for both general and specific factors between groups of individuals, as well as the group comparisons between latent means of the general and specific factors. Second-order factor models are restricted in this context, as only the second-order means can be compared. A final advantage of bifactor models is that the general and specific factors may be used jointly to test predictive strength to other latent variables, whereas second-order variables treat the first-order factors as disturbances for prediction.

The specification of the bifactor model requires that the latent constructs are uncorrelated with each other, which achieves several purposes. First, by postulating that the constructs are independent of each other, the specific factors of vocabulary and MA (and SA in Study 2) represent what is uniquely measured by the respective constructs after controlling for what is shared by the general factor. As such, in a subsequent structural analysis, it is possible to quantify if the constructs of vocabulary, MA, and SA uniquely relate to a proximal reading comprehension outcome after controlling for what is shared between them (i.e., general language comprehension), thereby producing a strong test of the unique

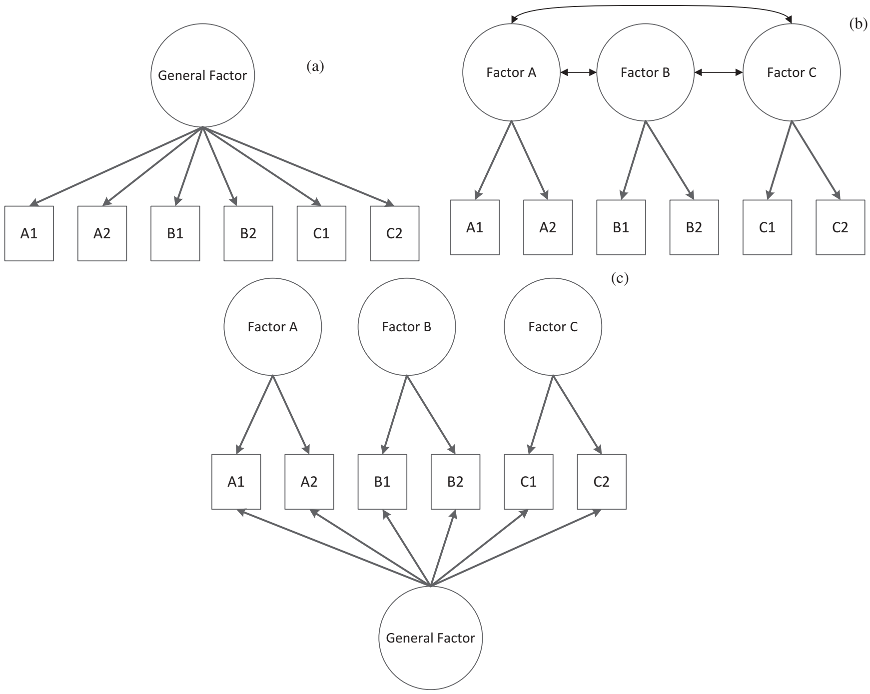

Figure 1. Graphical representation of three generalized measurement models as they can be applied to item-level data from three tasks (A, B, and C) designed to tap theoretically separate skills, including (a) unidimensional model; (b) multidimensional, correlated-factors model; and (c) multidimensional, bifactor model.

contributions of these component skills. Because the latent factors are specified to be orthogonal, results from such structural analyses do not suffer from the problems with multilinearity described in the example from Kieffer and Lesaux (2008) mentioned previously.

Across both studies, the comparative fit index (CFI; Bentler, 1990), Tucker-Lewis index (TLI; Bentler & Bonnett, 1980), root mean square error of approximation (RMSEA; Browne & Cudeck, 1992), and standardized root mean residual (SRMR) were used to evaluate model fit for each model. CFI and TLI values greater than or equal to 0.95 are considered to be minimally sufficient criteria for acceptable model fit, and RMSEA and SRMR estimates less than 0.05 are desirable (Browne & Cudeck, 1992). Moreover, for the nested model comparisons, a chi-difference test was used to evaluate parsimony, along with comparisons of the Akaike information criterion and Bayesian information criterion indices. Following an evaluation of the factor structure, a structural analysis was conducted to examine the predictive relation of the latent composite(s) with performance reading comprehension while accounting for word reading skills (i.e., silent word reading fluency in Study 1 and letter-word identification in Study 2). Due to the relatively small sample sizes, bootstrapping was used in Mplus to improve the model coefficient standard errors for both the measurement and structural analyses (Bollen & Stine, 1993).

Study 1 results

Preliminary analyses

Table 1 provides descriptive statistics, including means, standard deviations, and correlations for MA, vocabulary, silent word reading fluency, and reading comprehension. Participants were in the average range for silent word reading fluency and in the low-average range for reading comprehension, as demonstrated by the standard scores in Table 1. Missing data (from two participants on one measure) were accounted for using full-information maximum likelihood in all subsequent models.

Measurement model

Model comparisons among the one-factor, correlated-factor, and bifactor models suggested that the bifactor specification (Figure 2a) yielded the best fit, $\chi^2(493) = 508.85$, CFI = .98, TLI = .98, RMSEA = .015, 95% confidence interval [.000, .032], and was superior to both the one-factor model of language comprehension ($\Delta\chi^2 = 66.71$, $\Delta df = 34$, $p < .001$) and the correlated-factor model of vocabulary and MA ($\Delta\chi^2 = 48.94$, $\Delta df = 33$, $p < .05$). See supplementary online materials for more specific details of the measurement model results for the individual factor solutions for this study and for Study 2.

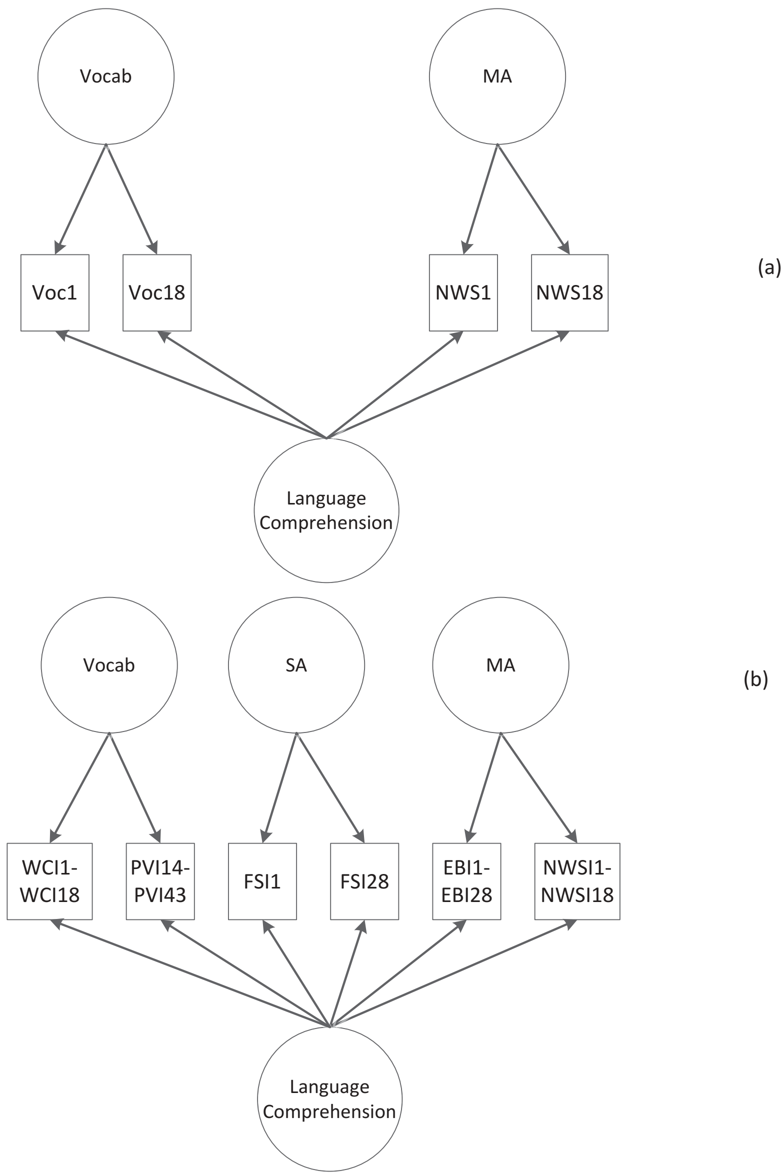

Figure 2. Final bifactor specifications for (a) Study 1 and (b) Study 2. *Note*. Vocab = item on test of academic vocabulary; SA = syntactic awareness; MA = morphological awareness; WCI = item on word classes; PVI = item on Woodcock picture vocabulary; FSI = item on CELF formulating sentences; EBI = item on extract-the-base; NWSI = item on nonword suffix choice.

Structural analysis

Informed by the measurement model comparisons, the three latent factors of vocabulary, MA, and oral language from the bifactor model tested the extent to which they explained individual differences in reading comprehension when controlling for individuals' silent word reading fluency. A hierarchical approach was taken with the structural analysis in order to evaluate whether the specific factors of MA and vocabulary uniquely related to reading comprehension after accounting for the effects of silent word reading fluency and the oral language factor. Results from the full model structural analysis, which included all three latent composites of the bifactor model as predictors of reading comprehension while controlling for silent word reading fluency, are displayed in Figure 3.

The estimates indicated that the strongest predictor of reading comprehension was the common factor of language comprehension (.73, p < .001). Once this factor and silent reading fluency were accounted for in the model, neither the specific factor of vocabulary (.30, p = .093) nor MA (.08, p = .360) uniquely predicted reading comprehension. The inclusion of all predictors in the full model resulted in 65% of the variance in reading comprehension explained. Following the full structural model, the paths of vocabulary and MA to reading comprehension were constrained to 0 to test for differential fit and change in the total R^2. The change in the chi-square between the two models was negligible ($\Delta\chi^2 = 1.55$, $\Delta df = 1$, $p > .05$), and there was no change in R^2, suggesting that the inclusion of the specific factors (vocabulary knowledge and MA) does not provide a benefit for explaining individual differences in reading comprehension above a model that excludes them.

Study 2 results

Preliminary analyses

Table 1 provides descriptive statistics and correlations for the MA, vocabulary, and SA scores, as well as word identification and reading comprehension. Participants were typical word readers

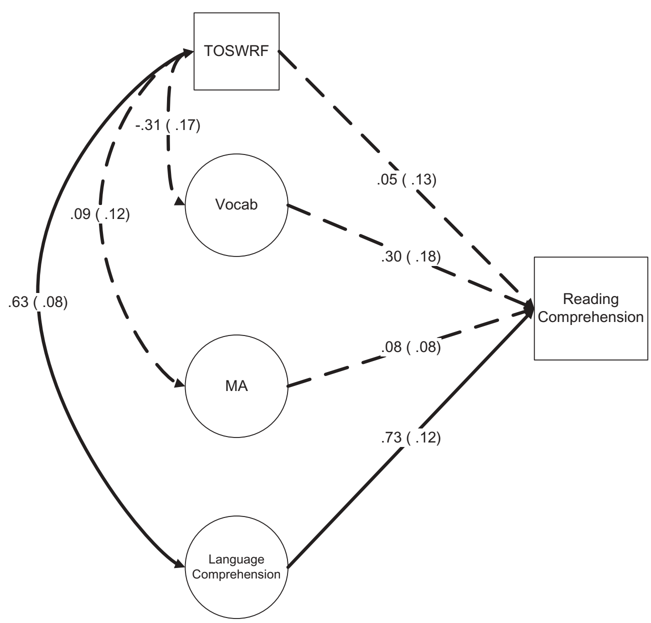

Figure 3. Study 1 structural model of morphological awareness (MA), vocabulary (Vocab), and language comprehension predicting reading comprehension, controlling for silent word reading fluency (Test of Silent Word Reading Fluency [TOSWRF]). *Note.* Significant relations (p < .05) are depicted as solid lines and nonsignificant relations are depicted as dotted lines.

(M = 100.64, SD = 14.03) and low average comprehenders on both the Woodcock–Muñoz (M standard score = 94.22, SD = 11.66) and on the Gates–MacGinite (M percentile rank = 37.29, SD = 26.65), as compared with national norms. Missing data (< 5%) were accounted for using full-information maximum likelihood in all subsequent models.

Measurement model

Similar to Study 1, one-factor, correlated-factor, and bifactor models were tested, with the bifactor model providing the best fit absolute fit to the data (Figure 2b), $\chi^2(6, 325) = 7{,}367$; CFI = .95, TLI = .95, RMSEA = .023, 95% CI [.021, .025], as well as relative fit when compared to the one-factor ($\Delta\chi^2 = 66.71$, $\Delta df = 34$, $p < .001$) and correlated-factor ($\Delta\chi^2 = 66.71$, $\Delta df = 34$, $p < .001$) specifications.

Structural analysis

The latent factors from the bifactor model were used to predict individual differences in reading comprehension when controlling for letter-word identification skills. Model fit results demonstrated excellent fit, $\chi^2(5, 448) = 5{,}746$; CFI = .97, TLI = .97, RMSEA = .013, 95% CI [.008, .019], with standardized coefficients for this model (Figure 4) showing that the general oral language factor had the strongest effect (i.e., .77, $p < .001$) on reading comprehension when accounting for the other variables in the model, followed by letter-word identification (.27, $p < .001$), and the specific factor of MA (.16, $p = .001$). Neither vocabulary (.03, $p > .500$) nor SA (–.03, $p > .500$) provided unique information above that accounted for by the general and MA factors. The linear combination of predictors in the model resulted in 99% of the variance explained in the latent factor of reading comprehension. Similar to Study 1, a constrained version of the structural model was run to test for degradation in model fit and the R^2. Although results showed no difference in fit ($\Delta\chi^2 = 6.12$, $\Delta df = 3$, $p > .05$), the significant path of MA suggests that the unique relation bears further research and replication. As a post hoc analysis, the influence of age via student grade level was included as a covariate in the structural analysis. Fit from this model (see supplemental online materials) was not

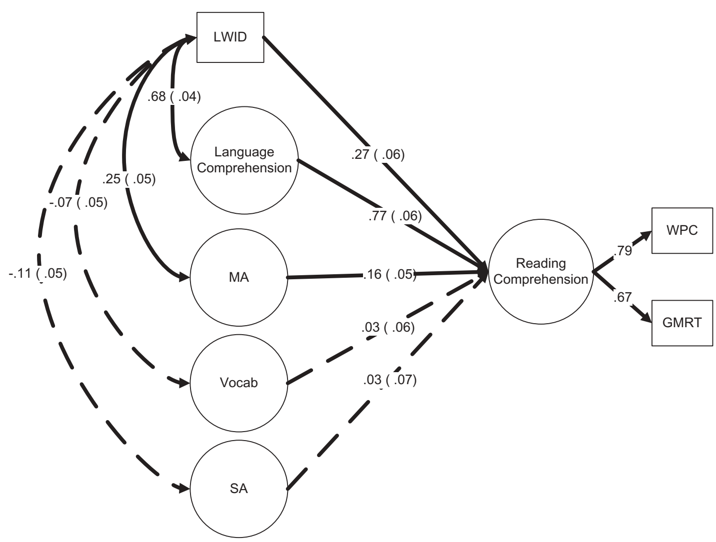

Figure 4. Study 2 structural model of general language comprehension, morphological awareness (MA), vocabulary (Vocab), and syntactic awareness (SA), predicting reading comprehension as measured by Gates–MacGinitie Reading Test (GMRT) and Woodcock Passage Comprehension (WPC), controlling for letter-word identification (LWID). *Note*. Significant relations (p < .05) are depicted as solid lines and nonsignificant relations are depicted as dotted lines.

statistically differentiated from that in Figure 2 ($\Delta\chi^2$ = 218, Δdf = 207, p = .286). The pattern of results with age included as a covariate was consistent with Figure 2 with the exception that the partial, unique effect of decoding changed to .12 and was no longer statistically significant.

Discussion

Across the two studies, we applied an innovative modeling approach to produce converging and complementary insights into the complex roles of language comprehension in reading comprehension in the upper elementary grades. The student samples in both data sets reflect an increasingly typical heterogeneity of students that includes children from home language backgrounds distinctive from the language of schooling. In research such as this, the result of such sample diversity is a broadening of distributions (Suárez-Orozco & Suárez-Orozco, 2001) that benefits inferential analyses (Hoover & Gough, 1990). Including students who were labeled as English language learners and thus limited in English proficiency through to students with above-average English language and reading comprehension skills reflects the linguistic realities of the 21st-century classroom. In both studies, the data were best represented with a bifactor model that included a common language comprehension factor along with specific, uncorrelated factors for the unique skills involved in vocabulary and morphology (Studies 1 and 2), and syntax (Study 2) tasks. Across both data sets, the general language proficiency factor tapped across the language comprehension tasks made very strong contributions to reading comprehension, after controlling for word reading skills. The strength of this contribution was strikingly similar (.73 in Study 1 and .77 in Study 2), whether reading comprehension was measured with scores from one instrument (Study 1) or when modeled as a latent factor using scores from two instruments (Study 2), and despite differences in the grade levels of the two samples.

In addition, in Study 2, the MA-specific factor made a unique contribution to reading comprehension; this contribution was significant and nontrivial in magnitude (.16), after controlling for the general language comprehension factor, word reading skills, the vocabulary-specific factor, and the SA-specific factor. Because this factor captured the portion of variance in MA that was uncorrelated with variance in vocabulary and SA, this constitutes a stronger test of the role of skills unique to MA than prior studies that used multiple regression with similar control measures. In Study 1, this path was positive, though smaller and not significant, and the stronger component effect was from vocabulary (.30 for vocabulary vs. .08 for MA; both $ps > .05$). Neither the vocabulary-specific factor nor the SA-specific factor uniquely predicted reading comprehension in Study 2. These findings have implications for theoretical models of language and reading comprehension as well as future research about how to measure and improve language comprehension.

Toward more robust theoretical and empirical models of language comprehension

The findings from both studies support the importance of modeling and conceptualizing general language comprehension as a common proficiency underlying what is shared across tasks involving vocabulary, morphology, and syntax. As such, this finding supports the focus on the integration of language processes when comprehending written texts (e.g., Gough & Tunmer, 1986; Perfetti et al., 2005; Scarborough, 2001). Consistent with these models, a robust view of language comprehension would attend not only to the ways in which readers use lexical knowledge, and morphological and syntactic awareness in isolation, but also to the ways in which they integrate such knowledge and awareness when making meaning from text. Analyzing language comprehension into its component parts is likely useful for understanding some of the essential ingredients of oral language proficiency, as well as for identifying more specific instructional targets. However, our findings also suggest that limiting our focus exclusively to these component parts may not be the most thorough means by which to conceptualize and measure individual differences that matter for reading comprehension. The superior fit of the bifactor models to both data sets suggested that a general indicator of oral language proficiency, alongside discrete linguistic components, was a more robust way to capture students' individual differences than a set of discrete, correlated skills (as represented by the correlated factor models).

Empirically, this finding converges with recent evidence (e.g., Foorman et al., 2015; Spencer et al., 2015) in supporting the value of general language composites. The consistent predictive strength of the general language comprehension factor across both studies suggests that a single composite score may be more useful for identifying students whose reading comprehension is constrained by their language comprehension than a profile of individual scores on vocabulary, MA, and SA measures. This is not say that the assessment of language comprehension can be limited to only one of these aspects of language; such composite scores will likely be most predictive when they are based on items tapping a range of language comprehension skills, that is, when they have strong construct coverage for the broad domain of language comprehension.

Such general language comprehension composites might also be useful as outcomes for intervention research. Researchers frequently design interventions that target component skills, but do so in the service of improving overall language comprehension. Thus, even if an evaluation of such an intervention indicates significant effects on proximal, researcher-designed measures of component skills and/or on distal measures (e.g., a standardized, norm-referenced listening or reading comprehension test), it is not clear if this goal has been met. The distal measure may have been selected to generalize to a broader domain of skills but also may be insensitive to instruction or have inadequate coverage of the construct as conceptualized by the researchers, by virtue of being externally developed. Using a bifactor model to form a general language comprehension factor with data from researcher-developed measures may have the advantages of providing a better test of the effects of an intervention on language comprehension broadly, as well as the specificity of its effects on component skill. In addition, by investigating effects of a treatment on both general and specific factors, this approach would minimize the inflation of Type 1 error that occurs when conducting multiple tests on related outcomes, because the general and specific factors are specified a priori to be unrelated to one another.

For Study 1, the relatively strong effect of the general language comprehension factor reported here can be viewed alongside the nonsignificant contribution of word reading fluency (and of word reading accuracy, in the post hoc analysis for Study 2). To a certain extent, this finding aligns with the developmental predictions of theoretical models of reading that frame the current research, in which decoding skills become more automatic by later grades (Scarborough, 2001) and thus explain less variation in reading comprehension (Hoover & Gough, 1990). Some research with upper elementary school-age children has also reported nonsignificant effects of decoding similar to those reported here. Lesaux, Crosson, Kieffer, and Pierce (2010), in their work with fourth and fifth graders, found a composite of word reading made a nonsignificant contribution to reading comprehension, and Foorman et al. (2015) reported nonsignificant effects of a unique latent decoding factor on reading comprehension. Be that as it may, this finding may also be related to the means by which MA and vocabulary were operationalized in Study 1. Students had access to the written text of the assessment when completing both language measures. Although each item was read aloud to the students to minimize decoding demands, their scores (and thereby the general language comprehension factor) might nonetheless have tapped decoding to some extent. It is also possible that this result might have been different if a different reading comprehension measure was used, given evidence that reading comprehension measures tap decoding skills to varying degrees (e.g., Cutting & Scarborough, 2006; Keenan, Betjemann, & Olson, 2008).

Mixed evidence for the contributions of morphology, syntax, and vocabulary

Our findings also provide mixed evidence for a unique role of MA in reading comprehension, above and beyond the role of general language comprehension. The significant unique contribution of MA found in Study 2 converges with evidence from prior SEM studies that had controlled for vocabulary knowledge (e.g., Kieffer & Lesaux, 2012b; Nagy et al., 2003; Nagy et al., 2006) while providing a more conservative test of this contribution by controlling for the common variance across language comprehension tasks as well as the variance specific to SA and vocabulary tasks. There are multiple theoretical explanations as to why MA may have a unique role above and beyond general language comprehension. One is that MA may be a particularly important aspect of metalinguistic awareness and that such metalinguistic awareness may be independently useful to reading for comprehension, beyond language skills involved in more implicit language comprehension tasks (e.g., Nagy, 2007). A second is that completing our MA tasks—as in using MA during real reading—involves the integration of orthographic, semantic, and phonological information in an integrative process that mirrors the integrative processes involved in reading comprehension (e.g., Kieffer & Lesaux, 2012a; Kuo & Anderson, 2006; Nagy, 2007), whereas the other language tasks required less integration across different sources of information.

By contrast, the unique contribution of MA-specific abilities to reading comprehension was not significant in Study 1. Although this may be due, in part, to differences in the participants' ages or in the control measures, it is also possible that the use of two MA instruments drawn from different paradigms (i.e., extract the base and nonword suffix choice) in Study 2 produced a more robust measure of MA than the single instrument used in Study 1. Three of the prior SEM studies that found significant contributions for MA beyond vocabulary have used multiple MA measures with differing task demands (Kieffer & Lesaux, 2012b; Nagy et al., 2003; Nagy et al., 2006), whereas one SEM study (Goodwin et al., 2013) that found no significant direct contribution of MA (only an indirect contribution via reading vocabulary) used data from a single MA task, divided into item bundles. Together with the known difficulties in adequately measuring MA (e.g., Carlisle, 2010; Kuo & Anderson, 2006), this evidence suggests that modeling the shared variance across MA items with different formats (and different construct-irrelevant demands) may be essential to capturing the MA construct and its true relationship to reading outcomes.

Regarding SA, Study 2 did not support a unique contribution of SA to reading comprehension, beyond the contribution of general language comprehension. This finding diverges from prior studies using multiple regression that controlled for vocabulary knowledge (e.g., Bowey, 1986; Geva & Farnia, 2012; Proctor et al., 2012; Silverman et al., 2015; Simard et al., 2014; Swanson et al., 2008; but see also Cain, 2007). This divergence may be due, in part, to differences in measures used; the CELF Formulated Sentences test used in Study 2 may tap important metalinguistic skills to a lesser extent than the tasks used in prior research (e.g., Bowey's sentence correction task or Geva & Farnia's grammatical judgment task). However, it may also be due, in part, to the more conservative test for the unique contribution of SA offered by our modeling approach. Our findings for SA largely echo those from Foorman et al. (2015), which used a similar bifactor measurement model. They found that a syntax-specific factor did not uniquely predict latent reading comprehension in six of the seven grades studied (i.e., Grades 5–10), though it was a significant predictor in Grade 4. Together these findings suggest the need for further research on not only how to appropriately measure SA but also how to model it in relation to other language comprehension skills with which it overlaps both conceptually and empirically.

Finally, neither study produced evidence that variance specific to vocabulary uniquely predicted reading comprehension, beyond general language comprehension. This is perhaps surprising given the ubiquity of reporting on vocabulary's uniquely predictive role in reading comprehension studies, including in multiple regression models that also included SA (Geva & Farnia, 2012; Proctor et al., 2012; Silverman et al., 2015; Simard et al., 2014; Swanson et al., 2008) or MA (e.g., Deacon et al., 2014; Kieffer et al., 2013; Kieffer & Box, 2013; Kieffer & Lesaux, 2008; Ku & Anderson, 2003; Pasquarella et al., 2012). Although it is conceivable that this divergence is due to limitations in the measures we selected, this is unlikely, because this finding was confirmed across two studies that used different operationalizations of vocabulary (a synonym task in Study 1; picture naming and semantic association tasks in Study 2). A more likely explanation is that the general language comprehension factor captured the essential core vocabulary skills that are important to reading comprehension, so the remaining variation specific to the vocabulary items was more noise than signal. This hypothesis is supported by Foorman et al. (2015)'s finding that a similar vocabulary-specific factor did not significantly predict reading comprehension in six out of seven grades studied (i.e., with the exception of Grade 7). It also generally converges with recent studies using other modeling approach that have found little independent prediction from vocabulary after accounting for listening comprehension (e.g., Language and Reading Research Consortium, 2015b; Protopapas, Simos, Sideridis, & Mouzaki, 2012; but see also Braze, Tabor, Shankweiler, & Mencl, 2007). In addition, because all of the MA and SA measures required some degree of knowledge of word meanings (like all such meaning-based measures), vocabulary skills were likely at the heart of the general language comprehension factor. Thus, this finding does not disconfirm the importance of vocabulary to reading comprehension but rather suggests that its importance may be better captured when assessed across language tasks than by a synonym or picture-naming task alone.

Limitations

There are some limitations to be noted that can inform future research on these questions. First, as secondary analyses of data sets drawn from larger projects, the two studies were not designed sequentially to build on one another. The two projects focused on similar student populations and had similar broad goals, and findings from each study complement the other in several ways (e.g., confirming results across different measures, Study 2 expands on Study 1 by incorporating SA). Nonetheless, the two studies differed from each other in ways (e.g., differences in grade level, differences in the word reading control measure) that might not have been the case if we had designed them sequentially. Second, although the sample sizes were adequate for whole-group analyses, they could not support subgroup analyses to investigate whether results differed by English language learner status in either study or by grade level in Study 2. Although a post hoc analysis for Study 2 indicated that controlling for grade level did not change our key findings, future research is needed to fully investigate the extent to which the dimensionality of language comprehension, the best approaches to measuring language skills, and predictive relationships might differ among these grades. Third, in Study 2, multiple measures of MA, vocabulary, and reading comprehension were used, but SA was assessed with only one instrument; given the challenges in assessing this construct (Cain, 2007) and the benefits of multiple measures discussed earlier, future studies should incorporate multiple measure of SA. Fourth, given our focus on meaning-based language comprehension skills, rather than language skills broadly, measures of phonological awareness were not incorporated into either study. Finally, as with other correlational findings from cross-sectional studies, our findings cannot shed light on the direction of causality among these skills; given the likelihood of reciprocal relations between language comprehension and reading comprehension, longitudinal and experimental studies using similar latent variable approaches to model these skills are needed.

Conclusion

Across two studies, we found that modeling the common variance underlying performance on vocabulary, morphological awareness, and syntactic awareness tasks may be useful to appropriately capturing language comprehension as it predicts reading comprehension. Findings from one study also suggest that the variance unique to MA tasks (but not necessarily SA or vocabulary tasks) may also be independently important. More broadly, our findings suggest that future research on the roles of language in reading would benefit from considering that the whole may be more valuable than the sum of its parts.

Acknowledgments

Michael J. Kieffer, Yaacov Petscher, and C. Patrick Proctor made equal and complementary contributions to this article and are thus listed in alphabetical order. The authors would like to thank William E. Nagy for his insightful comments on an early presentation of this research and Catherine D. Box for her help with data collection for Study 1.

Funding

Funding was provided by the Institute of Education Sciences, U.S. Department of Education (R305A090152, R305A150199), the National Institute of Child Health and Human Development (P50HD052120), and Spencer Foundation. The opinions expressed are those of the authors and do not represent the views of these funding agencies or the departments of which they are a part.

References

Adlof, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing: An Interdisciplinary Journal*, *19*, 933–958. doi:10.1007/s11145-006-9024-z

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.

August, D., Kenyon, D., Malabonga, V., Louguit, M., & Caglarcan, S. (2001). *Extract the base*. Washington, DC: Center for Applied Linguistics.

Bentler, P. M. (1990). Fit indexes, lagrange multipliers, constraint changes, and incomplete data in structural models. *Multivariate Behavioral Research*, *25*, 163–172. doi:10.1207/s15327906mbr2502_3

Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606. doi:10.1037/0033-2909.88.3.588

Bollen, K. A., & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 111–135). Newbury Park, CA: Sage.

Bowey, J. A. (1986). Syntactic awareness in relation to reading skill and ongoing reading comprehension monitoring. *Journal of Experimental Child Psychology*, *41*, 282–299. doi:10.1016/0022-0965(86)90041-X

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary reading skill differences in young adults. *Journal of Learning Disabilities*, *40*(3), 226–243. doi:10.1177/00222194070400030401

Britz, E., & Batalova, J. (2013, January). Frequently requested statistics on immigrants and immigration in the United States. *Migration Information Source*. Retrieved from http://www.migrationpolicy.org/article/frequently-requested-statistics-immigrants-and-immigration-united-states/#7

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*, 230–258. doi:10.1177/0049124192021002005

Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, *28*(4), 679–694. doi:10.1017/S0142716407070361

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, *12*, 169–190. doi:10.1023/A:1008131926604

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, *45*(4), 464–487.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, *12*, 471–492. doi:10.1207/s15328007sem1203_7

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277–299. doi:10.1207/s1532799xssr1003_5

Deacon, S. H., Kieffer, M. J., & Laroche, A. (2014). The relation between morphological awareness and reading comprehension: Evidence from mediation and longitudinal models. *Scientific Studies of Reading*, *18*, 432–451. doi:10.1080/10888438.2014.926907

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, *25*(02), 223–238. doi:10.1017/S0142716404001110

Farnia, F., & Geva, E. (2013). Growth and predictors of change in English language learners' reading comprehension. *Journal of Research in Reading*, *36*, 389–421.

Foorman, B. R., Koon, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). Examining general and specific factors in the dimensionality of oral language and reading in 4th–10th grades. *Journal of Educational Psychology*, *107*, 884–899. doi:10.1037/edu0000026

Foorman, B. R., Petscher, Y., & Bishop. (2012). The incremental variance of morphological knowledge to reading comprehension in grades 3–10 beyond prior reading comprehension, spelling, and text reading efficiency. *Learning and Individual Differences*, *22*(6), 792–798.

Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing: An Interdisciplinary Journal*, *25*, 1819–1845. doi:10.1007/s11145-011-9333-8

Goodwin, A. P., Huggins, A. C., Carlo, M. S., August, D., & Calderon, M. (2013). Minding morphology: How morphological awareness relates to reading for English language learners. *Reading and Writing: An International Journal*, *26*, 1387–1415. doi:10.1007/s11145-012-9412-5

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education*, *7*, 6–10.

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

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, *21*(2), 85–97. doi:10.1080/02702710050084428

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*(3), 281–300. doi:10.1080/10888430802132279

Kieffer, M. J. (2014). Morphological awareness and reading difficulties in Spanish-speaking language minority learners and their classmates. *Journal of Learning Disabilities*, *47*, 44–53.

Kieffer, M. J., Biancarosa, G., & Mancilla-Martinez, J. (2013). Roles of morphological awareness in the reading comprehension of Spanish-speaking language minority learners: Exploring partial mediation by vocabulary and reading fluency. *Applied Psycholinguistics*, *34*, 697–725. doi:10.1017/S0142716411000920

Kieffer, M. J., & Box, C. D. (2013). Derivational morphological awareness, academic vocabulary, and reading comprehension in Spanish-speaking language minority learners and their classmates. *Learning and Individual Differences*, *24*, 168–175. doi:10.1016/j.lindif.2012.12.017

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphological awareness in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing: An Interdisciplinary Journal*, *21*, 783–804. doi:10.1007/s11145-007-9092-8

Kieffer, M. J., & Lesaux, N. K. (2012a). Direct and indirect roles of morphological awareness in the English reading comprehension of native Spanish, Filipino, Vietnamese, and English speakers. *Language Learning*, *62*, 1170–1204. doi:10.1111/j.1467-9922.2012.00722.x

Kieffer, M. J., & Lesaux, N. K. (2012b). Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing: An Interdisciplinary Journal*, *25*, 347–373. doi:10.1007/s11145-010-9272-9

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389–410. doi:10.1007/s11145-010-9276-5

Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy*, *42*(2), 75–82. doi:10.1111/read.2008.42.issue-2

Ku, Y., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing: An Interdisciplinary Journal*, *16*, 399–422. doi:10.1023/A:1024227231216

Kuo, L.-J., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, *41*, 161–180. doi:10.1207/s15326985ep4103_3

Language and Reading Research Consortium. (2015a). The dimensionality of language ability in young children. *Child Development*, *86*, 1948–1965. doi:10.1111/cdev.12450

Language and Reading Research Consortium. (2015b). Learning to read: Should we keep things simple? *Reading Research Quarterly*, *50*, 151–169. doi:10.1002/rrq.99

Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, *47*, 596–632. doi:10.3102/0002831209355469

Lesaux, N. K., Crosson, A., Kieffer, M. J., & Pierce, M. (2010). Uneven profiles: Language minority learners' word reading, vocabulary, and reading comprehension skills. *Journal of Applied Developmental Psychology*, *31*, 475–483.

MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L. G. (2000). *Gates-MacGinitie Reading Test* (4th ed.). Itasca, IL: Riverside.

Mather, N., Hammill, D. D., Allen, E. A., & Roberts, R. (2004). *Test of silent word reading fluency*. Austin, TX: PRO-ED.

Mokhtari, K., & Thompson, H. B. (2006). How problems of reading fluency and comprehension are related to difficulties in syntactic awareness among fifth graders. *Reading Research and Instruction*, *46*, 73–94. doi:10.1080/19388070609558461

Nagy, W. E. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52–77). New York, NY: Guilford Press.

Nagy, W. E., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, *98*, 134–147. doi:10.1037/0022-0663.98.1.134

Nagy, W. E., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade and at-risk fourth-grade writers. *Journal of Educational Psychology*, *95*, 730–742. doi:10.1037/0022-0663.95.4.730

Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*, 342–356. doi:10.1111/jrir.2004.27.issue-4

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read* (NIH Pub. No. 00-4769). Washington, DC: Authors.

Neugebauer, S. R., Kieffer, M. J., & Howard, E. R. (2015). Exploring multidimensionality and mediation in the roles of lexical knowledge in reading comprehension for Spanish-speaking language minority learners. *Learning & Individual Differences*, *39*, 24–38. doi:10.1016/j.lindif.2015.03.003

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

Pasquarella, A., Chen, X., Lam, K., Luo, Y. C., & Ramirez, G. (2012). Cross-language transfer of morphological awareness in Chinese-English bilinguals. *Journal of Research in Reading*, *34*, 23–42. doi:10.1111/j.1467-9817.2010.01484.x

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell.

Pinker, S. (1999). *Words and rules: The ingredients of language*. London, UK: Orion.

Proctor, P., Silverman, R., Harring, J. R., & 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: An Interdisciplinary Journal*, *25*, 1635–1664. doi:10.1007/s11145-011-9336-5

Protopapas, A., Simos, P. G., Sideridis, G. D., & Mouzaki, A. (2012). The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology*, *33*(3), 217–240.

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Savage, R. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading*, *10*, 143–164. doi:10.1207/s1532799xssr1002_2

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (Dis) abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). New York, NY: Guilford Press.

Semel, E. M., Wiig, E. H., & Secord, W. (2003). *Clinical evaluation of language fundamentals*. San Antonio, TX: The Psychological Corporation.

Silverman, R. D., Proctor, C. P., Harring, J. R., Hartranft, A. M., Doyle, B., & Zelinke, S. B. (2015). Language skills and reading comprehension in English monolingual and Spanish–English bilingual children in grades 2–5. *Reading and Writing*, *28*(9), 1381–1405. doi:10.1007/s11145-015-9575-y

Simard, D., Foucambert, D., & Labelle, M. (2014). Examining the contribution of metasyntactic ability to reading comprehension among native and non-native speakers of French. *International Journal of Bilingualism*, *18*, 586–604. doi:10.1177/1367006912452169

Singson, M., Mahoney, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing: An Interdisciplinary Journal*, *12*, 219–252. doi:10.1023/A:1008196330239

Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C., … Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing: An interdisciplinary Journal*, *28*, 959–988. doi:10.1007/s11145-015-9557-0

Suárez-Orozco, C., & Suárez-Orozco, M. M. (2001). *Children of immigration*. Cambridge, MA: Harvard University Press.

Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology*, *46*, 413–429. doi:10.1016/j.jsp.2007.07.002

Tomblin, J. B., & Zhang, X. (2006). The dimensionality of language ability in school-age children. *Journal of Speech, Language, and Hearing Research*, *49*, 1193–1208. doi:10.1044/1092-4388(2006/086)

Wolf, M. (2007). *Proust and the squid: The story and science of the reading brain*. New York, NY: HarperCollins.

Woodcock, R., Muñoz-Sandoval, A., Reuf, M., & Alvarado, C. (2005). *Woodcock-Muñoz language survey-revised*. Itasca, IL: Riverside.