Journal of Educational Psychology

Academic Language and Listening Comprehension—Two Sides of the Same Coin? An Empirical Examination of Their Dimensionality, Relations to Reading Comprehension, and Assessment Modality

Young-Suk Grace Kim, Yaacov Petscher, Paola Uccelli, and Benjamin Kelcey

Online First Publication, December 19, 2019. http://dx.doi.org/10.1037/edu0000430

Journal of Educational Psychology

ISSN: 0022-0663

http://dx.doi.org/10.1037/edu0000430

Academic Language and Listening Comprehension—Two Sides of the Same Coin? An Empirical Examination of Their Dimensionality, Relations to Reading Comprehension, and Assessment Modality

Young-Suk Grace Kim
University of California, Irvine

Yaacov Petscher
Florida State University

Paola Uccelli
Harvard University

Benjamin Kelcey
University of Cincinnati

Two widely studied language skills in relation to reading comprehension are listening comprehension skill and academic language proficiency. Although their constituent skills and theoretical accounts of how they are related to reading comprehension share a large overlap, they have been studied in separate lines of work. In this study, we investigated the dimensionality of listening comprehension and academic language proficiency tasks, their relations to reading comprehension, and the impact of assessment modality (reading vs. oral language) of academic language proficiency, using data from children in Grade 2 (N = 350). Two cohorts of children from the same schools were assessed on the same set of listening comprehension, word reading, reading comprehension, and academic language tasks. Whereas the first three constructs were assessed in identical manner across the 2 cohorts, academic language tasks were assessed in different modalities (1 cohort in a reading context and the other cohort in an oral language context). Academic language proficiency and listening comprehension skill tasks were best described as having a general oral language construct that captured common variance among all the tasks as well as having specific residual factors. Students' average performance on academic language tasks was lower in the reading context, wherein students' reading skill was also captured beyond the academic "language" proficiency. Across assessment modalities, it was the general oral language construct, not the specific factors, that was reliable, and consistently and most dominantly related to reading comprehension after accounting for word reading.

Educational Impact and Implications Statement

Academic language proficiency and listening comprehension skill largely capture an overlapping common ability, not disparate abilities, and it is this common ability that is important to reading comprehension. Assessment of academic language proficiency should consider modality (oral or reading context) considering assessment purpose and children's developmental phase in reading. The findings demonstrate the importance of integrating evidence across lines and subfields of research (e.g., academic language, listening comprehension) to advance our theoretical understanding of language and literacy development.

Keywords: academic language, assessment modality, dimensionality, listening comprehension, reading comprehension

Supplemental materials: http://dx.doi.org/10.1037/edu0000430.supp

Editor's Note. Scott Baker served as the action editor for this article.—SG

Young-Suk Grace Kim, School of Education, University of California, Irvine; Yaacov Petscher, College of Social Work, Florida State University; Paola Uccelli, Graduate School of Education, Harvard University; Benjamin Kelcey, College of Education, Criminal Justice, and Human Services, University of Cincinnati.

We thank participating schools, teachers, and children. This research was supported by a grant from the Institute of Education Sciences (R305A130131). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agency.

Correspondence concerning this article should be addressed to Young-Suk Grace Kim, School of Education, University of California, Irvine, 3455 Education Building, Irvine, CA 92697. E-mail: youngsk7@uci.edu

Oral language is a multifaceted construct and has been described and classified in multiple ways by grain sizes (sublexical [phonemes, morphemes], lexical [vocabulary], discourse [listening comprehension]) or aspects (e.g., phonology, semantics, pragmatics). By now, a large body of literature indicates the roles of oral language skills in reading acquisition. Among these skills, one prominent oral language construct is listening comprehension. *Listening comprehension* is oral language comprehension at the discourse level—including multi-utterance conversations, narrative stories, and informational oral texts (Kim, 2016; Kim & Pilcher, 2016). Listening comprehension skill is necessary for reading comprehension according to the simple view of reading (Hoover & Gough, 1990; see Florit & Cain, 2011, for a meta-analysis). In another line of work, a construct called *academic language proficiency* has been studied. *Academic language*, although definitions vary, refers to a type of language (or the constellation of language features) most prevalent in school or academic learning (e.g., Bailey & Butler, 2002; Schleppegrell, 2001; Snow, 1983; Snow & Uccelli, 2009). Studies have shown the relation of students' academic language skills to their reading comprehension (Carlisle, Beeman, Davis, & Spharim, 1999; Phillips Galloway & Uccelli, 2019; Snow, Cancini, Gonzalezm, & Shriberg, 1989; Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015).

In the present study, we build on and extend the previous studies on listening comprehension skill and academic language proficiency in four ways. First, we examined listening comprehension skill and academic language proficiency simultaneously to investigate their dimensionality—whether they are best described as dissociable constructs or skills (multidimensional) or as a single construct (unidimensional). Prior conceptualization and empirical work on listening comprehension skill and academic language proficiency, despite apparent shared commonalities (see the literature review below), have been conducted in separate lines of work, and therefore, these two constructs, let alone their dimensionality, have not been considered simultaneously. This is an important effort to integrate theory and evidence across lines of work to bring coherence in the field. Second, we investigated the relations of identified dimension(s) of oral language skills (across listening comprehension and academic language proficiency tasks) to reading comprehension. Third, we extend prior work on academic language proficiency with preadolescents and adolescents by examining academic language proficiency with children in primary grades (i.e., second grade). Finally, we examined a potential impact of assessment modality—whether academic language proficiency is assessed in oral versus reading contexts—on performance level and its relation to reading comprehension for second graders.

Listening Comprehension Skill and Reading Comprehension

The role of listening comprehension skill in reading comprehension has been hypothesized in the simple view of reading, which states that reading comprehension relies on word reading and listening comprehension skills (Gough & Tunmer, 1986). *Listening comprehension skill* is "the ability to comprehend oral language at the discourse level—including multi-utterance conversations, stories, informational oral texts—that involves the processes of extracting, constructing, [and integrating] meaning" (Kim & Pilcher, 2016, p. 160; [integrating] has been added to the original). Listening comprehension skill is typically operationalized or assessed by tasks where children listen to oral connected texts such as passages, stories, and extended conversations (Adlof, Catts, & Little, 2006; Hoover & Gough, 1990; Kim, 2015, 2017; Tunmer & Chapman, 2012). A large body of literature has supported the role of listening comprehension skill in reading comprehension across languages with varying writing systems (e.g., see Florit & Cain, 2011, for a meta-analysis). In recent studies using latent variables, listening comprehension and word reading skills explained the vast majority of variance in reading comprehension (over 90%; Adlof et al., 2006; Kim, 2015, 2017; Lonigan, Burgess, & Schatschneider, 2018). Furthermore, for students in middle school, listening comprehension skill explained almost all the variance in reading comprehension (Catts, Adlof, & Ellis Weismer, 2006).

Despite robust evidence on the role of listening comprehension skill in reading comprehension, listening comprehension was not clearly defined in the simple view of reading.[1] It was only recently that the role of listening comprehension skill has been clearly defined and reasons why listening comprehension skill is robustly related to reading comprehension have been unpacked. This line of work has shown that listening comprehension skill is a discourse-level skill like reading comprehension. As such, listening comprehension involves the same complex discourse processes as reading comprehension, including constructing mental representations (i.e., surface code, textbase representation, situation model)—the only additional process involved in reading comprehension is the decoding process (Kim, 2016, 2017, 2019). That is, the construction and integration processes in discourse comprehension (Kintsch, 1988; McNamara & Magliano, 2009) are involved in listening comprehension, and, therefore, listening comprehension skill draws on highly similar language and cognitive component skills as does reading comprehension, such as vocabulary, syntactic knowledge, higher order cognitive skills and regulation (reasoning, inference, perspective taking, and comprehension monitoring), and background knowledge (content knowledge and discourse knowledge [e.g., text structure, genre, register]) as well as domain-general memory (working memory) and attentional resources (e.g., Daneman & Merikle, 1996; Florit, Roch, & Levorato, 2011, 2014; Kendeou, Bohn-Gettler, White, & van den Broek, 2008; Kim, 2016, 2017; Kim & Phillips, 2014; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012; see Kim, 2016, 2019, for a theoretical model). Consequently, these language and cognitive component skills are related to reading comprehension indirectly via listening comprehension (Kim, 2015, 2017, 2019).

[1] Gough and Tunmer (1986) defined *linguistic* or *listening comprehension* as "the process by which, given lexical (i.e., word) information, sentences and discourses are interpreted" (p. 7). However, this definition was ambiguous and did not provide a clear picture about mechanisms, its operationalization, or component skills. Consequently, linguistic comprehension has been operationalized in various ways (e.g., vocabulary knowledge, or story comprehension; see Florit & Cain, 2011). Recent efforts clarified listening comprehension as a discourse-level oral language skill and also clarified its component skills (see Kim, 2016, 2017).

Academic Language Proficiency and Reading Comprehension

Another line of work that has garnered much attention in practice and research, specifically for school-age children, is academic language (August & Shanahan, 2006; Gersten et al., 2007; Snow & Uccelli, 2009). In educational practice, for example, the widely adopted Common Core State Standards and similar state standards in the United States require explicit instruction of "academic . . . words and phrases" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). In research, academic language has been investigated in several lines of work and has been referred to with various terms, such as academic English, literary language, or language of school (Bailey & Butler, 2002; Nippold, 1999; Scarcella, 2003; Schleppegrell, 2004; Snow, 1983; Uccelli, Barr, et al., 2015, Uccelli, Galloway, et al., 2015). Although precise definitions and nuances differ among the various terms, in general, they refer to a type of linguistic register that is prevalent in academic contexts or in schoolwork (Bailey & Butler, 2002; Gersten et al., 2007; Scarcella, 2003; Schleppegrell, 2001, 2004; Snow, 1983; Snow, Tabors, & Dickinson, 2001; Snow & Uccelli, 2009; see Table 1), and include lexical (vocabulary), syntax, register, and text features (e.g., arrangement of ideas beyond the sentence level).

Academic language is typically contrasted with informal or conversational language, which relies on shared physical context, shared background, and paralinguistic cues for cohesion (e.g., prosodic cues or gestures). The language of school texts tends to be decontextualized where meaning is established by lexical and syntactic features rather than contextual cues (Schleppegrell, 2004; Snow, 1990). The following paragraph on water cycle in a science textbook for fourth graders illustrates this point.

> When water evaporates into water vapor, it is pure water. All the sediments and minerals that were in the water, while the water was on the ground, stay on the ground. This means that the mud in puddle, the salt in the oceans (and the Great Salt Lake), and any pollutants in water stay on the surface of Earth. (Utah State Board of Education, 2017, p. 15)

To understand this text successfully, the child needs to understand the concept in the text primarily via language itself such as technical terms (e.g., *evaporation*, *vapor*, *sediments*, *minerals*) and syntactic features (e.g., complex noun phrases and complex sentence structure; e.g., "All the sediments and . . ., while the water was on the ground, stay on the ground") rather than shared social and physical contexts (thus, decontextualized). It is important to note here that although academic language features are typically

Table 1
A Comparison Between Listening Comprehension and Academic Language: A Selected Review

Area	Listening comprehension skill	Academic language[a] proficiency
Definition	"The ability to comprehend oral language at the discourse level—including [multi-utterance] conversations, stories, informational oral texts—that involves the processes of extracting, constructing and [integrating] meaning" (Kim & Pilcher, 2016, p. 160; [integrating] has been added to the original)	"Knowledge and deployment of a repertoire of language forms and functions that co-occur with oral and written school learning tasks across disciplines" (Uccelli, Barr, et al., 2015, p. 1079)
Key components	Kim (2016, 2019) • Background knowledge (content knowledge and discourse knowledge [e.g., text structure knowledge; genre knowledge, register knowledge]) • Higher order cognitions and regulation (e.g., reasoning, inference, perspective taking, monitoring) • Vocabulary and grammatical knowledge (morphosyntactic and syntactic knowledge), which interacts with morphology and phonology • Domain general cognitions such as working memory, inhibitory and attentional control, shifting	According to CALS, Uccelli, Barr, et al. (2015) • Recognizing academic register (register knowledge) • Organizing analytic texts (text structure knowledge) • Interpreting writers' viewpoints (perspective taking) • Tracking participants and themes (referential inference) • Unpacking dense information (syntactic knowledge) • Connecting ideas logically (knowledge of connectives) • Understanding metalinguistic vocabulary (vocabulary knowledge)
Measurement	Listening comprehension skill: • Comprehension of narrative and expository oral texts (Kim, 2016, 2017, 2019) • Story or narrative comprehension (Florit et al., 2011, 2014; Lepola, Lynch, Laakkonen, Silvén, & Niemi, 2012; Tompkins, Guo, & Justice, 2013)	Academic language proficiency: • CALS-I that includes the component skills above (Phillips Galloway & Uccelli, 2019; Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015) • Knowledge of connectives (Crosson & Lesaux, 2013) • Vocabulary knowledge (Kieffer & Lesaux, 2012) • Register awareness (Carlisle, Beeman, Davis, & Spharim, 1999; Snow et al., 1989)

Note. CALS = Core Academic Language Skills.
[a] Academic language itself (not academic language proficiency) is defined as follows: "language of schooling" that includes lexical and grammatical features and stance expected in school that is conceptualized broadly as formal learning (Schleppegrell, 2001), and "vocabulary, syntax, discourse, and language functions as they cut across different contexts of use and cognitive and textual demands" (Bailey & Butler, 2002, pp. 7–8).

found in written texts, particularly expository texts in content areas (e.g., social studies, science), they are not limited to them. Instead, academic language features are found in both oral language and written language contexts for academic purposes across genres (e.g., literary work in narratives; see Figure 1).

The construct of academic language has been studied and operationalized broadly as vocabulary, syntax, and stance (e.g., detached and authoritative stance; Bailey & Butler, 2002; Scarcella, 2003; Schleppegrell, 2004); register (academic and nonacademic register; e.g., Solomon & Rhodes, 1995); and demands in contextual support (i.e., contextualized and decontextualized language) and metacognition (i.e., register awareness; e.g., Snow, 1983, 1990). Recently, in an effort to integrate these previous efforts, Uccelli, Barr, et al. (2015; Uccelli, Galloway, et al., 2015) conceptualized academic language proficiency as Core Academic Language Skills (CALS), which includes the text (e.g., perspective taking, text structure knowledge, referential inference), lexical (vocabulary), syntax or morphosyntactic, and register (register awareness) components.

It was argued that the type of language prevalent in school texts (i.e., academic language) can be challenging as most children enter school dependent on conversational exchanges (Bereiter & Scardamalia, 1987; McCutchen, 2006; Snow, 1983, 1991), and therefore, accessing and encoding complex ideas in school written texts (i.e., reading comprehension) require academic language proficiency (Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015). Indeed, studies have shown the relation of academic language proficiency to reading comprehension (e.g., Carlisle et al., 1999; Phillips Galloway & Uccelli, 2019; McKeown, Crosson, Moore, & Beck, 2018; Snow, 1990; Snow, Cancino, De Temple, & Schley, 1991; Uccelli, Galloway, et al., 2015). For example, Carlisle et al. (1999) and Snow and her colleagues (Snow, 1990; Snow et al., 1991) found that academic language skill operationalized as a metacognition (academic register awareness) predicted reading comprehension over and above vocabulary for children in elementary school grades. Other studies focused on linguistic features such as children's knowledge of connectives, academic vocabulary, and complex syntax, and their relations to reading comprehension (e.g., Crosson & Lesaux, 2013; McKeown et al., 2018). Uccelli and her colleagues found that CALS was related to reading comprehension for children in upper elementary grades (Uccelli, Galloway, et al., 2015), and that growth in CALS from Grade 6 to Grade 7 was related to growth in reading comprehension (Phillips Galloway & Uccelli, 2019).

Gaps in the Literature and the Present Study

Although previous work on listening comprehension skill and academic language proficiency has shown the roles of language skills in reading development, there are several important gaps in the literature. The first gap is the relation between listening comprehension skill and academic language proficiency as they have been conceptualized and studied in separate lines of work. This is a critical gap and may create confusion in the field. As suggested in the review above and detailed below, there is a clear and large conceptual overlap between these constructs (particularly in terms of component skills; see Table 1). However, the apparent overlap has not been acknowledged or explored. Figure 1 is a heuristic representation of the academic language layer (continuum of academic and nonacademic language, and their exemplars) and the discourse language layer in which language is described in oral (listening comprehension and oral production) and written modalities (reading comprehension and written composition) as well as receptive (listening comprehension and reading comprehension)

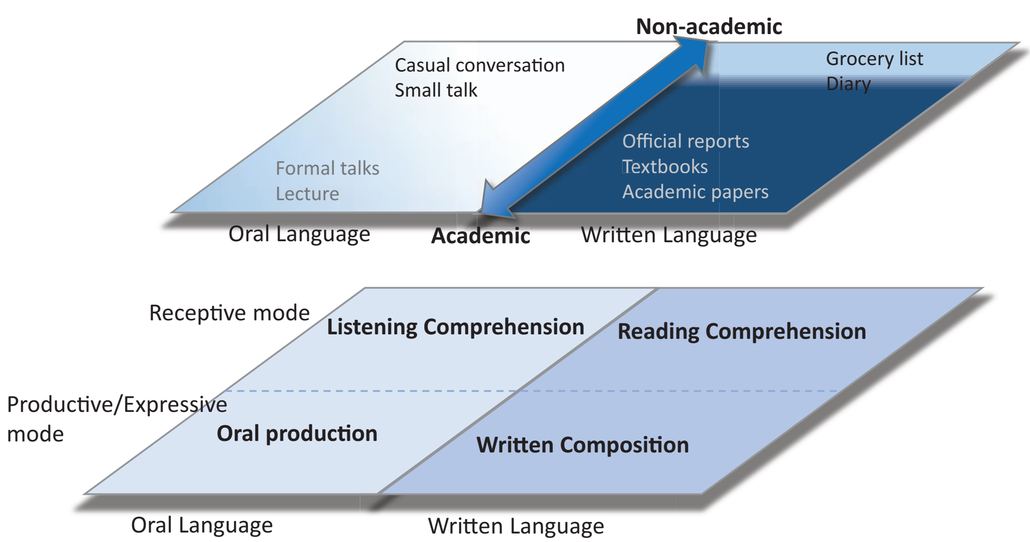

Figure 1. A heuristic representation of the academic language layer (continuum of academic and nonacademic language) with illustrative examples and the discourse language layer in different modalities (oral and written language and receptive and productive mode). Note that the academic language layer here includes written texts. However, this is for description of academic language features, and academic language *proficiency* is operationalized as "oral" language proficiency (see Phillips Galloway & Uccelli, 2019; Uccelli, Barr, et al., 2015) to access academic *written* texts. See the online article for the color version of this figure.

and productive modalities (oral production and written composition). Clear in Figure 1 are similarities and differences between listening comprehension and academic language. Listening comprehension skill is a general ability (general in the sense that it is not specific to academic contexts) to comprehend discourse-level oral language *across* the continuum of academic (e.g., lecture or formal talks) and nonacademic language contexts (e.g., casual conversation). Academic language proficiency is one's ability to use language for academic purposes specifically, but it is not confined to the receptive or productive oral modalities (comprehension or production). Note that Figure 1 includes academic language in both oral language and written language modalities. Prior work on description of academic language features focused on written texts (e.g., Bailey & Butler, 2002; Schleppegrell, 2001). Then, the argument is that children's academic language proficiency or academic language skills in oral language are necessary to access these academic language features in *written* texts (i.e., reading comprehension; Phillips Galloway & Uccelli, 2019; Uccelli, Galloway, et al., 2015).

From a theoretical point of view, there is a clear, large overlap between listening comprehension skill (Kim, 2015, 2016; Kim & Pilcher, 2016) and academic language proficiency, specifically CALS (Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015). Table 1 summarizes component skills identified in listening comprehension skill and CALS. Listening comprehension skill, as a discourse comprehension skill of oral language texts, draws on foundational cognitions (working memory, attention, and inhibitory control), foundational oral language skills (vocabulary and grammatical/morphosyntactic knowledge), higher order cognitive skills and regulation (inference, perspective taking, and comprehension monitoring), and background knowledge (content and discourse knowledge, which includes text structure, genre, and register knowledge; see Kim, 2016, 2019 for details about theoretical conceptualization). Academic language proficiency according to Uccelli, Barr, et al. (2015) also includes similar skills. Specifically, it requires foundational oral language skills, namely vocabulary and syntactic knowledge, which are also component skills of listening comprehension (Kim, 2016). Anaphor resolution, one of the CALS domains, taps into referential inferencing ability that is also important in listening comprehension (Florit et al., 2011; Graesser, Singer, & Trabasso, 1994; Kim, 2016). Connecting ideas logically, another CALS domain, would require making inferences about the chain of causes and consequences, which is likewise necessary for listening comprehension (Kim, 2016, 2017; Lepola et al., 2012; Tompkins, Guo, & Justice, 2013). Interpreting and understanding authors' viewpoints in CALS require perspective taking, one of the higher order skills needed for listening comprehension as well (Kim, 2015, 2016, 2019).

If the constituent skills of listening comprehension skill and academic language proficiency largely overlap, then a question that arises is the dimensionality of these two skills. In other words, if they draw on a highly similar set of skills, then are the listening comprehension skill and academic language proficiency tasks best described as a unidimensional skill, related but dissociable skills (correlated multidimensions), or a common ability with specific aspects (e.g., a general language ability and unique residual listening comprehension and academic language abilities over and above the common general language ability; multidimensional with a bifactor structure; Reise, Moore, & Haviland, 2010). Previous studies have examined dimensionality of oral language skills such as vocabulary, syntax, and listening comprehension (Kieffer, Petscher, Proctor, & Silverman, 2016; Language and Reading Research Consortium, 2015; Tomblin & Zhang, 2006). In addition, Uccelli, Barr, et al. (2015) reported that academic language proficiency tasks largely capture a single dimension for children in upper elementary and middle schools when using an exploratory factor analysis. However, to our knowledge, no previous studies, regardless of age groups, have investigated the dimensionality of academic language proficiency and listening comprehension skill tasks simultaneously.

Another vital question is the relations of identified dimensions across listening comprehension skill and academic language proficiency tasks to reading comprehension, the second gap. As reviewed above, separate lines of work have demonstrated the relations of listening comprehension skill and academic language proficiency to reading comprehension, respectively. Given the overlap that is outlined above, listening comprehension skill and academic language proficiency tasks might be largely capturing a common ability. Alternatively, if listening comprehension and academic proficiency tasks are found to be multidimensional constructs—either two related but dissociable constructs, or a common ability with additional specific constructs (something uniquely specific to listening comprehension and to academic language proficiency over and above what is captured in the common ability)—then, how are these identified multiple dimensions related to reading comprehension? Answers to this would help integrate work in separate lines and bring clarity and coherence to our understanding about the relation of two widely studied constructs, listening comprehension skill and academic language proficiency, to reading comprehension skill.

The third gap is work on academic language proficiency for primary grade children. Academic language proficiency has been widely examined for preadolescents or adolescents, which is reasonable given the increased language demands in texts in upper elementary and secondary school grades (Meneses et al., 2018; Schleppegrell, 2001, 2004; Snow & Uccelli, 2009; Uccelli, Barr, et al., 2015). If academic language proficiency is considered an important skill that can be measured and that contributes to reading skills for adolescents, a crucial corollary is its measurement and relation to reading comprehension for younger children. Academic language is not an all-or-nothing phenomenon but instead emerges and develops over time; therefore, measurement of academic language proficiency for primary grade children is an important effort to advance the field.

A measurement aspect that needs to be considered especially when working with younger students is modality of administration (independent reading vs. oral language), the fourth gap in the literature. Studies with children in upper elementary grades or in secondary schools assessed academic language proficiency in *reading* contexts (i.e., children were asked to read questions and items; Phillips Galloway & Uccelli, 2019; Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015). This assumes and works only if students' reading skills are proficient enough not to influence their performance on academic language proficiency tasks. If reading skills are not proficient, students' performance level on academic language proficiency tasks would be depressed to some extent, particularly for developing readers or anyone who is struggling with decoding, because decoding skill constrains one's performance. Consequently, the academic language proficiency construct itself would not be purely academic "language" proficiency but would reflect decoding skill as well. This has important implications for validity of assessment of academic language proficiency depending on the assessment purpose. If the goal of the assessment is to accurately measure performance level of academic "language" proficiency, assessment in reading contexts may not be appropriate, particularly for developing readers. In contrast, if the inference from the academic language proficiency tasks is students' academic language performance that includes their decoding skill or if the goal of the assessment is prediction of one's reading performance (success or failure in reading comprehension), assessment in reading contexts may be appropriate because decoding is an important skill for reading comprehension.

In the present study, we sought to examine the effect of assessment modality of academic language proficiency in two ways. First, we estimated mean differences in students' performance on academic language proficiency tasks in oral versus reading contexts. One can certainly speculate that mean performance in a reading condition would be lower than that in an oral language condition for developing readers, but it is important to have concrete empirical estimates of performance differences. Another way of identifying the extent to which students' performance on academic language proficiency tasks is influenced by word reading skill is to examine the unique contribution of word reading to reading comprehension over and above academic language proficiency. If academic language proficiency captures word reading skill (i.e., assessment in reading contexts), the unique, independent contribution of word reading skill over and above academic language proficiency would be smaller or reduced compared to when academic language proficiency is measured in oral language contexts. To our knowledge, no previous studies have examined the impact of assessment modality of academic language proficiency (i.e., reading vs. oral contexts) regardless of students' reading developmental phrases.

To summarize, in the present study we address gaps in the literature about academic language and listening comprehension skills, and their relations to reading comprehension with the following three research questions.

1. Is student performance on listening comprehension and academic language tasks best described as unidimensional or multidimensional constructs?

2. Does second graders' performance on academic language tasks vary by assessment modality (reading vs. oral language context)?

3. What is the average unique variance explained in reading comprehension by identified dimension(s) of listening comprehension and academic language proficiency tasks, and word reading? Do the relations vary as a function of assessment modality?

These questions were addressed, using data from two cohorts of English-speaking children in Grade 2. Academic language tasks were adapted for primary grade children (i.e., Grade 2) from Uccelli, Barr, et al.'s (2015) CALS Instrument (CALS-I) designed for upper primary and middle schoolchildren (see below for details). The two cohorts were administered listening comprehension, word reading, and reading comprehension tasks in an identical manner with an exception for CALS-I, which was administered in a reading context in one cohort and in an oral language context in the other cohort. Dimensionality was systematically examined by fitting a series of alternative models, including unidimensional and multidimensional models (see Data Analysis Strategy section).

We hypothesized that the academic language and listening comprehension tasks would be best captured by a general ability that captures common variance across all tasks along with factors that are specific to academic language and listening comprehension skills or lexical, text, or register factors (i.e., specific factors). We also hypothesized that overall performance levels in academic language tasks would be lower when assessed in a reading context because of the constraint placed by children's word reading skills, given that second graders, on average, are developing their word reading skills. Finally, we anticipated that across assessment modality (oral or reading contexts), the overall general oral language factor, not the specific factors, would be most strongly and consistently related to reading comprehension after accounting for word reading skill, and that the unique relation of word reading to reading comprehension would be weaker when academic language proficiency is assessed in the reading condition.

Method

Participants

Participants were 350 second-grade students from 30 classrooms and seven schools in the Southeastern United States (M_{age} = 7.54; SD = .64). The sample was drawn from a larger longitudinal study of children's language and literacy development (see Kim, 2017). The sample was composed of two cohorts of children that were approximately equivalent in size (N = 165 for Cohort 1; N = 185 for Cohort 2). Participants were recruited from the same schools in two consecutive years. The two cohorts differed in the way academic language assessments were administered. In Cohort 1, the academic language assessments were administered as reading tasks (i.e., children had to read items and mark answers) whereas in Cohort 2, identical tasks were administered in an oral language condition (i.e., an assessor read items to them while children followed along with the assessor then answered in oral language). The reading condition was used first in Cohort 1 as a default approach, following previous studies by Phillips Galloway & Uccelli, 2019; Uccelli, Barr, et al. (2015; Uccelli, Galloway, et al., 2015). All other assessments were administered identically, and both cohorts were comparable in all other areas tested as well as in their sociodemographic composition. Human subject approval was obtained from the Florida State University (HSC NO. 2014.13495). Participating schools used *Imagine it!* (Bereiter, 2010) as a reading program or curriculum.

Sex distribution was approximately equal in both cohorts (51% male in Cohort 1; 53% male in Cohort 2). Race/ethnicity characteristics of the students were similar in Cohorts 1 and 2: White (54% and 52%, respectively), Black (35% and 34%), Hispanic/Latino (6% and 7%), Asian (1% and 1%), Native American (1% and 1%), and multiracial (3% and 5%). A large proportion of the children were eligible for free or reduced-price lunch in both cohorts (72% and 76% for Cohorts 1 and 2, respectively). According to school records, 1% of Cohort 1 and 3% of Cohort 2 students were identified as English language learners, which is consistent with the demographic characteristics of the region. English language proficiency status is determined by the state-wide assessments that are conducted annually.

Measures

Children were assessed on the constructs of academic language proficiency, listening comprehension skill, word reading skill, and reading comprehension skill, using multiple tasks for each construct. Unless otherwise noted, children's responses were scored dichotomously (1 = correct; 0 = incorrect) for each item. All the reported reliabilities, including those for normed tasks, are based on the present sample.

Academic language. The following seven tasks were used to capture the lexical, text, and academic register factors of academic language ability: academic vocabulary, academic register, organizing texts, tracking referents (or anaphor resolution), connecting ideas, understanding viewpoints: modal verbs, and understanding arguments. All tasks except for academic vocabulary were adapted from Uccelli and her colleagues' (Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015) CALS-I designed for older children (i.e., Grades 4–6). Adaptation was led by the first author and two doctoral students who were former classroom teachers on the research team. Adaptation primarily involved developing easier items, appropriate for younger children in the present study, for the constructs noted above. For instance, the Connecting Ideas task in Uccelli, Barr, et al.'s (2015) work included connectives such as *nonetheless* and *in spite of that*, but in the adapted task, easier connectives such as *in addition* and *unless* were included. In the adapted Organizing Texts task, text lengths were shorter (limited to three to five sentences compared to four to six sentences in CALS-I), and many items included transition words for clear signposting (see the example item in Appendix A). In the Understanding Viewpoints: Modal Verbs task, the prompt sentences included modal verbs, adjectives, and adverbs that are likely to be familiar to children in lower primary grades (e.g., *will, might, possible*) instead of more sophisticated ones used in CALS-I (e.g., *unquestionable, unlikely*). The adapted tasks were piloted with children in Grades 2 to 4 (approximately 30–40 children per grade) and revised (e.g., changing directions, changing items) by the authors based on psychometric properties and assessors' feedback from the field. Below is a brief description of each task (detailed descriptions of many of the tasks are found in Uccelli, Barr, et al., 2015; Uccelli, Galloway, et al., 2015). Note that for each task, there were two practice items, followed by the experimental items. Appendix A presents example items and the total number of experimental items for each task used in the present study.

In the Academic Vocabulary task (14 items), which was adapted from Kieffer and Lesaux (2012), the child was asked to select the meaning of a target word out of four options. In the Academic Register task (10 items), the child was presented with definitions of a familiar word (e.g., *car*) and was asked to identify whether the definition provided was written for children or adults. In the Organizing Texts task (six items), the child was provided with three to four sentences and was asked to come up with a reasonable, logical sentence that fit the context. Texts explicitly included conventional markers (e.g., first, next) or paragraph organizations with clear cue words (e.g., Weather in Alaska is very *different* from weather in Florida.). In the Tracking Referents task (seven items), the child was asked to identify the expression to which an anaphor (e.g., it, this) referred in a provided short text. In the Connecting Ideas task (10 items), the child was asked to select a missing connective to complete a sentence from among four options. In the Understanding Viewpoints: Modal Verbs task (16 items), the child was presented with a sentence containing a modal verb (e.g., There *will* be a quiz tomorrow; There *might* be a quiz tomorrow) and was asked to indicate the perspective of the student expressed in the sentence. In the Understanding Arguments task (11 items), the child was provided with a sentence that contained a claim (e.g., Students need recess to play and relax at school) and a student's response to the argument (e.g., Yes. At recess we play games like tag, soccer, or basketball). The child was asked to identify the nature of the student's response among four options (e.g., giving examples, excuses, definitions, or exaggerations; see Appendix A).

Overall internal consistency of scores from the academic language tasks was .88. Because our second research question is centered on the dimensionality (factor structure) of item responses from the set of academic language tasks, coefficient H (McNeish, 2018) for the subsequent factors are reported in our section on confirmatory factor analysis.

Listening comprehension. Children's listening comprehension was measured by two standardized, normed tasks and one experimental informational task. The normed tasks included the Narrative Comprehension subtest of the Test of Narrative Language (TNL; Gillam & Pearson, 2004) and the Listening Comprehension Scale of the Oral and Written Language Scales–II (OWLS-II; Carrow-Woolfolk, 2011). In the TNL Narrative Comprehension subtest, the child heard three narrative stories and was asked comprehension questions for each story (e.g., What should they [characters] do?). There were a total of 30 questions across the three stories ($\alpha = .74$). In the OWLS-II Listening Comprehension task, the child listened to stimulus sentences and short passages and was asked to point to one of four pictures that corresponded to the sentence ($\alpha = .94$). The experimental comprehension task was composed of three informational passages (140 words, 200 words, and 282 words, respectively) from the Qualitative Reading Inventory-5 (QRI; Leslie & Caldwell, 2011) and the Analytical Reading Inventory, 9th ed. (Woods & Moe, 2011). Titles of the three passages were *Matter*, *Whales and Fish*, and *Where Do People Live?* After listening to each passage, the child was asked comprehension questions (e.g., According to the story, how are whales and fish different?). There were eight questions in each passage for a total of 24 questions ($\alpha = .72$). The items in all the tasks were scored dichotomously except for a few items in TNL, which were scored 0 to 3 following the manual (Gillam & Pearson, 2004).

Word reading. Three standardized and normed tasks were used to measure children's word reading proficiency: the Letter Word Identification (LWID) of the Woodcock Johnson, 3rd ed. (WJ-III; Woodcock, McGrew, & Mather, 2001) and two forms (A & B) of the Sight Word Efficiency subtask of the Test of Word Reading Efficiency-II (TOWRE-II; Wagner, Torgesen, & Rashotte, 2012). In the WJ-III LWID task, the child was asked to read aloud a list of words of increasing difficulty ($\alpha = .91$). In the TOWRE-II Sight Word Efficiency task, the child was asked to read words of increasing difficulty with accuracy and speed in 45 s (test-rest reliability = .93; Wagner et al., 2012).

Reading comprehension. The following two standardized and normed tasks were used to assess children's reading comprehension: Passage Comprehension of the WJ-III (Woodcock et al., 2001) and Reading Comprehension of the Wechsler Individual Achievement Test-Third Edition (WIAT-III; Wechsler, 2009). In the WJ-III Passage Comprehension task, the child read sentences and short narrative and informational passages and filled in blanks ($\alpha = .83$). In the WIAT-III Reading Comprehension task, the child read narrative and informational passages and answered multiple choice questions ($\alpha = .82$).

Procedure

Children were individually assessed by rigorously trained research assistants in a quiet space in the school. The assessment battery was administered in several sessions, each of which lasted 30 to 40 min. As noted above, all assessment procedures were identical for both cohorts of children with an exception for academic language proficiency tasks, which were administered in a reading context for Cohort 1 and an oral language context for Cohort 2.

Data Analysis Strategy

Research Question 1: Dimensionality of academic language and listening comprehension tasks. The extent to which the item responses from the academic language and listening comprehension tasks converged on a single construct or multidimensional constructs was evaluated systematically via confirmatory factor analysis (CFA). Because the items across tasks were dichotomous, it was necessary to use a CFA of categorical indicators, an analysis that is structurally equivalent to item response theory (Kamata & Bauer, 2008). As mentioned above, multiple alternative models were tested (see Figure 2). First, a unidimensional model (Figure 2a; Model 1) was tested. In this model, all items from the seven academic language subscales along with items from the three listening comprehension tasks (i.e., the TNL, OWLS-II, and experimental informational text comprehension) were specified to load on a general oral language factor.

A second set of models (Models 2–5) tested a bifactor structure. The second model (see Figure 2b; Model 2) was a bifactor model where items loaded on a general construct (i.e., oral language [OL in Figure 2]) that captures common variance across all the items, as well as specific factors over and above the general construct—text factor, lexical factor, and register factor—from the seven academic language proficiency tasks, and specific factors for the three listening comprehension tasks. In this model, variation in item responses across all tasks is due to a general and shared trait, and residual variance in item responses beyond the general factor (i.e., specific factors) may be further due to something shared within the tasks such as text, lexical, and register awareness factors or passage-specific listening comprehension. Text-level processes involve multiple text-based inferences such as anaphor resolution, identification of themes, inferences about textual stance or viewpoint, and intent and goals (Graesser et al., 1994; Kim, 2016, 2017; Kintsch & Rawson, 2005). Register awareness is also one of the hallmarks of academic language proficiency (academic register vs. informal/conversational register) because it entails the ability to recognize the constellation of academic language features as appropriate for academic contexts (Carlisle et al., 1999; Scarcella, 2003; Snow & Uccelli, 2009). The lexical factor was indicated by the academic vocabulary items; the text factor was indicated by organizing texts, tracking referents, connecting ideas, modal verbs, and understanding arguments tasks; and the register factor was indicated by the items in the academic register task. These specific factors reflect variance that was captured among the items but that is statistically not relevant to the general factor from a malleable factor perspective. For the items that corresponded to the three TNL, three experimental informational texts, and OWLS-II passages, items within a passage may include response-dependencies that can be modeled via latent factors. These latent factors are not believed to have inherent value for predictive purposes but serve to capture potential item dependence within an item bundle. The general factor is hypothesized to maintain moderate to strong loadings in the measurement model to comprise a reliable factor and is further hypothesized to reflect a substantive hypothetical construct compared to the specific factors. Advantages of bifactor models include testing for the presence of specific constructs beyond the general constructs, and allowance of both general and specific factors as explanatory variables of individual differences in other outcomes (Reise et al., 2010).

A variant of the bifactor model in Figure 2b replaced the theoretical specific factor of text with task-specific factors (Figure 2c; Model 3). Model 3 replaced the single, latent text factor (shown in Figure 2b) with multiple, specific factors corresponding to each task-based item set (i.e., a latent variable each for items associated with organizing texts, tracking referents, connecting ideas, modal verbs, and understanding arguments tasks; see Figure 2c). This model was tested because it is plausible that the residual variances across items beyond the general oral language factor are not best captured by a theoretical specific text factor but rather multiple task-based factors that have no meaningful variance beyond the general factor. The fourth model was also a bifactor model, but instead of one general oral language factor (shown in Figure 2b), two general factors of academic language and listening comprehension were specified (see Figure 2d; Model 4). The academic language factor tested the extent to which the academic language items characterized a general factor beyond the text, lexical, and register specific factors. Likewise, the general listening comprehension factor tested whether the items from the TNL, informational texts, and OWLS-II passages loaded on a general factor in addition to the passage-specific effects. Model 5 (Figure 2e) tested a variant of Model 4 where the text factor was replaced by the multiple task-based factors.

Model 6 (Figure 2f) and Model 7 (Figure 2g) are trifactor models. A trifactor model is similar in its conceptualization to the bifactor model; in the bifactor model, each measured variable (or item in this study) loads onto one specific factor and one general factor. The trifactor, as used by Bauer et al. (2013), allows each measured variable to load on three target factors. For example, Model 6 in our study was a trifactor model (Figure 2f) that included a factor of oral language (independent of the other factors such as academic language, listening comprehension, text, lexical, register, and passage factors), factors of academic language and listening comprehension (over and above the general language factor), and the text, lexical, register, and listening comprehension passage factors. In the bifactor model, such as in Figure 2c, conventional statistical modeling practices refer to the oral lan-

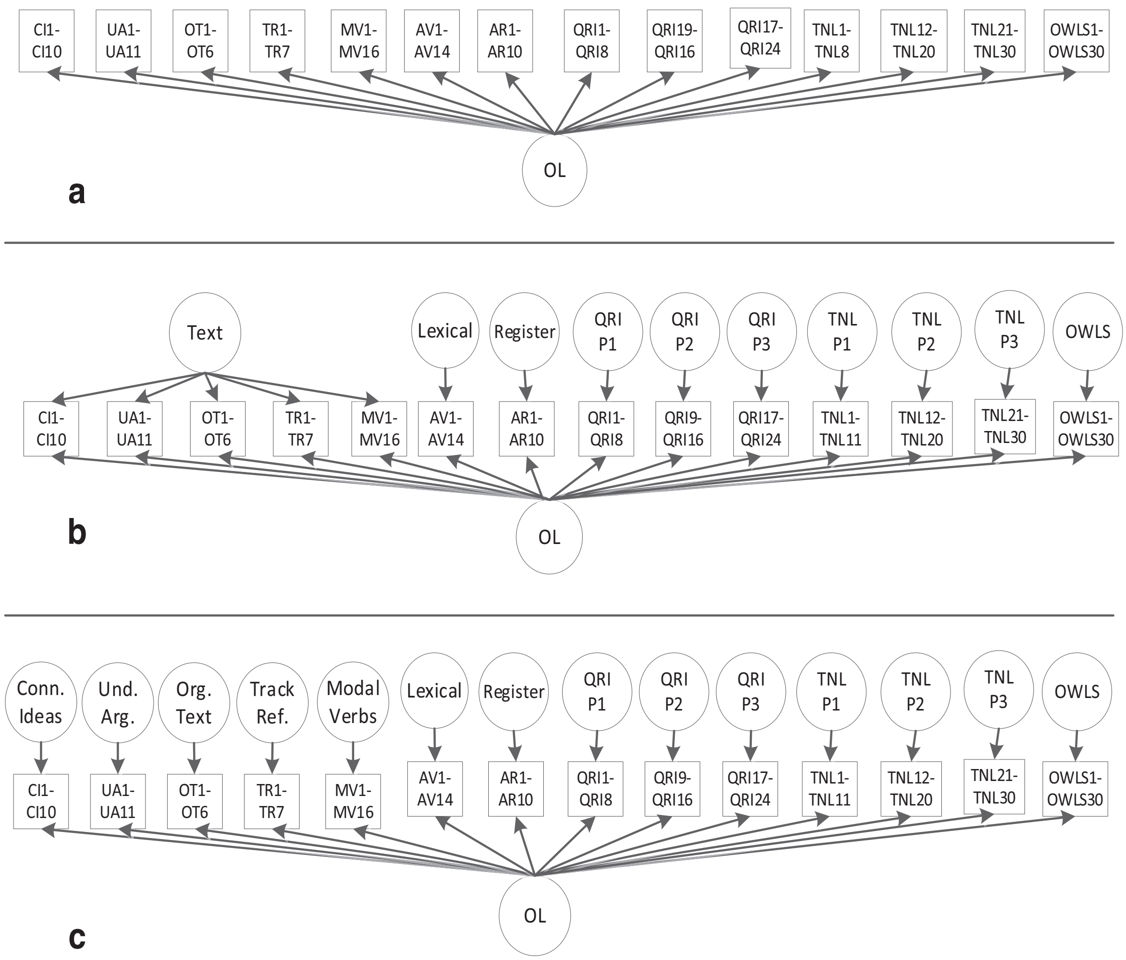

Figure 2. Measurement model sample specifications for (a) a unidimensional academic language model (Model 1); (b) a bifactor model of general oral language (OL) and specific text, lexical, register, and listening comprehension passage factors (i.e., QRI, TNL, and OWLS; Model 2); (c) a bifactor model of general OL and specific connecting ideas, understanding arguments, organizing texts, tracking referents, modal verbs, lexical, register, and listening comprehension passage factors (Model 3); (d) a bifactor model of general factors of academic language (AL) and listening comprehension (LC), and specific text, lexical, register, and listening comprehension passage factors (Model 4); (e) a bifactor model of general factors of AL and LC, and specific factors of connecting ideas, understanding arguments, organizing texts, tracking referents, modal verbs, lexical, register, and listening comprehension passage (Model 5); (f) a trifactor model of general OL, local factors of AL and LC, and specific text, lexical, register, and listening comprehension passage factors (Model 6); and (g) a trifactor model of general OL, local factors of AL and LC, and specific connecting ideas, understanding arguments, organizing texts, tracking referents, modal verbs, lexical, register, and listening comprehension passage factors (Model 7). CI = connecting ideas; UA = understanding arguments; OT = organizing texts; TR = tracking referents; MV = modal verbs; AV = academic vocabulary; AR = academic register; TNL = Test of Narrative Language; OWLS = Listening Comprehension subtest of the Oral and Written Language Scales; QRI = Qualitative Reading Inventory-5 used as expository text comprehension.

guage factor as a general factor and the task-based factors are called specific factors. As previously described for Models 2–5, this terminology reflects that the general factor of oral language is theoretically expected to capture the *general* common variance across all items and the remaining factors capture residual variance *specific* to the tasks. In the trifactor model, such as Figure 2f, we again describe oral language as a general factor and the task-based factors are still referred to as specific factors. The introduction of the academic language and listening comprehension factors for the trifactor model necessitates a terminological discussion. These two factors are not general factors because they do not capture the variance across all items, yet they are not quite specific factors either as academic language is comprised of items from text, lexical, and register items, and listening comprehension is comprised of items from the QRI, TNL, and OWLS-II passages. We opt to refer to these latent representations as *local factors* to reflect that they are not general but also not hypothesized or measured in the same way as the specific factors. The local factors are theo-

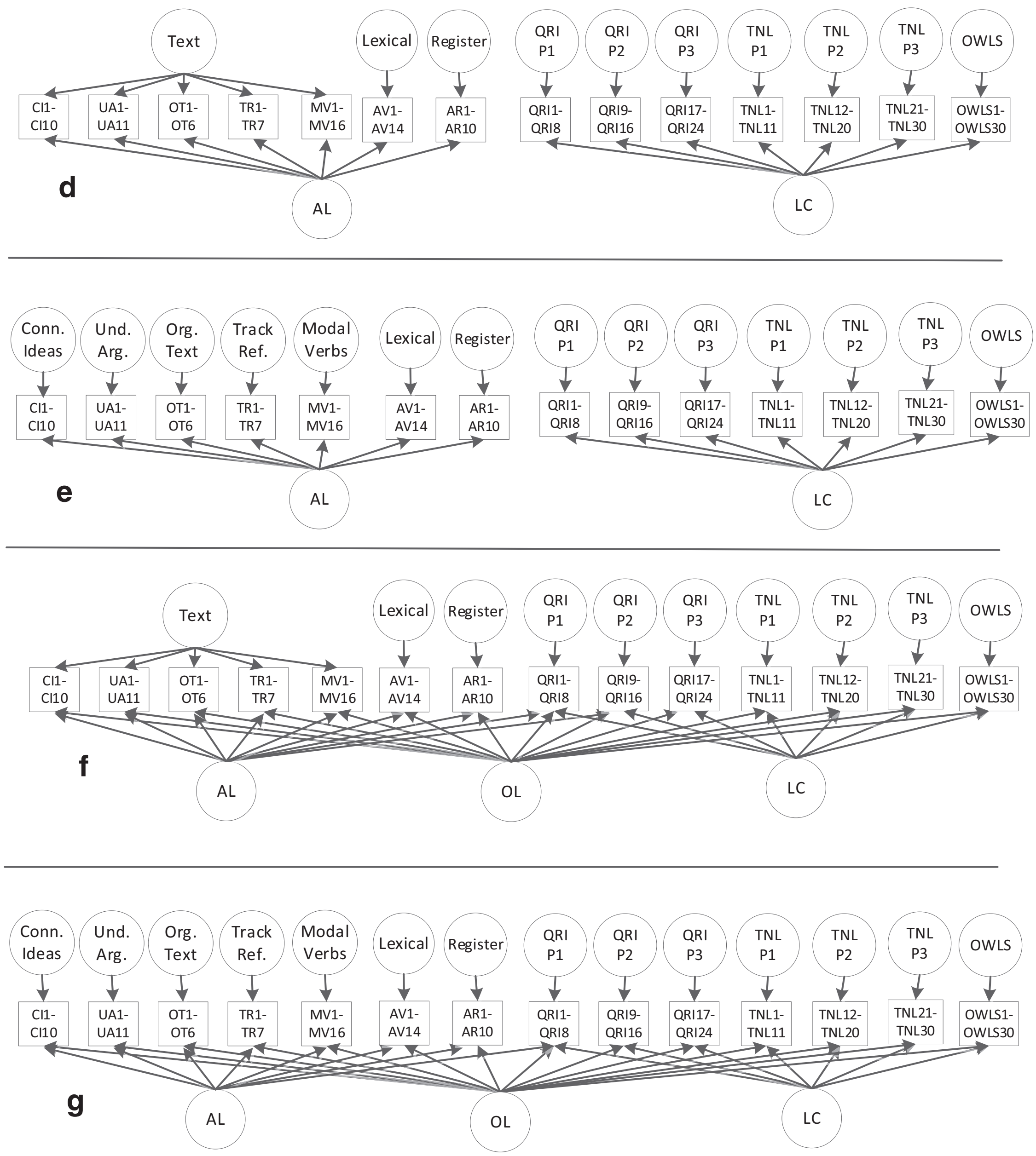

Figure 2. (continued)

retical aggregates of content-specific areas but do not cross each other (i.e., academic language items only load on the academic language local factor and listening comprehension items only load on the listening comprehension local factor). Using the term local factor allows us to delineate important theoretical and statistical differences between the bifactor and trifactor models. Model 7 was also a trifactor model (Figure 2g) that replaced the text factor with multiple task-specific factors.

Models 8 and 9 tested second-order factor models (Figures not shown) where Model 8 explored the presence of a second-order factor over the text, lexical, register, and listening comprehension factors, and Model 9 explored the presence of a second-order factor over the organizing texts, tracking referents, connecting ideas, modal verbs, understanding arguments tasks, and listening comprehension factors. In addition to these nine primary models, we explored several variations that allowed for correlated factors. Because of parsimony in reporting, details on these ancillary models and results are reported in the online supplemental materials.

All models were estimated using Mplus 8.1 software with the logit-link via maximum likelihood estimation. Relative model fit among the specifications was compared using the Akaike information criterion (AIC) and the sample-adjusted Bayes information criterion (nBIC). The latter index was used as simulation studies have suggested that the nBIC is useful for model comparisons

(Enders & Tofighi, 2008). For the model demonstrating best relative fit, incremental fit was evaluated with the comparative fit index (CFI; > .95 as acceptable), Tucker-Lewis index (TLI; >.95 as acceptable), and root mean square error of approximation (RMSEA; < .10 as acceptable). The explained common variance index (ECV; Rodriguez, Reise, & Haviland, 2016) was used to evaluate the variance related to the general factors by taking the ratio of the variance explained by the general factors to the total variance across all general or specific factors. The value estimated by the ECV represents the percentage of the common variance that is extracted among all the indicators that is due to the general factor(s) with the 1 − ECV value as the percentage of the common variance that is due to specific factors. Moreover, because of the complexity and multidimensionality of the models that were tested with 128 items in our sample of 350, each factor model was bootstrapped 100 times so that the standard errors could be better approximated.

Research Question 2: Performance level in the academic language proficiency tasks as a function of assessment modality. Students' mean performances on the academic proficiency tasks (i.e., oral language vs. reading modality) were compared, using Cohen's *d*.

Research Question 3: The relations of identified dimension(s) to reading comprehension and moderation by assessment modality. The primary analytic strategy to address this question was dominance analysis, using results from the second research question. Furthermore, prior to conducting dominance analysis, CFA models for word reading and reading comprehension latent variables were added to the trifactor model to obtain the covariance matrix needed in the dominance analysis (see the online supplemental materials). A latent factor of word reading was composed of WJ-III LWID and two forms of the TOWRE-II, and a latent factor of reading comprehension was formed by WJ-III Passage Comprehension and the WIAT-III. Separate factor models were run for each cohort. Because of known limitations related to using estimated factor scores, such as factor indeterminacy (Beauducel & Hilger, 2017) that is impacted by sample size, reliability of measurement, and missing data, we opted to use the actual covariance matrix from the full CFA to conduct a secondary-analysis via a dominance analysis. Using R macro scripts developed by Nimon (2015), data sets with sample sizes of 165 and 185 for Cohorts 1 and 2, respectively, were simulated according to their separate covariance matrices.

Dominance analysis (Azen, 2013; Azen & Budescu, 2003) allows a rank ordering by importance of the contribution of predictors to outcome by modeling all possible main effects regression models and evaluating the unique contributions a predictor makes when considering models which have 0 to *n* variables already included. These authors note that several types of measures exist for evaluating predictor importance including (a) a regression coefficient, (b) a partially standardized regression coefficient, (c) R^2, or (d) various other indices such as the Pratt index (Pratt, 1987) or the structure coefficient (Courville & Thompson, 2001). The limitation of these indices in traditional regression is that they are limited by interpretation (Azen & Budescu, 2003), but more importantly how predictor importance should be evaluated in a set of models. Azen and Budescu argued that one predictor is valued as more important if it has a greater contribution in all other possible subset models where only one of the two variables is entered into the equation. For example, suppose that listening comprehension skill ($X1$), decoding ($X2$), and academic language proficiency ($X3$) were all used as predictors of reading comprehension. Dominance analysis uses $2^{(p-2)}$ subset models to listening comprehension skill all possible R^2 from the included predictors (p). To look at the unique value of listening comprehension, one would take the R^2 from any subset model that includes listening comprehension skill and subtract the R^2 from a model not including listening comprehension skill. The unique value of listening comprehension skill when only 1 other predictor in the model would be evaluated by looking at the following model comparisons using R^2:

Modeling Process 1: $(X1+X2) - (X2) =$ unique value of listening comprehension skill above decoding.

Modeling Process 2: $(X1+X3) - (X3) =$ unique value of listening comprehension skill above academic language proficiency.

The averaged R^2 across the two modeling processes represents the unique value of listening comprehension skill when one other predictor is in the model. This process repeats looking at listening comprehension skill when no other predictors are in the model or all other predictors are in the model. The averaged R^2 across all possible models reflects the average unique contribution of the variable. Without dominance analysis, a researcher is left with relatively few options for predictor importance as the unique R^2 for any model is confounded by its order of entry. In the example above, the importance of listening comprehension skill differs if decoding is entered first into the model compared to if academic language proficiency is entered first. As the number of predictors increases for a regression analysis, dominance analysis becomes increasingly important as a tool to test the relative importance of variables. The average R^2 contribution and unique contribution to size *k* models was evaluated and reported as was the level of dominance of each predictor in pairwise contrasts of the predictors.

Furthermore, levels of dominance were established by reviewing two statistics from 1,000 bootstrapped estimates to establish one of three levels of dominance among the predictors. First, the dominance of one predictor (*i*) over another (*j*) is viewed via the D*ij* coefficient that denotes whether the level of dominance between each *ij* pairing can be observed. A D*ij* coefficient of 0.5 means that neither variable dominates the other; 1 indicates that the variable *i* dominates variable *j*; and 0 means that the variable *j* dominates variable *i.* The reproducibility coefficient, denoted R, indicates the percent of bootstrapped conditions where the level of dominance was observed; values of 1.00 are reflective of the observed dominance occurring in 100% of bootstrapped samples, and lower values show weaker confidence in the level of observed sample dominance based on the bootstraps. The reason the D*ij* and R coefficients are helpful is that the D*ij* gives the direct, summative comparison of how one predictor compares to another in explaining unique variance. The R^2 values from the dominance is helpful for reviewing contributions of predictors across all subset models, but the D*ij* succinctly evaluates the head-to-head comparison of predictors. Similarly, the R coefficient is a useful statistic for looking at the generalizability of D*ij*. That is, if a D*ij* value of 1.0 is observed along with an R of .48 when testing for complete dominance, then we may conclude that predictor *i* completely dominates predictor *j* with the sample data. However, when the sample data were bootstrapped 1,000 times, complete dominance was only observed in 480 of the 1,000 bootstraps. Subsequently, the generalizability of the complete dominance would be weaker in the presence of a lower R coefficient. Complete dominance describes where a predictor *i* explains more unique variance for *every* subset model than predictor *j* (or vice versa). Conditional dominance is where the average additional contribution *within* model size *k* (e.g., one predictor only, two predictors only, etc.) for predictor *i* is greater than predictor *j* (or vice versa). General dominance is the weakest form of dominance where the average additional contribution *across* all model sizes for predictor *i* is greater than predictor *j* (or vice versa). Because the bootstrap process for dominance analysis evaluates pairwise dominance relations using variable *i* compared to variable *j*, each of the predictors can be represented as either an *i* or a *j* variable. SAS software was used to estimate the dominance analysis of latent variables.

Results

Data and Preliminary Analysis

Prior to the estimation of the descriptive statistics and correlations, missing data were evaluated for the measures across the two cohorts. Missing data ranged from 0% on the TNL to 4% for the items on the Organizing Texts, Tracking Referents, Academic Vocabulary, and Academic Register tasks. Little's testing of missing completely at random was not statistically significant for either cohort—Cohort 1: $\chi^2(14) = 22.55$, $p = .068$; Cohort 2: $\chi^2(1) = 0.29$, $p = .593$—indicating that the hypothesis of data missing completely at random was not rejected. We used full-information maximum likelihood, which is efficient in handling missing data (Enders & Bandalos, 2001).

Descriptive statistics for the sample by cohort are reported in Table 2. Children's mean performances on the normed measures of language and reading (i.e., TNL, OWLS-II, TOWRE-II, WJ-III LWID and Passage Comprehension, and WIAT-III Reading Comprehension) were in the low average to average range across both cohorts. For example, in the listening comprehension tasks, students' mean performances on TNL ranged from 8.30 to 8.65 (in TNL, mean of standardized score is 10 with $SD = 3$) and mean performances on OWLS-II ranged from 97.05 to 98.41. Students' performances on reading tasks ranged from 96.36 on WJ Passage Comprehension to 103.53 on WJ-III LWID in Cohort 1 and from 96.82 on WIAT-III Reading Comprehension to 104.76 on WJ-III LWID in Cohort 2. Children's mean performances on the academic language tasks had sufficient variations around means. There was a slight floor effect in the Organizing Texts task, but skewness was not severe (.72). Raw scores were used in subsequent analyses.

Table 2
Descriptive Statistics for Cohorts 1 and 2

	Cohort 1 (reading condition: *N* = 165)				Cohort 2 (oral language condition: *N* = 185)				
Measure	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max	*d*
Academic language									
Academic vocabulary	7.23	2.79	1	13	8.52	2.77	1	13	−.46
Academic register	5.37	2.15	0	10	5.64	1.87	2	10	−.13
Organizing texts	1.69	1.56	0	6	1.53	1.54	0	6	.10
Tracking referents	3.72	1.88	0	7	3.93	1.62	0	7	−.12
Connecting ideas	5.19	2.37	1	10	5.36	1.98	0	10	−.08
Understanding viewpoints: Modal verbs	7.98	3.37	0	16	8.79	3.34	1	16	−.24
Understanding arguments	5.41	2.46	1	10	6.23	2.12	0	11	−.36
Listening comprehension									
TNL Raw	26.47	4.95	13	36	25.34	4.97	5	33	.23
TNL SS[a]	8.65	3.07	1	15	8.30	2.89	1	15	.12
OWLS Raw	76.90	12.89	37	103	76.08	13.16	38	101	.06
OWLS SS	98.41	14.32	44	124	97.05	15.15	44	124	.09
QRI Raw	10.25	3.49	2	20	9.21	3.40	1	20	.30
Word reading									
WJ LWID Raw	42.01	6.51	25	63	42.01	6.47	18	63	.00
WJ LWID SS	103.53	12.43	65	134	104.76	13.31	47	135	−.10
TOWRE Form A Raw	50.99	11.94	18	73	51.64	12.08	7	75	−.05
TOWRE Form A SS	97.39	15.38	55	128	100.32	16.64	55	131	−.18
TOWRE Form B Raw	51.55	11.71	17	78	52.01	12.18	9	78	−.03
TOWRE Form B SS	98.00	15.50	55	131	100.33	17.21	55	135	−.14
Reading comprehension									
WJ Passage Comp Raw	22.97	4.22	12	83	22.98	4.22	9	33	.00
WJ Passage Comp SS	96.36	11.23	57	119	97.33	12.11	44	122	−.08
WIAT Reading Comp Raw	51.74	11.35	3	83	50.94	11.30	3	83	.07
WIAT Reading Comp SS	96.58	13.23	40	136	96.82	13.18	58	138	−.02

Note. TNL = Test of Narrative Language; SS = standard score; OWLS = Listening Comprehension subtest of the Oral and Written Language Scales; QRI = Qualitative Reading Inventory-5 used as expository text comprehension; WJ = Woodcock Johnson-III; LWID = Letter Word Identification; TOWRE = Test of Word Reading Efficiency-II; Comp = comprehension; WIAT = Wechsler Individual Achievement Test-third edition; *d* = Cohen's *d* standardized mean differences.
[a] Norm mean is 10 with *SD* of 3 for the Test of Narrative Language.

Table 3
Correlations Among Measured Variables for Cohort 1 (N = 165; Reading Condition; Lower Diagonal) and Cohort 2 (N = 185; Oral Language Condition; Upper Diagonal)

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Academic vocabulary	—	.17	.39	.31	.50	.53	.41	.47	.50	.46	.33	.25	.26	.43	.38
2. Academic register	.19	—	.14†	.06†	.17	.16	.16	.15	.03†	.04†	.03†	.00†	.00†	.13†	.06†
3. Organizing texts	.46	.08†	—	.36	.41	.36	.33	.38	.31	.44	.33	.21	.22	.38	.44
4. Tracking referents	.61	.18	.44	—	.28	.32	.12†	.16	.18	.31	.21	.12†	.14†	.23	.18
5. Connecting ideas	.69	.09†	.44	.60	—	.50	.46	.41	.40	.36	.41	.33	.33	.48	.41
6. Viewpoints: modal verbs	.52	.11†	.42	.47	.51	—	.32	.37	.42	.37	.31	.23	.22	.38	.40
7. Understand arguments	.55	.14†	.44	.57	.63	.59	—	.45	.40	.32	.32	.22	.18	.39	.42
8. TNL	.41	.12†	.45	.35	.39	.45	.41	—	.44	.59	.30	.19	.17	.46	.41
9. OWLS	.32	.11†	.31	.27	.22	.21	.26	.44	—	.48	.32	.19	.14†	.35	.32
10. QRI	.45	.18	.46	.38	.49	.40	.46	.59	.37	—	.15	.03†	.04†	.29	.32
11. WJ LWID	.52	.07†	.35	.44	.62	.53	.49	.32	.16	.33	—	.87	.86	.89	.67
12. TOWRE A	.47	.01†	.30	.42	.63	.47	.50	.27	.07†	.26	.84	—	.96	.82	.63
13. TOWRE B	.43	.04†	.24	.37	.61	.40	.45	.17	.02†	.23	.84	.94	—	.80	.62
14. WJ Passage Comp	.55	.11†	.42	.53	.61	.52	.55	.45	.28	.44	.86	.77	.75	—	.68
15. WIAT Reading Comp	.47	.03†	.44	.44	.58	.47	.49	.42	.31	.43	.67	.64	.62	.71	—

Note. TNL = Test of Narrative Language; OWLS = Listening Comprehension subtest of the Oral and Written Language Scales; QRI = Qualitative Reading Inventory-5 used as expository text comprehension; WJ = Woodcock Johnson-III; LWID = Letter Word Identification; TOWRE = Test of Word Reading Efficiency-II; Comp = comprehension; WIAT = Wechsler Individual Achievement Test-third edition.
† $p > .05$.

Table 3 presents bivariate correlations. Patterns of bivariate correlations showed similar patterns between many measures for the two cohorts. However, some differences were found in magnitudes between academic language tasks and the other tasks. For example, Tracking Referents was more strongly related to other tasks in Cohort 1 (reading condition; $.18 \leq rs \leq .61$) than in Cohort 2 (oral language condition; $.06 \leq rs \leq .36$).

Research Question 1: Dimensionality of Academic Language and Listening Comprehension Tasks

As shown in Table 4, among the seven primary models tested, the trifactor model with general oral language, local factors of academic language and listening comprehension, and task-specific factors (Model 7) maintained the lowest *n*BIC (66,893.93) and the second lowest AIC (66,353.89). The model with the lowest AIC index was Model 3 (66,155.83; ΔAIC = −198.06), yet Model 7 outperformed Model 3 in terms of the *n*BIC (i.e., Δ*n*BIC = 1,318.08), our primary relative model fit comparison index. The model with the next lowest *n*BIC compared to Model 7 was Model 4 with Δ*n*BIC = 28.02, suggesting stronger relative fit. Therefore, Model 7 or Figure 2g was chosen as the final model (loadings, standard errors, and *p* values reported in the online supplemental materials). The model fit for Model 7 was excellent: $\chi^2(18{,}330) = 18{,}737.75$, $p = 1.00$, CFI = .97, TLI = .97, RMSEA = .008 (90% CI [.004, .011]), and the ECV value of .71 indicated that 71% of the common variance extracted from all items was attributed to the oral language (general) factor and academic language and listening comprehension (local) factors, and the remaining 29% of the common variance was spread across or due to the other 12 latent construct representations in Figure 2g.

Table 4
Confirmatory Factor Analysis Model Fit

Model	LL	ECV	AIC	nBIC	ΔnBIC
1. One factor (Figure 2a)	−34,043.07		68,888.15	69,161.92	
2. Bifactor model of general oral language with text and task-specific factors (Figure 2b)	−32,663.97	.33	66,519.94	66,926.85	2,235.07[a]
3. Bifactor model of general oral language with task-specific factors (Figure 2c)	−32,481.91	.49	66,155.83	68,212.01	−1,285.16[b]
4. Bifactor model with uncorrelated general factors with text and task-specific factors (Figure 2d)	−32,398.61	.52	66,381.22	66,921.95	4.90[b]
5. Bifactor model with uncorrelated general factors with task-specific factors (Figure 2e)	−32,687.73	.55	66,567.46	66,974.37	−47.52[b]
6. Trifactor model of general oral language with local factors, theoretical and task-specific factors (Figure 2f)	−32,422.75	.57	66,427.50	66,967.54	−40.69[b]
7. Trifactor model of general oral language with local factors, task-specific factors only (Figure 2g)	−32,385.94	.71	66,353.89	66,893.93	32.92[b]
8. Second-order model with text and task-specific factors	No convergence				
9. Second-order model with correlated local factors	No convergence				

Note. LL = log likelihood; ECV = explained common variance; AIC = Akaike information criterion; nBIC = sample-adjusted Bayes information criterion. Local factors = academic language and listening comprehension in trifactor models. ECV for Models 2 and 3 is based on the general oral language factor. ECV for Models 4 and 5 is based on the AL and LC factors. ECV for Models 6 and 7 is based on the AL, LC, and OL factors.
[a] Model compared to Model 1. [b] Model compared to Model 2.

A review of the specific factors such as text, lexical, register, the experimental informational comprehension, TNL, and OWLS-II factors suggested that they were largely residual variance unrelated to the general and local constructs as indicated by weak or nonsignificant standardized loadings (i.e., −.26 to .25). The exception to this case was the register factor where the loadings were all statistically significant (i.e., −.24, −.19, .25 to .56). As such, we hypothesized that this construct may have relevance for the purpose of explaining individual differences in reading comprehension. Coefficient H was computed for the general factor of oral language (H = .95), the two local factors of academic language (H = .88) and listening comprehension (H = .85), and the register factor (H = .73).

Research Question 2: Mean Performances on the Academic Language Tasks as a Function of Assessment Modality

We examined the academic language scores across Cohort 1 (assessed in the reading context) and Cohort 2 (assessed in the oral language context). As displayed in Table 2, some, mostly small, standardized differences were observed between cohorts in the academic language tasks, with somewhat lower performances in the reading condition (Cohort 1). Cohen's d estimates were as follows: −.08 in Connecting Ideas (connectives), −.12 in Tracking Referents, −.13 in Academic Register, −.24 in Understanding Viewpoints: Modal Verbs, −.36 in Understanding Arguments, −.46 in Academic Vocabulary, and .10 in Organizing Texts. Note again that only the academic language tasks were administered in different conditions (reading vs. oral language) for Cohort 1 and Cohort 2; all other tasks (i.e., listening comprehension, word reading, and reading comprehension) were measured identically between cohorts.

Research Question 3: Nature of the Relation of Identified Dimension(s) of Oral Language to Reading Comprehension

CFA models for word reading and reading comprehension by each cohort fit the data well (see the online supplemental materials), and the latent correlation matrix for each of the cohorts is provided in Appendix B. As shown in Appendix B, in Cohort 1 (academic language measured in reading condition), word reading had the strongest correlation with reading comprehension (φ = .88), followed by general oral language (φ = .62), local academic language (φ = .46), local listening comprehension (φ = .25), and specific register (φ = .08). The correlations in Cohort 2 (academic language measured in oral language condition) also showed that the strongest bivariate correlate to reading comprehension was word reading (φ = .91), followed by general oral language (φ = .64), and then to a lesser degree local academic language (φ = .17), specific register (φ = .12), and local listening comprehension factor (φ = −.07).

Cohort 1 dominance analysis (academic language assessment in reading context). Results for Cohort 1 (see Table 5) show that word reading appeared to contribute the most unique variance for all k predictor-size models. The unique variance explained was 52% on average and ranged from 27% when the other four predictors were in the model to 77% when there were zero additional predictors in the model. The general construct of oral language contributed the next greatest amount of unique variance, averaging 24% across all k models and ranging from 10% to 38%. The local academic language explained an average of 11% unique variance in reading comprehension, ranging from 0% when all other variables were in the model to 21%. The local listening comprehension explained an average of 6% unique variance with a range of 5% to 6%. The specific register factor did not uniquely contribute to reading comprehension.

Reproducibility coefficients in Table 6 show the result of 1,000 bootstrapped estimates to establish one of three levels of dominance among the predictors. The 1 to 5 numeration in columns I and J correspond to the predictors in the following way: I_1/J_1 = local factor of Academic Language, I_2/J_2 = local factor of Listening Comprehension, I_3/J_3 = general factor of Oral Language, I_4/J_4 = specific factor of Register, and I_5/J_5 = Word Reading. Results for Cohort 1 show that word reading (I_5/J_5) completely dominated local factors of academic language (Contrast 2) and listening comprehension (Contrast 4), general oral language (Contrast 7), and the specific factor of register (Contrast 11) with a reproducibility value of 1.00 for each of these contrasts. General oral language completely dominated local factors of academic

Table 5
Dominance Analysis Average Predictor Contributions of Each Predictor to k-Predictor Model Size by Cohort

Cohort and predictor	Average R-squared contribution					
	Overall average	0 predictor	1 predictor	2 predictors	3 predictors	4 predictors
Cohort 1						
Word reading	.52	.77	.65	.52	.39	.27
General oral language	.24	.38	.31	.24	.17	.10
Local academic language	.11	.21	.16	.11	.05	.00
Local listening comprehension	.06	.06	.06	.05	.05	.05
Specific register	.00	.00	.00	.00	.00	.00
Cohort 1						
Word reading	.68	.83	.76	.68	.61	.54
General oral language	.28	.41	.35	.28	.22	.15
Local academic language	.02	.03	.02	.02	.01	.00
Local listening comprehension	.01	.01	.01	.01	.01	.00
Specific register	.01	.00	.01	.01	.01	.00

Table 6
Predictor Dominance Relations and Reproducibility by Cohort

	Cohort 1					Cohort 2				
Contrast	Level of dominance	*I*	*J*	D*ij*	R	Level of dominance	*I*	*J*	D*ij*	R
1	Complete	1	2	.5	1.00	Complete	1	2	.5	.89
2	Complete	1	5	0	.00	Complete	1	5	0	.04
3	Complete	2	1	.5	1.00	Complete	3	1	1	1.00
4	Complete	2	5	0	.00	Complete	3	2	1	1.00
5	Complete	3	1	1	.98	Complete	3	5	0	.00
6	Complete	3	2	1	1.00	Complete	4	1	.5	.00
7	Complete	3	5	0	.00	Complete	4	2	.5	.00
8	Complete	4	1	0	.00	Complete	4	3	0	.00
9	Complete	4	2	0	.00	Complete	4	5	0	.00
10	Complete	4	3	0	.00	Complete	5	2	1	.07
11	Complete	4	5	0	.00	Conditional	1	2	1	.53
12	Conditional	1	5	0	.00	Conditional	1	5	0	.09
13	Conditional	2	1	.5	.99	Conditional	3	1	1	1.00
14	Conditional	2	5	0	.00	Conditional	3	2	1	1.00
15	Conditional	3	1	1	.99	Conditional	3	5	0	.00
16	Conditional	3	2	1	1.00	Conditional	4	1	0	.00
17	Conditional	3	5	0	.00	Conditional	4	2	1	1.00
18	Conditional	4	1	0	.00	Conditional	4	3	0	.00
19	Conditional	4	2	0	.00	Conditional	4	5	0	.00
20	Conditional	4	3	0	.00	Conditional	5	2	1	.36
21	Conditional	4	5	0	.00	General	1	2	1	.88
22	General	1	2	1	.95	General	1	5	0	.30
23	General	1	5	0	.00	General	3	1	1	1.00
24	General	2	5	0	.00	General	3	2	1	1.00
25	General	3	1	1	1.00	General	3	5	0	.00
26	General	3	2	1	1.00	General	4	1	0	.00
27	General	3	5	0	.00	General	4	2	1	1.00
28	General	4	1	0	.00	General	4	3	0	.00
29	General	4	2	0	.00	General	4	5	0	.00
30	General	4	3	0	.00	General	5	2	1	.74
31	General	4	5	0	.00	—	—	—	—	—

Note. $I = \chi i$; J = variable χj; D*ij* = dominance of χi over χj; R = reproducibility; *I*1/*J*1 = local academic language; *I*2/*J*2 = local listening comprehension; *I*3/*J*3 = general oral language; *I*4/*J*4 = specific register; *I*5/*J*5 = word reading.

language (Contrast 5, R = .98) and listening comprehension (Contrast 6, R = 1.00), and specific register factor (Contrast 10, R = 1.00). The local factor of academic language completely dominated the specific factor of register (Contrast 8, R = 1.00) and generally dominated the local factor of listening comprehension (Contrast 22, R = .95). The local factor of listening comprehension generally dominated the specific factor of register (Contrast 29, R = 1.00).

Cohort 2 dominance analysis (academic language assessment in oral language context). Similar to Cohort 1, word reading explained the most unique variance compared to the other predictors (i.e., 68% average; Table 5), ranging from 54% to 83%. When comparing Cohorts 1 and 2, it can be observed that when all other predictors were included in the model (i.e., four-predictor model), the amount of variance explained in reading comprehension by word reading was twice as much in Cohort 2 (54%) compared to Cohort 1 (27%). General oral language was the next most important predictor with an overall average of 28% and range of 15–41% of unique variance explained, similar to the magnitudes of variance explained by general oral language for each *k* model in Cohort 1. The local factor of academic language's contribution in Cohort 2 was descriptively different from Cohort 1. Whereas the average unique contribution in Cohort 1 was 11%, in Cohort 2 it was 2%. In the one-predictor model, the local factor of academic language explained 16% variance in Cohort 1 compared to 2% in Cohort 2, and in the 2-predictor model, the local factor of academic language explained 11% in Cohort 1 versus 2% in Cohort 2. The local factor of listening comprehension explained 1% on average in Cohort 2 compared to 6% in Cohort 1, and the specific factor of register explained 1% on average in Cohort 2 compared to 0% in Cohort 1.

The reproducibility for Cohort 2 (see Table 6) showed that word reading completely dominated the local factors of academic language (Contrast 2, R = .04) and listening comprehension (Contrast 10, R = .07), general oral language factor (contrast 4, R = 1.00), and the specific factor of register (Contrast 9, R = 1.00). General oral language completely dominated local factors of listening comprehension (Contrast 4, R = 1.00) and academic language (Contrast 3, R = 1.00), and specific factor of register (Contrast 8, R = 1.00). The local factor of academic language conditionally dominated local factor of listening comprehension (Contrast 11, R = .53) and specific factor of register (Contrast 16, R = 1.00). The specific factor of register conditionally dominated the local factor of listening comprehension (Contrast 17, R = 1.00).

Discussion

Language proficiency is essential for many aspects of modern life, including literacy acquisition. In the present study, we examined two widely studied language constructs, academic language proficiency and listening comprehension skill, by investigating their dimensionality, the relations of identified dimensions to reading comprehension, and the impact of the assessment modality of academic language proficiency on performance level and on the relation to reading comprehension, using data from second-grade students.

Our findings indicate that listening comprehension tasks and academic language proficiency tasks—as measured by the seven tasks in the present study modeled after Uccelli, Barr, et al. (2015; Uccelli, Galloway, et al., 2015)—are best described as a multidimensional construct with a trifactor structure. Children's performances on the listening comprehension and academic language tasks yielded a common underlying factor (or the general language factor) and the academic language and listening comprehension local factors as well as lexical, text, and register specific factors. In other words, the various listening comprehension and academic language tasks captured a common ability, and at the same time, these tasks differentially tapped children's skills in local academic language proficiency and listening comprehension skill as well as specific residual aspects of the lexical, text, and register constructs (Figure 2g). The specific residual factors mostly tapped into residual variance not accounted for by the general and local factors. An exception was the academic register specific factor, which indicates that awareness of academic register is a somewhat unique ability over and above what was common in the other academic language tasks and listening comprehension tasks (the general language construct) and the other specific factors; and it suggests that it captures metacognitive awareness about language use rather than language ability (e.g., Carlisle et al., 1999; Snow et al., 1989).

Having emerged from separate lines of work, listening comprehension skill and academic language proficiency have not been considered together, and their commonality or overlap was not recognized. The present findings indicate that academic language proficiency (at least CALS) is isomorphic with listening comprehension skill empirically. These results are in line with theoretical accounts of listening comprehension skill (Kim, 2016, 2017, 2019) and academic language proficiency (Uccelli, Galloway, et al., 2015). Listening comprehension is a discourse-level skill and as such, draws on lower level or foundational oral language skills such as vocabulary and grammatical/morphosyntactic knowledge as well as higher order cognitive processes necessary to connect and integrate propositions across texts (e.g., reasoning, inferences, comprehension monitoring, and perspective taking; Florit et al., 2014; Kendeou et al., 2008; Kim, 2015, 2016, 2017; Lepola et al., 2012; Strasser & del Rio, 2014; Tompkins et al., 2013). Conceptualization of academic language proficiency also is rooted in discourse theory and includes similar skills (Uccelli, Galloway, et al., 2015).

However, these findings should not be interpreted as indicating that listening comprehension and academic language proficiency are identical constructs. As shown in Figure 1, there is overlap *and* uniqueness between listening comprehension skill and academic language proficiency: the former is a discourse *comprehension* skill in oral language inclusive of academic and informal contexts whereas the latter focuses on language demands for academic purposes without being limited to comprehension and production modalities. Thus, listening comprehension skill and academic language proficiency are different ways of carving up language proficiency studied in separate lines of work, and the unique aspects of the constructs reflect differences in focus.

When it comes to the relation of identified dimensions of oral language skills to reading comprehension, word reading skill and the general oral language skill (not academic language, listening comprehension, or other specific factors) were the two primary contributors to reading comprehension, explaining the vast majority of variance in reading comprehension (see Table 5; also see bivariate correlations between latent variables shown in Appendix B). This was the case across assessment modalities such that in either assessment modality, the local academic language proficiency did not add any unique variance in reading comprehension once all other variables, including word reading and general oral language construct, were accounted for (see Table 5). These results make sense given the overlap in the component skills of listening comprehension and academic language proficiency, and provide further support for the importance of word reading and oral language skills (again not local academic language or listening comprehension factors but common ability) in reading comprehension (Adlof et al., 2006; Catts et al., 2006; Florit & Cain, 2011; Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015; Hoover & Gough, 1990; Kendeou, van den Broek, White, & Lynch, 2009; Kim, 2015, 2017). These results extend previous studies by integrating the literature on the role of listening comprehension and that on academic language.

Another important question, particularly given that our sample was second graders who are rapidly developing reading skills, was the impact of assessment modality on the results. As noted earlier, one cohort of children was administered academic language tasks in the context of independent reading whereas the other cohort was administered these tasks in an oral language context. Our findings revealed that assessment modality influenced performance level in the academic language tasks to a varying degree such that mean differences between cohorts in many tasks were small (e.g., tracking referents task and connecting ideas [connectives]), but mean differences were larger in tasks where the task had greater decoding demands (e.g., academic vocabulary and understanding arguments tasks). For example, the academic vocabulary task, which, by nature, included sophisticated, often multisyllabic words (e.g., *debate, remove),* tended to present greater challenges in decoding ($d = .46$).

Beyond performance level, assessment modality has implications for the nature of the academic language proficiency construct—when assessed in reading contexts, it captures and reflects academic language proficiency *and* decoding skill. The fact that the local academic language factor when assessed in a reading context reflects decoding skill was confirmed in the dominance analysis. The local academic language factor explained approximately 11% of variance in reading comprehension when academic language proficiency was measured in a reading context, whereas it explained approximately only 2% of variance when academic language proficiency was measured in an oral language context (see Table 5). The effect of assessment modality was also shown in the unique amount of variance explained in reading comprehension by word reading: The amount of variance explained by word reading was smaller when academic language proficiency was measured in a reading context (52%; see Table 5) than in an oral language context (68%) because in the former, some of the variance in reading comprehension is explained by the academic language proficiency construct, which taps into decoding. These results offer an important implication for assessment of academic language proficiency. If the goal is to gauge the level of students' academic language proficiency without a confounding role of word reading skill, administration in the oral language condition would be appropriate for precise measurement, particularly for tasks that include greater decoding demands (e.g., academic vocabulary with multisyllables). If the goal is to measure academic language proficiency that includes decoding skill or to measure academic language proficiency for predictive purposes in the context of reading, assessments in a reading context would be appropriate.

Limitations, Future Directions, and Implications

Several limitations and future directions are worth noting. First, the findings are based on the seven tasks for academic language proficiency included in the present study with children in second grade. As noted above, academic language skills have been operationalized in various ways, and in the present study, we adapted from Uccelli and her colleagues' CALS-I for primary grade children. Therefore, future work can extend the current study with other previously used academic language tasks. Second, the present findings are a snapshot for English-speaking children in Grade 2, and thus, generalizability of the present findings is limited to populations similar to students in the present study. Future replications with children at different developmental stages (e.g., Grade 3 or 4) or those from diverse linguistic backgrounds (e.g., those who speak languages other than English, second language learners) are necessary. Another direction for future studies includes an examination of dimensionality in conjunction with assessment modality. For example, dimensionality may differ when students' academic language proficiency is assessed in an oral language or reading mode. As well, with a much larger sample, predictive relations examined in this study may be replicated, and measurement invariance across modalities can be empirically tested to evaluate any systematic differences in loadings, intercepts/means, and variances/residual variances. Future research should also consider the extent to which clustering in classrooms and schools relates to the factor structure estimated in the present study (Pornprasertmanit, Lee, & Preacher, 2014). Finally, it is an open question whether instructional practices influence dimensionality of oral language skills, and future studies can shed light on this question.

Both listening comprehension skill and academic language proficiency constructs recognize and emphasize the importance of knowledge and higher order cognitions beyond vocabulary for successful reading comprehension—vocabulary has received the most attention empirically (e.g., Coyne, McCoach, & Kapp, 2007; Silverman, 2007) and in schools. We believe that listening comprehension skill and academic language proficiency can be useful for different purposes or contexts. Listening comprehension skill is highly useful in explaining and understanding the language and cognitive demands for discourse-level language (including both listening and reading comprehension) while academic language proficiency is useful in highlighting the demands of texts for academic purposes. This difference in utility, however, should not come at the expense of misconceiving these as independent constructs or at the expense of ignoring their clear and large overlap.

Overall the present findings advance the field by adding clarity and coherence about two widely studied language skills, listening comprehension skill and academic language proficiency. These constructs are important for theory and practice, and thus, we believe that our findings add to our growing knowledge and, at the same time, open a number of areas of future research.

References

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

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Erlbaum.

Azen, R. (2013). Using dominance analysis to estimate predictor importance in multiple regression. In Y. Petscher, C. Schatschneider, & D. L. Compton (Eds.), *Applied quantitative analysis in education and social sciences* (pp. 34–64). New York, NY: Routledge.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods, 8,* 129–148. http://dx.doi.org/10.1037/1082-989X.8.2.129

Bailey, A. L., & Butler, F. A. (2002). *An evidentiary framework for operationalizing academic language for broad application to K–12 education: A design document* (CSE Report 611). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.

Bauer, D. J., Howard, A. L., Baldasaro, R. E., Curran, P. J., Hussong, A. M., Chassin, L., & Zucker, R. A. (2013). A trifactor model for integrating ratings across multiple informants. *Psychological Methods, 18,* 475–493.

Beauducel, A., & Hilger, N. (2017). On the bias of factor score determinacy coefficients based on different estimation methods of the exploratory factor model. *Communications in Statistics Simulation and Computation, 46,* 6144–6154. http://dx.doi.org/10.1080/03610918.2016.1197247

Bereiter, C. (2010). *Imagine it!* Columbus, OH: SRA/McGraw-Hill.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics, 20,* 459–478. http://dx.doi.org/10.1017/S0142716499004014

Carrow-Woolfolk, E. (2011). *Oral and written language scales* (2nd ed.). Torrance, CA: Western Psychological Services.

Catts, H. W., Adlof, S. M., & Ellis Weismer, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49,* 278–293. http://dx.doi.org/10.1044/1092-4388(2006/023)

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: Beta is not enough. *Educational and Psychological Measurement, 61,* 229–248. http://dx.doi.org/10.1177/0013164401612006

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

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connective play a unique role of in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36,* 241–260. http://dx.doi.org/10.1111/j.1467-9817.2011.01501.x

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433. http://dx.doi.org/10.3758/BF03214546

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8,* 430–457. http://dx.doi.org/10.1207/S15328007SEM0803_5

Enders, C. K., & Tofighi, D. (2008). The impact of misspecifying class-specific residual variances in growth mixture models. *Structural Equation Modeling, 15,* 75–95. http://dx.doi.org/10.1080/10705510701758281

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23,* 553–576. http://dx.doi.org/10.1007/s10648-011-9175-6

Florit, E., Roch, M., & Levorato, M. C. (2011). Listening text comprehension of explicit and implicit information in preschoolers: The role of verbal and inferential skills. *Discourse Processes, 48,* 119–138. http://dx.doi.org/10.1080/0163853X.2010.494244

Florit, E., Roch, M., & Levorato, M. C. (2014). Listening text comprehension in preschoolers: A longitudinal study on the role of semantic components. *Reading and Writing: An Interdisciplinary Journal, 27,* 793–817. http://dx.doi.org/10.1007/s11145-013-9464-1

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. http://dx.doi.org/10.1037/edu0000026

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practice guide (NCEE 2007–4011)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://files.eric.ed.gov/fulltext/ED497258.pdf

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language*. Austin, TX: PRO-ED.

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

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395. http://dx.doi.org/10.1037/0033-295X.101.3.371

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

Kamata, A., & Bauer, D. J. (2008). A note on the relation between factor analytic and item response theory models. *Structural Equation Modeling, 15,* 136–153. http://dx.doi.org/10.1080/10705510701758406

Kendeou, P., Bohn-Gettler, C. M., White, M. J., & van den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31,* 259–272. http://dx.doi.org/10.1111/j.1467-9817.2008.00370.x

Kendeou, P., van den Broek, P., White, M. J., & Lynch, J. S. (2009). Predicting reading comprehension in early elementary school: The independent contributions of oral language and decoding skills. *Journal of Educational Psychology, 101,* 765–778. http://dx.doi.org/10.1037/a0015956

Kieffer, M. J., & Lesaux, N. K. (2012). Effects of academic language instruction on relational and syntactic aspects of morphological awareness for sixth graders from linguistically diverse backgrounds. *The Elementary School Journal, 112,* 519–545. http://dx.doi.org/10.1086/663299

Kieffer, M. J., Petscher, Y., Proctor, C. P., & Silverman, R. D. (2016). 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. *Scientific Studies of Reading, 20,* 436–454.

Kim, Y.-S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development, 86,* 128–144. http://dx.doi.org/10.1111/cdev.12293

Kim, Y.-S. G. (2016). Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *Journal of Experimental Child Psychology, 141,* 101–120. http://dx.doi.org/10.1016/j.jecp.2015.08.003

Kim, Y.-S. G. (2017). Why the simple view of reading is not simplistic: Unpacking the simple view of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading, 21,* 310–333. http://dx.doi.org/10.1080/10888438.2017.1291643

Kim, Y.-S. G. (2019). Hierarchical and dynamic relations of language and cognitive Skills to reading comprehension: Testing the direct and indirect effects model of reading (DIER). *Journal of Educational Psychology*. Advance online publication. http://dx.doi.org/10.1037/edu0000407

Kim, Y.-S., & Phillips, B. (2014). Cognitive correlates of listening comprehension. *Reading Research Quarterly, 49,* 269–281. http://dx.doi.org/10.1002/rrq.74

Kim, Y.-S. G., & Pilcher, H. (2016). What is listening comprehension and what does it take to improve listening comprehension? In R. Schiff & M. Joshi (Eds.), *Handbook of interventions in learning disabilities* (pp. 159–173). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-31235-4_10

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163–182. http://dx.doi.org/10.1037/0033-295X.95.2.163

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 211–226). Oxford, UK: Blackwell. http://dx.doi.org/10.1002/9780470757642.ch12

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

Lepola, J., Lynch, J., Laakkonen, E., Silvén, M., & Niemi, P. (2012). The role of inference making and other language skills in the development of narrative listening comprehension in 4- to 6-year old children. *Reading Research Quarterly, 47,* 259–282. http://dx.doi.org/10.1002/rrq.020

Leslie, L., & Caldwell, J. S. (2011). *Qualitative Reading Inventory* (5th ed.). Boston, MA: Pearson.

Lonigan, C. J., Burgess, S. R., & Schatschneider, C. (2018). Examining the simple view of reading with elementary school children: Still simple after all these years. *Remedial and Special Education, 39,* 260–273. http://dx.doi.org/10.1177/0741932518764833

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. MacArthur & S. G. J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 115–130). New York, NY: Guilford Press.

McKeown, M. G., Crosson, A. C., Moore, D. W., & Beck, I. L. (2018). Word knowledge and comprehension effects of an academic vocabulary intervention for middle school students. *American Educational Research Journal, 55,* 572–616. http://dx.doi.org/10.3102/0002831217744181

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model for comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 297–384). New York, NY: Elsevier. http://dx.doi.org/10.1016/S0079-7421(09)51009-2

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23,* 412–433. http://dx.doi.org/10.1037/met0000144

Meneses, A., Uccelli, P., Santelices, M. V., Ruiz, M., Acevedo, D., & Figueroa, J. (2018). Academic language as a predictor of reading comprehension in monolingual Spanish-speaking readers: Evidence from Chilean early adolescents. *Reading Research Quarterly, 53,* 223–247. http://dx.doi.org/10.1002/rrq.192

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards*. Washington, DC: Author.

Nimon, K. (2015). Secondary data analyses from published descriptive statistics: Implications for human resource development theory, research, and practice. *Advances in Developing Human Resources, 17,* 26–39. http://dx.doi.org/10.1177/1523422314559805

Nippold, M. A. (1999). Word definition in adolescents as a function of reading proficiency: A research note. *Child Language Teaching and Therapy, 15,* 171–176. http://dx.doi.org/10.1177/026565909901500205

Phillips Galloway, E., & Uccelli, P. (2019). Examining developmental relations between core academic language skills and reading comprehension for English learners and their peers. *Journal of Educational Psychology, 111,* 15–31. http://dx.doi.org/10.1037/edu0000276

Pornprasertmanit, S., Lee, J., & Preacher, K. J. (2014). Ignoring clustering in confirmatory factor analysis: Some consequences for model fit and standardized parameter estimates. *Multivariate Behavioral Research, 49,* 518–543. http://dx.doi.org/10.1080/00273171.2014.933762

Pratt, J. W. (1987). Dividing the indivisible: Using simple symmetry to partition variance explained. In T. Pukkila & S. Puntanen (Eds.), *Proceedings of the Second International Conference in Statistics* (pp. 245–260). Tampere, Finland: University of Tampere.

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92,* 544–559. http://dx.doi.org/10.1080/00223891.2010.496477

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21,* 137–150. http://dx.doi.org/10.1037/met0000045

Scarcella, R. (2003). *Academic English: A conceptual framework* (Technical Report 2003–1). Berkeley: The University of California Linguistic Minority Research Institute.

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education, 12,* 431–459.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. New York, NY: Routledge. http://dx.doi.org/10.4324/9781410610317

Silverman, R. (2007). Vocabulary development of English language and English-only learners in kindergarten. *The Elementary School Journal, 107,* 365–383. http://dx.doi.org/10.1086/516669

Snow, C. E. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53,* 165–189. http://dx.doi.org/10.17763/haer.53.2.t6177w39817w2861

Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language, 17,* 697–710. http://dx.doi.org/10.1017/S0305000900010953

Snow, C. E. (1991). The theoretical basis for relationships between language and literacy development. *Journal of Research in Childhood Education, 6,* 5–10. http://dx.doi.org/10.1080/02568549109594817

Snow, C., Cancini, H., Gonzalezm, P., & Shriberg, E. (1989). Giving formal definitions: An oral language correlate of school literacy. In D. Bloome (Ed.), *Classrooms and literacy* (pp. 233–249). Norwood, NJ: Ablex.

Snow, C. E., Cancino, H., De Temple, J., & Schley, S. (1991). Giving formal definitions: A linguistic or metalinguistic skill? In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 90–112). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511620652.007

Snow, C. E., Tabors, P. O., & Dickinson, D. K. (2001). Language development in the preschool years. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 93–110). Baltimore, MD: H. Brookes Publishing Company.

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511609664.008

Solomon, J., & Rhodes, N. (1995). *Conceptualizing academic language* (Research Rep. No. 15). Santa Cruz: National Center for Research on Cultural Diversity and Second Language Learning, University of California.

Strasser, K., & del Rio, F. (2014). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly, 49,* 169–187. http://dx.doi.org/10.1002/rrq.68

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. http://dx.doi.org/10.1044/1092-4388(2006/086)

Tompkins, V., Guo, Y., & Justice, L. M. (2013). Inference generation, story comprehension, and language in the preschool years. *Reading and Writing: An Interdisciplinary Journal, 26,* 403–429. http://dx.doi.org/10.1007/s11145-012-9374-7

Tunmer, W. E., & Chapman, J. W. (2012). The simple view of reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities, 45,* 453–466. http://dx.doi.org/10.1177/0022219411432685

Uccelli, P., Barr, C. D., Dobbs, C. L., Galloway, E. P., Meneses, A., & Sanchez, E. (2015). Core academic language skills: An expanded operational construct and a novel instrument to chart school-relevant language proficiency in preadolescent and adolescent learners. *Applied Psycholinguistics, 36,* 1077–1109. http://dx.doi.org/10.1017/S014271641400006X

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly, 50,* 337–356. http://dx.doi.org/10.1002/rrq.104

Utah State Board of Education. (2017). *Fourth Grade Science, 2017–2018*. Retrieved from https://eq.uen.org/emedia/file/6482d388-3b56-4186-8d92-a9a47e56ea25/1/Grade4RS.pdf

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2012). *Test of word reading efficiency* (2nd ed.). Austin, TX: Pro-ED.

Wechsler, D. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio, TX: Pearson.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests of Achievement*. Itasca, IL: Riverside.

Woods, M. L., & Moe, A. J. (2011). *Analytical Reading Inventory: Comprehensive standards-based assessment for all students including gifted and remedial* (9th ed.). Upper Saddle River, NJ: Pearson Education.

(*Appendices follow*)

Appendix A

Academic Language Tasks and Examples

Task (total number of items)	Example item
Academic vocabulary (14)	We had a **<u>debate</u>** about pizza. a. discussion b. question c. exam d. conclusion
Academic register (10)	"An umbrella is what you use so you do not get wet in the rain." Let's think about this one. Do you think it is written for children or for adults? (child version) "An umbrella is an object that protects you from the rain." Let's think about this one. Do you think it is written for children or for adults? (adult version) The child was asked to circle either 'children' or 'adults,' which were presented with illustrations of children and adults.
Organizing texts (6)	Title: How is paper made? Paper is made out of trees. First, trees are cut down into branches. Second, these branches are put in trucks. _______
Tracking referents (7)	Hot cocoa is a delicious treat on a cold day. It warms you up almost as much as a hat and mittens! In the sentence above, It refers to: a. cold day b. hot cocoa c. treat
Connecting ideas (10)	John hurt his knee ___ he was playing basketball. a. so b. while c. unless d. if
Understanding viewpoints: modal verbs (16)	"Two friends are trying to figure out if their teacher is going to give them a quiz. One friend is very sure there will be a quiz, and one friend is not sure." A boy says, "There will be a quiz tomorrow." A girl says, "There might be a quiz tomorrow." How sure is the girl about a quiz tomorrow? The child was asked how sure or unsure each person is about the situation by selecting Yes, Maybe yes, Maybe no, No.
Understanding arguments (11)	The newspaper said, "Students need recess to play and relax at school." Harold says, "Yes. At recess we play games like tag, soccer, or basketball." In the sentence above, Harold is giving . . . a. excuses b. examples c. definitions d. exaggerations

(*Appendix continues*)

Appendix B

Latent Correlation Matrix for Cohorts 1 and 2

	Cohort 1						Cohort 2					
Construct	1	2	3	4	5	6	1	2	3	4	5	6
1. Local[a] factor academic lang	1.00						1.00					
2. Local factor listening comp	.00	1.00					.00	1.00				
3. General oral lang	.00	.00	1.00				.00	.00	1.00			
4. Word reading	.59	.04	.36	1.00			.13	−.17	.29	1.00		
5. Reading comp	.46	.25	.62	.88	1.00		.17	−.07	.64	.91	1.00	
6. Specific factor register	.00	.00	.00	−.08	−.07	1.00	.00	.00	.00	.07	.12	1.00
M	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
SD	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
n	165	165	165	165	165	165	185	185	185	185	185	185

Note. General oral language refers to common shared variance across all the oral language tasks; Lang = language; Comp = comprehension.
[a] The term *local* is used for the academic language proficiency and listening comprehension skill latent variables in trifactor models, and specific register latent variable indicates a residualized construct, all of which are after taking out the variance in items due to the general oral language factor. The general oral language construct is deemed to be the most reliable portion of the total item variance as the construct is indicated by all items. The local academic language and listening comprehension factors exist after taking out the variance in items due to the general oral language factor. The register construct represents unique item variance in the construct after removing the variance shared with the other constructs. These features are important as one may be puzzled by seemingly low correlations, such as the .25 and −.07 correlations between listening and reading comprehension in Cohorts 1 and 2.

Received December 1, 2018
Revision received September 20, 2019
Accepted September 30, 2019 ■