DOI: 10.1111/cdev.13935

CHILD DEVELOPMENT

EMPIRICAL ARTICLE

Oral discourse skills: Dimensionality of comprehension and retell of narrative and expository texts, and the relations of language and cognitive skills to identified dimensions

Young-Suk Grace Kim

University of California, Irvine, Irvine, California, USA

Correspondence
Young-Suk Grace Kim, University of California at Irvine, Irvine, CA, USA.
Email: youngsk7@uci.edu

Funding information
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Number: P50HD052120; Institute of Education Sciences, Grant/Award Number: R305A180055, R305A130131 and R305A200312

Abstract

We examined the dimensionality of oral discourse skills (comprehension and retell of texts) and the relations of language and cognitive skills to the identified dimensions. Data were from 529 English-speaking second graders (M_{age} = 7.42; 46% female; 52.6% Whites, 33.8% African Americans, 4.9% Hispanics, 4.7% two or more races, .8% Asian Americans, .6% American Indians, .2% Native Hawaiians, 2.5% unknown; data from 2014–2015 to 2016–2017). Confirmatory factor analysis revealed that oral discourse skills are best described as four related but dissociable dimensions of narrative comprehension, narrative retell, expository comprehension, and expository retell (*r*s = .59–.84). Language and cognitive skills had different patterns of relations to the identified dimensions and explained larger amounts of variance in comprehension than in retell.

Oral discourse skills—the skills to engage in multi-utterance conversations; understand, retell, and produce stories and informational texts; and describe experiences and information in oral language contexts—are an integral part of daily life, academic learning, and workplace functioning. Discourse skills have received the lion's share of attention, particularly in the context of written texts (i.e., reading comprehension), and have been extensively studied in children's language production (e.g., utterances in parent–child interactions). Despite their importance, however, comprehension and retell of oral texts rarely have been examined together. In particular, one question that remains open is the dimensionality of oral discourse skills—whether comprehension and retell of oral texts in various genres reflect a unidimensional construct or multidimensional construct—and the relations of language and cognitive skills to the identified dimensions.

In the present study, we investigated the dimensionality of comprehension and retell of narrative and expository texts, and the relations of cognitive skills (working memory, attentional control, inference, theory of mind, comprehension monitoring) and language skills (vocabulary and grammatical knowledge) to the identified dimensions of oral discourse skills, using data from English-speaking children in Grade 2. Theoretically, comprehension, and retell of texts across genres are hypothesized to draw on essentially the same processes (e.g., Kintsch, 1988) and skills (e.g., Kim, 2016). However, the relative demands on the processes and skills might differ as a function of receptive and expressive modalities (comprehension and retell) and genres (narrative and expository). Identifying dimensionality and predictors of the dimensions can shed light on the processes and extent to which language and cognitive skills are similarly and differentially tapped for comprehension versus retell and for narrative versus expository genres. The results can have practical implications for assessment and teaching. For example, if a unidimensional structure describes

Abbreviations: CASL, Comprehensive Assessment of Spoken Language; CFI, comparative fit index; DIET, direct and indirect effects model of text comprehension and production; *n*BIC, sample size adjusted Bayesian Information Criterion; QRI-5, Qualitative Reading Inventory-5; RMSEA, root mean square error of approximation; SALT, Systematic Analysis of Language Transcription; SRMR, standardized root mean square residual; SWAN, Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale; TNL, Test of Narrative Language; WJ, Woodcock-Johnson.

comprehension and retell of narrative and expository texts, instruction of any single aspect (e.g., narrative comprehension) might support the development of the other aspects (e.g., narrative retell, expository comprehension, and expository retell). If a multidimensional structure is supported, each of the identified dimensions may need to be assessed and taught, respectively (see the implications section below).

Note that we draw on the literature on *oral* discourse to the extent evidence is available. For example, the literature on language and cognitive predictors of oral discourse comprehension (listening comprehension) is growing, and we draw on this literature base. We also draw on prior work on *written* discourse (e.g., reading comprehension), particularly about the effects of genre and modality because the vast majority of previous studies on these topics have been conducted in the context of reading—in fact, unless otherwise noted, studies on modalities and genres are from the reading comprehension literature. We acknowledge that reading comprehension involves processes and skills associated with word reading in addition to comprehension processes. However, theoretically, processes and skills for listening and reading comprehension are essentially the same with the exception of word reading (Kim, 2020; Gough & Tunmer, 1986), and evidence indeed supports this (for reading comprehension, see e.g., Ahmed et al., 2021; Cain & Oakhill, 2007; for listening comprehension, see e.g., Florit et al., 2011; Kendeou et al., 2008; Kim, 2016; Kim & Phillips, 2014; Lepola et al., 2012).

The dimensionality of oral language skills

Oral language skills include a range of grain sizes and aspects, such as phonology, morphology, vocabulary, syntax, sentence, pragmatics, and discourse, and receptive and expressive aspects. By now a sizable body of literature exists on the dimensionality of oral language skills of various grain sizes and aspects (see e.g., Anthony et al., 2014; Language and Reading Research Consortium, 2015, 2017; Lonigan & Milburn, 2017; Mouzaki et al., 2020; Tomblin & Zhang, 2006). For example, receptive and expressive vocabulary and grammatical knowledge were found to be two related but dissociable dimensions for children in Grades 2 and above, whereas a single dimension was supported for children in kindergarten (e.g., Tomblin & Zhang, 2006). In a study with preschoolers, vocabulary and grammatical knowledge formed a single dimension while speech articulation and perception were related but distinct dimensions (Anthony et al., 2014). When oral discourse skill, listening comprehension, was examined together with vocabulary and grammatical knowledge, listening comprehension was a related but dissociable dimension from the lexical and sentence dimension for children in preschool to Grade 3 (e.g., Language and Reading Research Consortium, 2017).

When it comes to the dimensionality of discourse skills, literature is limited and extant studies show different patterns. For example, Muijselaar et al. (2017) examined the dimensionality of reading comprehension of a narrative text and an expository text for German-speaking children in Grade 4 and found that reading comprehension was best described as a common factor across genres. When it comes to the dimensionality of *oral discourse* skills, the focus of the present study, to our knowledge, there are only two studies. Gillam and Pearson (2004) examined narrative comprehension, retell, and production, using data from English-speaking 5- to 11-year-olds, and found that narrative comprehension was strongly related to but dissociable from the retell and production construct ($r = .85$). Kim et al. (2015) examined narrative comprehension, narrative retell, and narrative production in Korean-speaking students in Grade 1. Their data supported a bifactor structure that was composed of a discourse language general factor that captured common variance across all the tasks and two specific factors, the narrative comprehension-specific factor that reflected the comprehension tasks and the oral narration-specific factor that reflected the retell and production tasks.

Oral discourse skills

According to the construction-integration model, discourse skills require constructing an integrated mental representation of texts called the situation model (Kintsch, 1988). Discourse skills and associated information processing encompass different modalities such as a receptive skill (comprehension) and a productive skill (recall and production) and different types of texts such as narrative and informational texts. The central process for discourse comprehension and production is constructing a rich and accurate situation model—one's representation of the situation of the text. The situation model is built on lower-order mental representations: the textbase representation, the representation of elementary and literal propositions, which, in turn, is built on the surface code, the representation of linguistic input of the text. The textbase representation tends to lack coherence and include inconsistencies; therefore, integration processes are needed to establish a globally coherent structure (Kintsch, 1988). Once a well-structured situation model is established, one can successfully answer comprehension questions and retell and produce coherent texts across genres (McNamara et al., 1996). Therefore, performance on comprehension and retell tasks of various genres should be similar inasmuch as they similarly tap into one's mental representation.

The above-noted construction and integration processes draw on language and cognitive skills and knowledge. According to the direct and indirect effects model of text comprehension and production (DIET; Kim, 2016), constructing the surface code and textbase representation draws on domain-general cognitive skills or executive function (e.g., working memory, inhibitory and attentional control, shifting), and language skills such as vocabulary, morphosyntactic, and syntactic knowledge as well as some degree of world/topic knowledge and inference. Constructing the situation model, which involves integrating propositions for global coherence, further relies on higher-order cognitions and regulation such as reasoning, inference, perspective taking, and comprehension monitoring (see Kim, 2016, for details and a review of empirical evidence on language and cognitive predictors of oral discourse comprehension). The language and cognitive skills and knowledge that contribute to oral discourse skills are hierarchically related such that domain-general cognitions support foundational language skills (vocabulary and grammatical/syntactic knowledge), which support higher-order cognitions and regulation, which, in turn, support discourse oral language skills. According to DIET, discourse skills are built on foundational oral language skills such as vocabulary and syntactic knowledge, but they are dissociable higher-order constructs from foundational oral language skills as they require further skills and knowledge such as higher-order cognitions and background knowledge such as discourse knowledge (Kim, 2016).

It should be noted that having essentially the same underlying processes may not entail symmetry of recruitment of processes across modalities such as comprehension and retell (Kim, 2020; McNamara et al., 1996). Comprehension is generally measured by questions that probe one's understanding of explicitly stated information in a text (i.e., literal comprehension) and implied information that is not explicitly stated (i.e., inferential comprehension). Different types of questions are designed to tap different degrees of comprehension, such as shallow comprehension of text recall and higher-order, deep comprehension of interpretation and evaluation of texts that require reorganization of one's understanding of the text as a whole. In retell, children are asked to reconstruct the information typically without specific prompts or probes (i.e., free retell or recall; see Barnes et al., 2014; Bellinger & Diperna, 2011; Shapiro et al., 2014; Spencer et al., 2019). Some argued that retell taps and provides insight into children's deep comprehension because retell captures students' interpretation of the text as well as recall (e.g., Kida et al., 2016; Loyd & Steele, 1986). On the other hand, retell has been argued to primarily capture recall of the content of the text (textbase representation) rather than deep comprehension (McNamara et al., 1996). The presence and absence of probing questions in comprehension versus retell may influence the extent to which language and cognitive skills are tapped because it has been shown, in the context of reading (reading comprehension), that unless probed or challenged with questions that encourage deep comprehension, comprehenders tend to settle for shallow comprehension (Graesser et al., 2005). If comprehension and retell tap language and cognitive skills differentially, then they are likely to be related but distinct constructs. In line with this speculation, a recent meta-analysis found that retell was moderately, not strongly, related to other measures of reading comprehension (Cao & Kim, 2021). Studies on differential relations of language and cognitive skills for comprehension versus retell are highly limited. An exception is Spencer et al.'s (2019) study in the context of written texts (i.e., reading comprehension) for English-speaking adolescents. Their results showed that working memory was independently related to retell, but not reading comprehension, whereas vocabulary and cohesive inferencing were uniquely related to reading comprehension, but not retell.

Asymmetry of recruitment of processes and skills also applies to the genre (Kim, 2016). Narrative texts and informational texts have different goals, text structures, language demands, and world and content knowledge demands (Best et al., 2008). Narrative texts generally tell a story evolving around characters and associated events, actions, and conflicts. In contrast, informational texts focus on factual information on a topic (Duke, 2000), and include multiple subgenres. Exposition is a subgenre of informational texts and focuses on the description of ideas and their logical interrelations (Berman & Nir-Sagiv, 2007). According to DIET, the extent to which language and cognitive skills contribute to discourse skills differs depending on text features and genres (also see Kintsch, 1988). For example, texts that focus on conveying information such as expository texts will place greater demands on world/content/topic knowledge (e.g., Best et al., 2008). Narrative texts typically evolve around different characters' motivations, viewpoints, and associated behaviors, and therefore, perspective taking—the ability to understand multiple perspectives—likely plays a greater role in successful comprehension of narrative texts than expository texts (Dore et al., 2018; Kim, 2016). Indeed, a recent study with children in Grade 4 found that perspective taking measured by theory of mind—the ability to understand one's own and others' beliefs, desires, and perspectives—was more strongly related to comprehension of narrative texts than expository texts in oral language context (Kim et al., 2021). In that study, theory of mind predicted the extent of mental state talk (e.g., *believe* and *decide*) in narrative texts and informational texts, which, in turn, was more strongly related to narrative comprehension than expository comprehension (Kim et al., 2021).

Studies that were conducted using written texts (reading comprehension) also suggest differential contributions of skills to narrative versus expository texts. World/topic knowledge was found to be particularly

important for comprehension of expository texts (Best et al., 2008; Kendeou & van den Broek, 2007), vocabulary was more important for comprehension of narrative texts (Eason et al., 2012), and inference skill was more important for comprehension of narrative and expository texts than functional texts (Eason et al., 2012). Furthermore, executive functions such as working memory and planning and organizing were related to reading comprehension and retell of expository texts but not narrative texts after accounting for vocabulary, phonological awareness, word reading, and socioeconomic status for children in elementary grades (Wu et al., 2020; also see Eason et al., 2012). It was speculated that the same words tend to be more frequently used or repeated to link ideas in expository texts than in narrative texts, and keeping track of the same words and their relations would place greater demands on memory and attentional control processes to establish deep comprehension (Wu et al., 2020).

Present study

Previous studies suggested that oral discourse skills such as listening comprehension are dissociable from other oral language skills (e.g., vocabulary and syntactic knowledge), and receptive and expressive modalities and genres might impact discourse processes and recruit language and cognitive skills differentially. To our knowledge, there is no previous study that examined the dimensionality of *oral discourse* skills across comprehension and retell modalities and narrative and expository genres, and the potential differential relations of language and cognitive skills to the identified dimensions. Identifying the dimensionality of oral discourse skills and the language and cognitive predictors of the dimensions can inform similarities and differences in the discourse process. If oral discourse skills do not reflect a single ability, it is reasonable to speculate that different dimensions recruit processes to a different extent, and language and cognitive skills make differential contributions to the identified dimensions of discourse skills (Kim, 2020). The following questions guided the present study:

1. Are children's skills in comprehending and retelling narrative texts and expository texts in oral language contexts best described as a unidimensional construct or multidimensional construct for English-speaking children in Grade 2?
2. How are language and cognitive skills (vocabulary, grammatical knowledge, working memory, attentional control, inference, theory of mind, and comprehension monitoring) related to the identified dimension(s)? If multidimensionality is supported, are the language and cognitive skills differentially related to the identified dimensions?

We hypothesized that a multidimensional structure would describe the data better. However, we did not have a clear hypothesis about which multidimensional structure would best characterize the data. Alternative multidimensional structures were systematically examined by considering modality (comprehension & retell) and genre (narrative & expository; see the Data Analytic Strategies section for details), informed by DIET and previous studies (see above). We also posited differential relations of language and cognitive skills to the identified dimensions such that theory of mind (a measure of perspective taking) and vocabulary would be more strongly related to narrative texts than to expository texts (Dore et al., 2018; Eason et al., 2012; Kim, 2016; Kim et al., 2021), and working memory and attentional control would be more strongly related to expository texts than to narrative texts (Eason et al., 2012; Wu et al., 2020). It was also posited that vocabulary would be more strongly related to comprehension than to retell, and working memory would be more strongly related to retell than to comprehension (Spencer et al., 2019). The study was exploratory in nature as it explored dimensionality and predictors of oral discourse skills.

METHOD

Participants

Data were from 529 English-speaking children in Grade 2 (46% female; mean age = 7.42, SD = .58) from nine schools in the southeastern part of the US. The sample was drawn from three cohorts who were assessed in three consecutive academic years (n = 179 in Cohort 1 in 2014–2015; n = 165 in Cohort 2 in 2015–2016; n = 185 in Cohort 3 in 2016–2017). These children were assessed using identical measures administered in the same sequence at the same time during the academic year. Data from Cohorts 2 and 3 were reported in a study that focused on predictors of reading comprehension (Kim, 2017), and some of the data from the three cohorts were reported in a study that examined text factors that influence comprehension (Kim & Petscher, 2021). These previous studies included working memory, vocabulary, grammatical knowledge, inference-making, theory of mind, and comprehension monitoring, which are used in the present study. Children's racial and ethnic backgrounds were as follows: 53% White children, 34% African American children, 5% Hispanic children, 5% two or more races, and less than 1% Asian American children and American Indian children, respectively. According to the district record, approximately 1% of the children were classified as limited English proficiency, 13% received speech services, 1% received services related to language impairment, .6% had learning disability, and 2% were identified as gifted children. All children were included in the data analysis.

Measures

Children were administered the following language and cognitive tasks. Reliability estimates are from the sample. Unless otherwise noted, all the items were scored dichotomously (1 = correct, 0 = incorrect).

Narrative comprehension and retell

Children's comprehension and retell of narrative texts were measured by the Test of Narrative Language (TNL; Gillam & Pearson, 2004), which was slightly adapted to systematically measure comprehension and retell using identical stories. Whereas TNL measures comprehension of stories in Tasks 1, 3, and 5, it measures story retell only for the story in Task 1. In our modified protocols, children were asked to retell the stories in Tasks 3 and 5, in addition to the story in Task 1. Children were asked to retell the story immediately after hearing it before being presented with comprehension questions. The story was read to the child one time, following the TNL protocol. Characteristics of the three stories in TNL are found in Online Supplemental Materials, including the number of words, number of sentences, text difficulty as measured by Lexile, number of literal and inferential comprehension questions, and narrativity (the extent of narrative features of a text such as characters and their interactions, and events) as measured by Coh-Metrix (Graesser et al., 2004). As expected, narrativity was consistently higher in the TNL texts (85% to 95%) than the expository texts (24% to 54%). There was a total of 30 comprehension questions: 11 items for Task 1, 9 items for Task 3, and 10 items for Task 5. The majority of items, 22 items, were scored 0 (incorrect) and 1 (correct); six items (two items in Task 1, two items in Task 3 and two items in Task 5) were scored on 0, 1, or 2 scale; and two items (one item in Task 1 and one item in Task 5) were scored on a 0, 1, 2, or 3 scale. Therefore, total possible maximum was 15 for Task 1, 11 for Task 3, and 14 for Task 5, and the total possible maximum across the tasks was 40. Cronbach's α was .76.

Children's retell of TNL stories was recorded by a digital recorder and was transcribed verbatim following the Systematic Analysis of Language Transcription (SALT; Miller & Iglesias, 2006) guidelines. All transcribers were trained until they met the minimum transcription reliability (exact agreement) of 95% or less than 5% of discrepancy with the master transcribers, who were a PhD student in education and a research staff member with a master's degree in education. Both master transcribers had extensive prior experiences with transcription in a previous project on children's language and literacy development. In addition, a minimum of 20% of the transcriptions were randomly checked by the master transcribers.

Retell quality was coded following previous work (see Gillam & Pearson, 2004, for a review of scoring retell) in terms of the following eight aspects: the extent to which retell included main characters, setting, main events, problem, resolution, introduction, conclusion, and logical sequencing of the story. Children received scores for partially correct responses. Most of the structural elements were rated on a scale of 0 (absence of relevant information) to 3 (precise information). For example, there were three main events in the MacDonald's story (Task 1). If the child's retell included all three events, then their score was 3 whereas retell of one or two events was assigned a score of 1 or 2, respectively. The resolution element was scored 0 to 2. Presence of introduction and conclusion was dichotomously scored (e.g., 1 = introduction was present [e.g., This story is about…]; 0 = introduction was absent). Logical sequencing of the story was also dichotomously scored (1 = order of mainline events was logical; 0 = order of mainline events was not logical).

Two coders, one graduate student in education and one undergraduate student majoring in speech and language pathology, were rigorously trained in coding retell quality. The training was composed of a series of meetings and practice sessions. In an initial meeting, the rubric was discussed, and coders practiced coding samples together. This was followed by several practice sessions where coders brought their scores for a practice set and discussed discrepancies and clarified points. After several iterations, they tried a reliability set, followed by a discussion of discrepancies; This continued until the minimum exact agreement rate of 90% was reached. Reliability was estimated by calculating percent of exact agreement between two coders across the multiple aspects of the narrative tasks (e.g., characters, setting, mainline events, and logical sequencing). Percent agreement ranged from .91 to .97, using 40 sample retells. All the retells were double scored, and final scores were determined after discussion of discrepant scores.

Expository comprehension and retell

Due to the absence of a normed listening comprehension measure of expository texts, we used texts from the Qualitative Reading Inventory-5 (QRI-5; Leslie & Caldwell, 2011). Three texts for Grade 2 (Leslie & Caldwell, 2011) were *Changing Matter*, *Whales and Fish*, and *Where Do People Live*? Characteristics of the three expository texts are presented in Online Supplemental Materials. The child heard each text once and was asked to retell the text. Then, the child was asked eight open-ended comprehension questions per text, and each was scored dichotomously for a total possible maximum score of 24 (8 × 3). Cronbach's α was .74.

Children's retell was transcribed verbatim following the SALT guidelines (Miller & Iglesias, 2006) in an identical manner as for narrative retell. Overall quality of retell was evaluated in terms of the extent to which main ideas and associated details were included,

following previous work (e.g., Wagner et al., 2011). Main ideas were scored for accuracy on a scale of 0 (inaccurate), 1 (partially accurate), and 2 (accurate), and each key detail was given a point. For example, for the main idea of the text, *Where Do People Live?*, the child who stated that people live in different places received a score of 1; and the child who stated that people live in different places depending on what they like the most (or what they value or what is important to them) received a score of 2. The number of key details varied depending on texts, and key details per text were a priori identified. Note that variation in the number of key details across expository texts does not present a problem in the data analysis because the research question is about covariation of children's performance on comprehension and retell across texts, not absolute level of performance depending on texts. Two coders, one PhD student in education and a research staff member with a master's degree in education were rigorously trained in coding retell quality, following the identical procedures as narrative retell. Exact agreement using 40 samples ranged from .90 to .95 (that is, both coders agreed 90% to 95% on coded main ideas and key details). All the retells were double scored, and final scores were determined after discussion of discrepant scores.

Inference

Knowledge-based inference (the ability to infer information based on background knowledge) was measured by the Inference task of the Comprehensive Assessment of Spoken Language (CASL; Carrow-Woolfolk, 1999). In this task, the child heard one- to three-sentence scenarios and was asked a question that required inference drawing on prior knowledge (e.g., "*Before Jane went outside, she put on a thick jacket. What was the weather like?*"). Cronbach's α was .90.

Theory of mind

Theory of mind was measured by false-belief tasks. First-order false-belief tasks involve an understanding of a character's thoughts or feelings, and second-order false-belief tasks involve an understanding of a character's thoughts or feelings regarding another character's beliefs, thoughts, or feelings (Mahy et al., 2017). Evidence indicates that first-order theory of mind typically develops between the ages of four and five, and second-order theory of mind around age 7 (e.g., Perner & Wimmer, 1985). Therefore, considering developmental stage of the participating children (i.e., $M_{age} = 7.42$) we used second-order false belief tasks using three scenarios in the context of a bake sale, a visit to a farm, and going out for a birthday party (Kim et al., 2021). The scenarios were presented orally accompanied by a series of corresponding illustrations and children were asked a total of 18 target questions with six questions per scenario (3 × 6). There were memory questions in relevant places, and if the child' answers to memory questions were incorrect, correct answers were provided. Corrective feedback was not provided for target questions. Children's responses to target questions, but not memory probes, were scored. Cronbach's α was .82.

Comprehension monitoring

Following previous work (e.g., Baker, 1984), an inconsistency detection task was used to measure comprehension monitoring (see Kim & Phillips, 2014). The child heard two- to four-sentence scenario and was asked whether the story made sense or not. "Not making sense" was explained as sentences not going together in two practice items. If the child indicated that the story did not make sense, they were asked to provide a brief explanation and to fix the story so that it made sense. Nine experimental items included three consistent stories and six inconsistent stories, and they were randomly ordered. For the six inconsistent stories, one point was given to correct explanation and repair of the story, respectively. Therefore, for an inconsistent story, the total maximum possible score for the item was 3—one point for correctly identifying inconsistency, one point for providing a correct explanation, and one point for an accurate repair. Cronbach's α was .69.

Vocabulary

The Picture Vocabulary subtest of the Woodcock-Johnson III (WJ; Woodcock et al., 2001) was used. Children were shown pictures and asked to name them. Cronbach's α was .69.

Grammatical knowledge

The Grammaticality Judgement task of CASL (Carrow-Woolfolk, 1999) was used. The child heard a sentence and was asked whether the sentence was grammatically correct. If the child stated that the sentence was grammatically incorrect, they were asked to correct the sentence. Cronbach's α was .94.

Working memory

A listening span task (Daneman & Merikle, 1996) was used. There were four practice items and 13 test items. The child heard a three- to four-word short sentence involving common knowledge familiar to children (e.g., Birds can

fly; Apples are blue) and was asked whether the heard sentence was correct or not (yes/no question). After hearing multiple sentences (i.e., two to four), the child was asked to identify the last word of each sentence. Children's yes/no responses regarding the veracity of the statements were not scored, but their responses on the last words in correct order were given a score of 0 to 2: 2 points were given for correct last words in correct order; 1 point was given for correct last words in incorrect order; and 0 point was given for incorrect last words. Cronbach's α was .73.

Attentional control

Participating children's teachers completed a behavioral checklist, the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale (SWAN; Swanson et al., 2006; see Arnett et al., 2013, for validity evidence). There were 30 items that are rated on a 7-point scale ranging from 1 (*far below average*) to 7 (*far above average*). Cronbach's α was .99.

Procedures

Children were individually assessed in a quiet space in the school. Assessments were administered in several sessions with each session lasting approximately 30 to 40 min. Assessment order was identical across the three cohorts and was as follows: working memory, WJ Picture Vocabulary, TNL, CASL Grammaticality Judgement, CASL Inference, expository text retell and comprehension, theory of mind, and comprehension monitoring.

Data analytic strategies

Primary data analytic strategies were confirmatory factor analysis and structural equation modeling. Confirmatory factor analyses were conducted to address the first research question on the dimensionality of comprehension and retell of narrative and expository texts. After examining distributional properties, full information maximum likelihood was used for estimation using Mplus 8.4 (Muthén & Muthén, 1998–2017). A total of eight alternative models shown in Figure 1 that systematically considered modality and genre were fitted to the data. Model 1 (Figure 1.1) is a unidimensional model where all the indicators across genres (narrative and expository) and modality (comprehension and retell) were hypothesized to reflect a single factor. Model 2 (Figure 1.2) is two-factor model by genre where comprehension and retell of narrative texts form a factor, and comprehension and retell of expository texts form another related factor. Model 3 (Figure 1.3) is a two-factor model by modality where comprehension of narrative and expository texts forms a factor, and retell of narrative and expository texts forms another related factor. Model 4 (Figure 1.4) is a four-factor model, composed of narrative comprehension, narrative retell, expository comprehension, and expository retell. Model 5 (Figure 1.5) has a second-order factor by genre, narrative and expository, along with the four first-order factors (i.e., narrative comprehension, narrative retell, expository comprehension, and expository retell). Model 6 (Figure 1.6) has a second-order factor by modality, comprehension, and retell, along with the four first-order factors (i.e., narrative comprehension, narrative retell, expository comprehension, and expository retell). Model 7 (Figure 1.7) is a bifactor model with a discourse language general factor and specific factors by genre (narrative and expository). Lastly, Model 8 (Figure 1.8) is a bifactor model with a discourse language general factor and specific factors by modality (comprehension and retell).

The second research question is about the relations of language and cognitive skills to the identified dimensions of oral discourse skills. This question was addressed by fitting a structural equation model shown in Figure 3. In this model, all the language and cognitive skills were allowed to predict the identified dimensions of oral discourse skills. Aligned with DIET, lower skills predicted higher skills (e.g., working memory and attentional control predicted vocabulary, grammatical knowledge, inference, theory of mind, and comprehension monitoring). Error variances were allowed to covary between the same tasks within genre because they involved the same texts (i.e., answering comprehension questions and retelling after hearing the same text). To examine potential differential contributions of cognitive skills, total effects (direct effects+indirect effects; regression weights) of language and cognitive predictors were estimated.

Model fit was evaluated using chi-square, comparative fit index (CFI), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and sample size adjusted Bayesian information criterion (*n*BIC). Some of the eight competing models for the first research question were nested whereas others were not. For example, the two-factor models (Figure 1.2,1.3) were nested within the single-factor model (Figure 1.1) whereas the second-order models (Figure 1.5,1.6) were not nested. To compare model fit, the chi-square difference test was conducted for nested models, and *n*BIC values were compared for non-nested models following Raftery (1995) as the guidelines for the magnitude of differences.

RESULTS

Descriptive statistics and preliminary analysis

Missingness was minimal, ranging from 0% in the working memory task to 3% in the attentional control task. Little's missing completely at random test revealed that

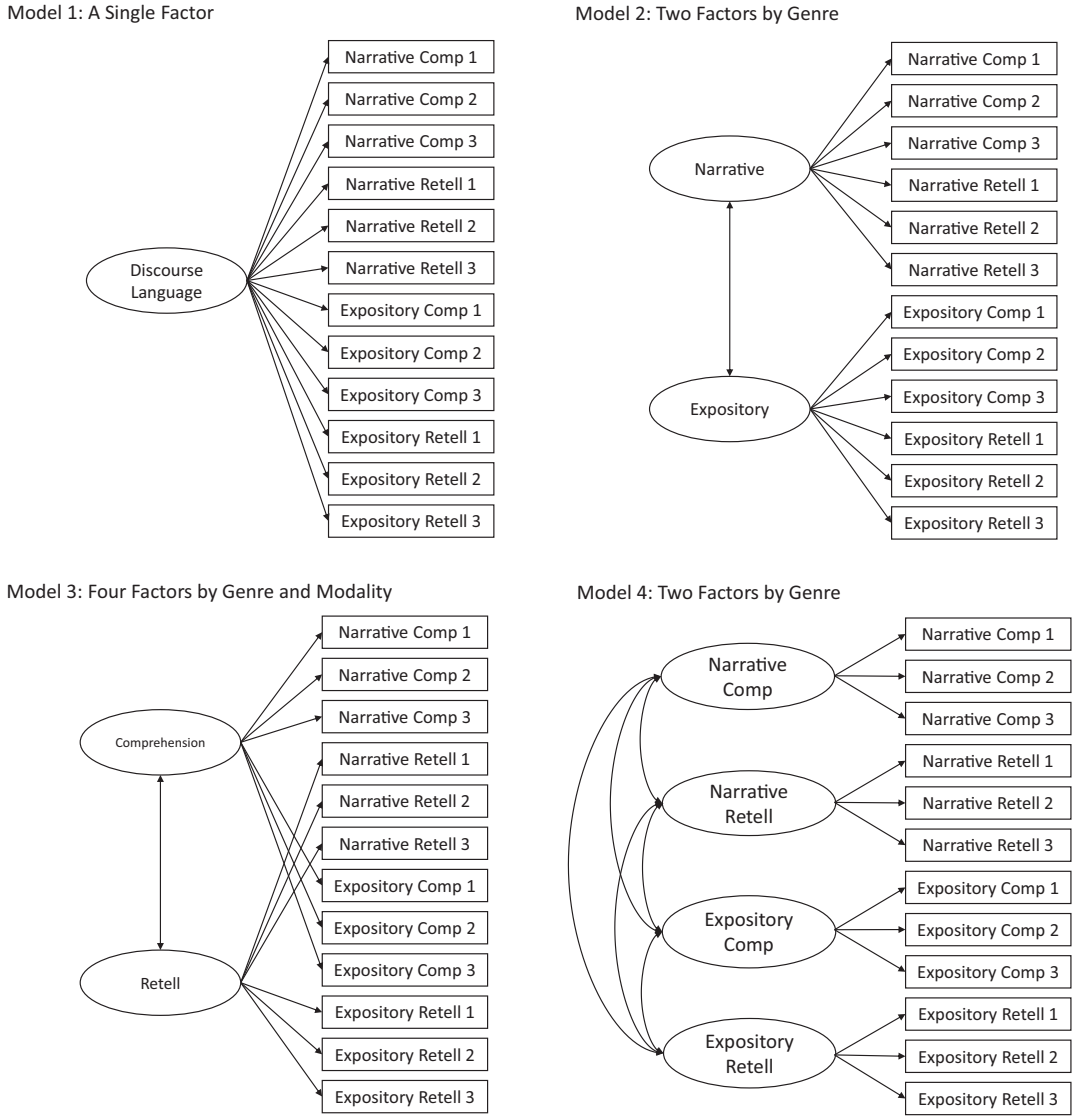


FIGURE 1 (Continued)

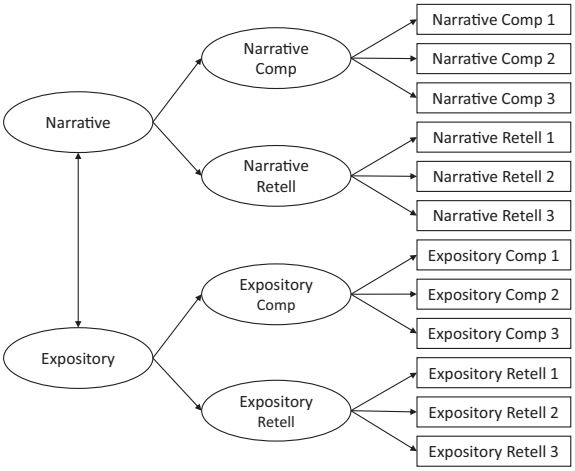
the null hypothesis of missing completely at random cannot be rejected ($\chi^2 = 136.452$, df = 152, $p = .81$). Descriptive statistics are presented in Table 1. Children's mean performances on narrative tasks were higher than those on expository tasks. For example, mean performance on the TNL Task 5 was 61% (8.53 out of possible 14 points) whereas mean performance on the QRI Task 3 was 36% (2.89 out of possible 8 points). These results are in line with previous evidence that comprehension of narrative texts is easier than expository texts (Best et al., 2008; Wolfe & Woodwyk, 2010). For the normed tasks, including the TNL comprehension, CASL Inference, WJ Picture Vocabulary, and CASL Grammaticality Judgement, the mean standard score was in the average range (see Table 1). Distributional properties in terms of skewness and kurtosis were all adequate for subsequent analysis. Raw scores were used in subsequent analysis.

Bivariate correlations are presented in Table 2. Children's performances among the three tasks for each of the narrative and expository comprehension and retell measures were moderately related to each other ($.38 \le rs \le .57$). Their performances on narrative and expository comprehension and retell tasks were weakly to moderately related to their language and cognitive skills ($.12 \le rs \le .52$).

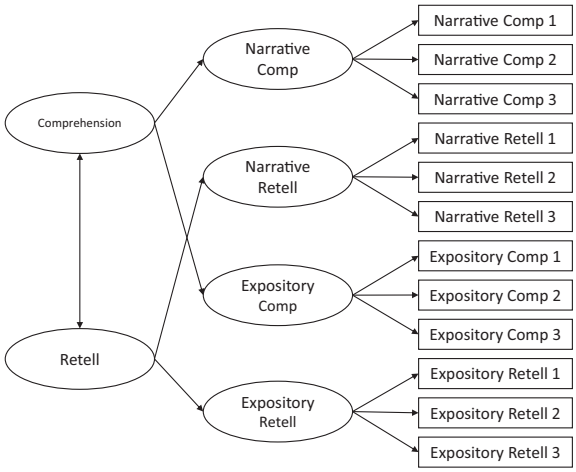
Dimensionality of comprehension and retell of narrative and expository texts

The eight alternative models shown in Figure 1 were fit to the data, and model fits are shown in Table 3. The majority of models had excellent fit to the data except for the single-factor model (Model 1) and the two-factor model by genre (Model 2). However, model comparison in Table 3 (see last column) indicated that Model 4, the four-factor model, fit the data best. Therefore, Model 4 was chosen as the final model. In other words, oral

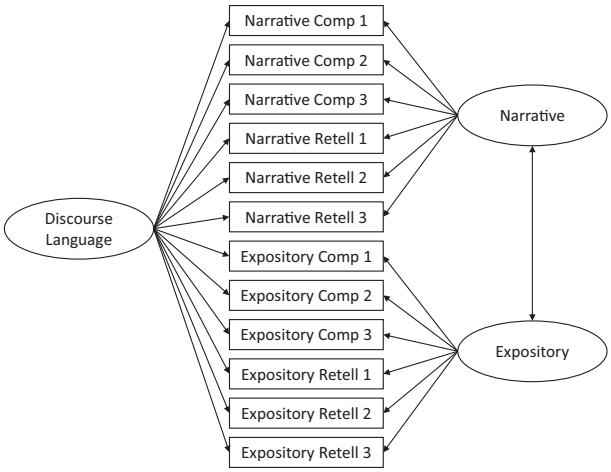
Model 5: Second-Order by Genre



Model 6: Second-Order by Modality



Model 7: Bifactor with Genre Specific Factors



Model 8: Bifactor with Modality Specific Factors

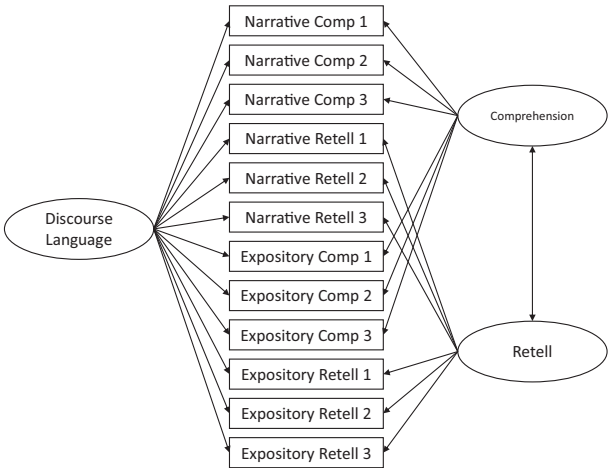


FIGURE 1 Alternative models of narrative and expository comprehension and retell. (1.1) Model 1: A Single Factor, (1.2) Model 2: Two Factors by Genre, (1.3) Model 3: Two Factors by Modality, (1.4) Model 4: Four Factors by Genre and Modality, (1.5) Model 5: Second-Order by Genre, (1.6) Model 6: Second-Order by Modality, (1.7) Model 7: Bifactor with Genre Specific Factors, (1.8) Model 8: Bifactor with Modality Specific Factors. Comp, Comprehension.

TABLE 1 Descriptive statistics.

Variable	*N*	*M*	SD	Min	Max	Skewness	Kurtosis
TNL Comp Task 1	527	8.50	2.56	1	15	−0.33	−0.09
TNL Comp Task 3	527	8.90	1.48	0	11	−1.32	4.18
TNL Comp Task 5	527	8.53	2.34	1	13	−0.75	0.46
TNL Retell Task 1	528	8.74	4.93	0	18	−0.27	−1.01
TNL Retell Task 3	526	11.37	3.98	0	24	−1.00	1.28
TNL Retell Task 5	522	10.67	5.36	0	21	−0.33	−0.77
TNL Comp SS	527	8.47	2.86	1	16	−.20	−.04
QRI Comp Task 1	524	2.53	1.29	0	7	0.81	0.84
QRI Comp Task 2	524	3.93	1.67	0	8	−0.24	−0.31
QRI Comp Task 3	524	2.89	1.54	0	8	0.42	−0.04
QRI Retell Task 1	523	2.11	2.39	0	12	1.21	1.19
QRI Retell Task 2	519	4.07	3.03	0	16	0.79	0.43
QRI Retell Task 3	525	4.42	3.71	0	20	1.19	1.81
CASL Inference	527	10.59	6.87	0	32	0.60	−0.35
CASL Inference SS	527	92.70	12.96	56	132	0.21	−0.29
Theory of Mind	525	7.93	4.09	0	18	0.03	−0.78
Comp Monitoring	523	6.72	2.97	1	16	0.46	−0.34
WJ Picture Vocabulary	528	20.23	2.93	7	29	−0.07	0.69
WJ Picture Vocabulary SS	528	96.78	10.29	43	126	−0.35	1.30
CASL Grammaticality	527	31.35	12.88	1	66	0.03	−0.24
CASL Grammaticality SS	527	95.50	13.16	50	134	−0.39	0.56
Working Memory	528	7.78	4.05	0	20	0.06	−0.07
Attentional Control	515	120.42	34.80	36	210	0.44	0.18

Note: TNL Standard Score has a mean of 10 with SD of 3, whereas mean and SD for CASL Inference, WJ Picture Vocabulary, and CASL Grammaticality Standard Scores are 100 and 15. In the vast majority of tasks, the percentage of children who scored 0 was small, ranging from 0% to 8%. An exception was the QRI Retell Task 1 (39%, $n=205$). Although this resulted in an asymmetrical distribution for the QRI Retell Task 1, skewness and kurtosis values were within acceptable ranges, and maximum likelihood estimator is robust unless distributional properties are severely nonnormally distributed.

Abbreviations: CASL, Comprehensive Assessment of Spoken Language; Comp, Comprehension; QRI, Qualitative Reading Inventory; SS, Standard Score; TNL, Test of Narrative Language; WJ, Woodcock-Johnson.

discourse skills are best characterized as four related but dissociable skills of narrative comprehension, narrative retell, expository comprehension, and expository retell. Standardized coefficients for Model 4 are presented in Figure 2. Loadings were all strong ($.62 \leq \lambda s \leq .78$, $ps < .001$). The narrative comprehension, narrative retell, expository comprehension, and expository retell factors were all fairly strongly to strongly related to each other ($.59 \leq rs \leq .84$).

Relations of language and cognitive skills to the identified dimensions

The structural equation model was fit to the data and the model fit was excellent: χ^2 (98) = 133.44, $p = .01$, RMSEA = .03 [90% CI = .01, .04], CFI = .99, and SRMR = .03. Standardized coefficients for statistically significant paths are displayed in Figure 3. Narrative comprehension was directly or independently predicted by inference (.26, $p < .001$), theory of mind (.32, $p < .001$), comprehension monitoring (.12, $p = .005$), vocabulary (.16, $p < .001$), and grammatical knowledge (.19, $p < .001$). Narrative retell was independently predicted by inference (.28, $p < .001$) and theory of mind (.17, $p = .001$). Expository comprehension was independently predicted by inference (.18, $p = .001$), theory of mind (.29, $p < .001$), comprehension monitoring (.14, $p = .002$), vocabulary (.17, $p = .001$), and grammatical knowledge (.17, $p = .001$). Expository retell was independently predicted by inference (.16, $p = .006$), theory of mind (.25, $p < .001$), and comprehension monitoring (.17, $p = .001$). Total variance explained by the included language and cognitive predictors was as follows: .71 for narrative comprehension, .26 for narrative retell, .60 for expository comprehension, and .41 for expository retell.

Standardized total effects of language and cognitive skills on narrative comprehension, narrative retell, expository comprehension, and expository retell are presented in Table 4. Total effects of vocabulary and grammatical knowledge ranged from .16 to 42, and the effects were larger for comprehension than for retell across narrative and expository texts. Total effects of cognitive skills (theory of mind, inference, comprehension monitoring,

TABLE 2 Bivariate correlations.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. TNL Comp Task 1	—																	
2. TNL Comp Task 3	.41	—																
3. TNL Comp Task 5	.48	.50	—															
4. TNL Retell Task 1	.36	.32	.34	—														
5. TNL Retell Task 3	.33	.43	.34	.56	—													
6. TNL Retell Task 5	.35	.34	.45	.57	.54	—												
7. QRI Comp Task 1	.36	.31	.41	.18	.28	.27	—											
8. QRI Comp Task 2	.32	.39	.44	.23	.31	.36	.40	—										
9. QRI Comp Task 3	.37	.37	.43	.28	.35	.32	.45	.46	—									
10. QRI Retell Task 1	.25	.27	.29	.27	.29	.36	.47	.36	.43	—								
11. QRI Retell Task 2	.27	.32	.28	.33	.37	.40	.28	.46	.37	.38	—							
12. QRI Retell Task 3	.34	.34	.37	.34	.38	.46	.40	.40	.53	.49	.48	—						
13. Inference	.52	.40	.49	.29	.36	.37	.41	.36	.44	.35	.24	.40	—					
14. Theory of Mind	.42	.44	.49	.24	.30	.32	.37	.43	.41	.35	.29	.38	.48	—				
15. Comp Monitoring	.36	.32	.35	.21	.28	.25	.30	.34	.33	.30	.24	.34	.47	.36	—			
16. Vocabulary	.40	.37	.42	.17	.25	.27	.39	.34	.35	.30	.20	.33	.49	.40	.31	—		
17. Grammar	.46	.39	.48	.21	.28	.29	.38	.36	.42	.32	.23	.36	.60	.41	.39	.50	—	
18. Working Memory	.26	.22	.33	.13	.16	.19	.20	.33	.24	.25	.18	.24	.28	.31	.21	.36	.39	—
19. Attentional Control	.18	.16	.27	.12	.12	.19	.20	.21	.19	.18	.13	.25	.23	.26	.27	.17	.26	.27

Note: All coefficients are statistically significant at $p < .001$.

Abbreviations: Comp, Comprehension; Grammar, Grammaticality task of Comprehensive Assessment of Spoken Language; QRI, Qualitative Reading Inventory; TNL, Test of Narrative Language; Vocabulary, Woodcock-Johnson Picture Vocabulary Task.

TABLE 3 Model fit and comparison of confirmatory factor analysis models shown in Figure 1.

Model	χ^2 (df), *p* value	CFI	RMSEA [90% CI]	SRMR	*n*BIC	Model comparison
1. A Single Factor (Figure 1.1)	357.20 (48), <.001	.86	.11 [.10, .12]	0.06	27,937.45	NA
2. Two Factors by Genre: Narrative & Expository (Figure 1.2)	267.45 (47), <.001	.90	.09 [.08, .11]	.05	27,850.80	Model 1 versus Model 2 $\Delta\chi^2 = 89.75$, $\Delta df = 1$, $p < .001$ Model 2 is superior to Model 1
3. Two Factors by Modality: Comprehension & Retell (Figure 1.3)	186.85 (47), <.001	.94	.08 [.06, .09]	.05	27,673.80	Model 2 versus Model 3 ΔnBIC = 177.00 (very strong) Model 3 is superior to Model 2
4. Four Factors by Genre & Modality: Narrative Comp & Retell, and Expository Comp & Retell (Figure 1.4)	56.71 (42), .06	.99	.03 [.00, .04]	.02	27,655.55	Model 3 versus Model 4 $\Delta\chi^2 = 130.14$, $\Delta df = 5$, $p < .001$ Model 4 is superior to Model 3
5. Second-Order by Genre: Narrative & Expository Factors (Figure 1.5)	99.14 (43), <.001	.98	.05 [.04, .06]	.03	27,694.88	Model 4 versus Model 5 ΔnBIC = 39.33 (very strong) Model 4 is superior to Model 5
6. Second-Order by Modality: Comp & Retell Factors (Figure 1.6)	78.06 (43), <.001	.98	.04 [.03, .05]	.03	27,673.80	Model 4 versus Model 6 ΔnBIC = 18.25 (very strong) Model 4 is superior to Model 6
7. Bifactor with Genre, Narrative & Expository, Specific Factors (Figure 1.7)	52.55 (35), .03	.99	.03 [.01, .05]	.02	27,673.06	Model 4 versus Model 7 ΔnBIC = 17.51 (very strong) Model 4 is superior to Model 7
8. Bifactor with Modality, Comp & Retell, Specific Factors (Figure 1.8)	42.91 (35), .17	1.00	.02 [.00, .04]	.02	27,663.43	Model 4 versus Model 8 ΔnBIC = 7.88 (strong) Model 4 is superior to Model 8

Abbreviations: CFI, comparative fit index; Comp, Comprehension; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual; *n*BIC, sample size adjusted Bayesian information criterion.

working memory, and attentional control) ranged from .10 to 35. Total effects of our focal cognitive skills for the research question, theory of mind, working memory, and attentional control, were as follows. Theory of mind had its largest total effect on narrative comprehension (.32), followed by expository comprehension (.29), expository retell (.25), and narrative retell (.17). The total effects of working memory had a similar pattern as that for theory of mind ranging from .18 to .35. Attentional control had similar sized total effects on narrative comprehension, expository comprehension, and expository retell (.21 to .22) whereas it was smaller for narrative retell (.15).

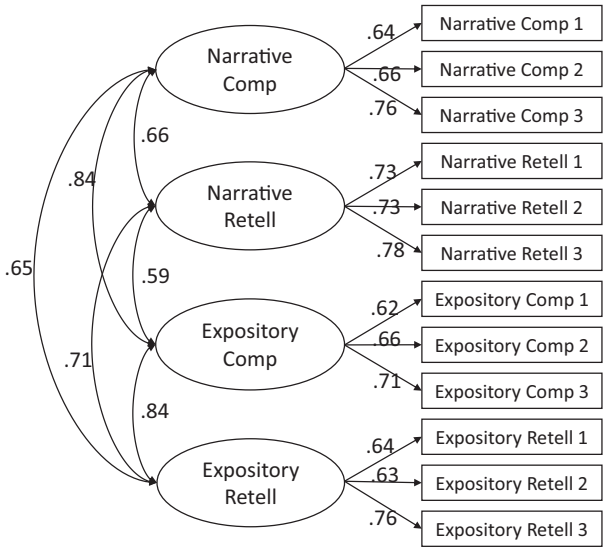


FIGURE 2 Standardized loadings and correlations among the factors. Comp, Comprehension.

DISCUSSION

There are multiple grain sizes and aspects of oral language (e.g., vocabulary, syntax, and discourse; receptive and expressive). In the present study, we focused on and systematically investigated the dimensionality of oral *discourse* skills in two modalities, comprehension and retell, and two genres, narrative and expository genres, using data from English-speaking children in Grade 2. To our knowledge, this is the first study that examined the dimensionality of oral discourse skills by including comprehension and retell of narrative and expository texts, and examined the relations of a relatively comprehensive set of language and cognitive skills to the identified dimensions.

Overall, the results revealed that comprehension and retell of oral discourse skills are best characterized as a multidimensional construct, consisting of four related but dissociable constructs of narrative comprehension, narrative retell, expository comprehension, and expository retell. The four dimensions were fairly strongly to strongly related to each other (*r*s = .59 to .84). These results indicate that children's skills to comprehend and retell narrative texts and expository texts are related but dissociable. In other words, although children who are strong in comprehension tend to be also strong in retell,

TABLE 4 Standardized direct, indirect, and total effects from structural equation model shown in Figure 3.

	Direct effect	Indirect effect	Total effect	Direct effect	Indirect effect	Total effect
Narrative comprehension outcome				Narrative retell outcome		
Theory of mind	.32 (.04)***	—	.32 (.04)***	.17 (.05)**	—	.17 (.05)**
Inference	.26 (.05)***	—	.26 (.05)***	.28 (.06)***	—	.28 (.06)***
Comprehension monitoring	.12 (.04)***	—	.12 (.04)**	.10 (.05)	—	.10 (.05)
Vocabulary	.16 (.04)***	.15 (.03)***	.32 (.05)***	.04 (.05)	.12 (.02)***	.16 (.05)**
Grammatical knowledge	.22 (.03)***	.19 (.05)***	.42 (.05)***	.03 (.06)	.19 (.03)***	.22 (.06)***
Working memory	.07 (.04)	.28 (.03)***	.35 (.05)***	.03 (.05)	.15 (.03)***	.18 (.05)***
Attentional control	.03 (.04)	.18 (.03)***	.21 (.05)***	.04 (.05)	.11 (.02)***	.15 (.05)**
Expository comprehension outcome				Expository retell outcome		
Theory of mind	.29 (.05)***	—	.29 (.05)***	.25 (.05)	—	.25 (.05)***
Inference	.18 (.06)**	—	.18 (.06)**	.16 (.06)**	—	.16 (.06)**
Comprehension monitoring	.14 (.05)**	—	.14 (.05)**	.17 (.05)**	—	.17 (.05)**
Vocabulary	.17 (.05)**	.13 (.02)***	.30 (.05)***	.09 (.05)	.12 (.02)***	.21 (.05)***
Grammatical knowledge	.17 (.05)**	.19 (.03)***	.36 (.05)***	.09 (.06)	.18 (.03)***	.27 (.05)***
Working memory	.08 (.05)	.25 (.03)***	.33 (.05)***	.08 (.05)	.19 (.03)***	.27 (.05)***
Attentional control	.05 (.05)	.16 (.03)***	.21 (.05)***	.08 (.05)	.14 (.03)***	.22 (.05)***

<.01; *<.001.

there are also children whose relative performances are discrepant. This also holds for genre: Although children who are strong in comprehension and retell of narrative texts tend to be strong in expository texts, there are children whose relative performances are discrepant. The substantial relations between comprehension and retell across the genres support theoretical models of discourse skills such as the construction-integration model (Kintsch, 1988), which hypothesizes that essentially the same underlying processes are involved in discourse comprehension and production. However, this does not entail that recruitment of processes and associated language and cognitive skills are identical for comprehension and retell (Kim, 2016; Kim, 2020), and the present findings of dissociability of comprehension and retell provide empirical evidence for it. The present findings of multidimensionality of oral discourse skills are in line with previous work using narrative texts (Gillam & Pearson, 2004), but discrepant from Kim et al.'s (2015) study, which found a bifactor structure for narrative comprehension, retell, and production. Note, however, that these previous studies only examined narrative texts, and thus, the present findings cannot be directly compared to them. Overall, the results expand our understanding of the dimensionality of oral discourse skills of comprehension and retell of narrative and expository texts.

Multidimensionality of oral discourse skills suggests that underlying skills and knowledge might differentially contribute to the identified dimensions, and results indeed revealed different patterns. The relation by modality was stronger for comprehension (.84) than for retell (.71), and the relation by genre was stronger for expository texts (.84) than for narrative texts (.66). Reasons for these findings are suggested in the differential contributions of language and cognitive skills shown in Table 4. For example, the contributions of language and cognitive skills to *comprehension* of narrative texts and expository texts had a more similar pattern than the pattern for *retell* of narrative texts and expository texts (compare left panel with right panel of Table 4). Specifically, the contributions or total effects of theory of mind, vocabulary, grammatical knowledge, working memory, and attentional control to *comprehension* were more similar and larger across narrative and expository genres compared to the pattern of contributions to *retell* of narrative texts versus expository texts, which was more divergent. One potential explanation, as noted above, is that comprehension is probed with different types of comprehension questions (literal and inferential), and accurately responding to them (i.e., comprehension) may elicit both shallow and deep comprehension and place greater demands on the language and cognitive skills. In contrast, retell does not include specific probes and this might

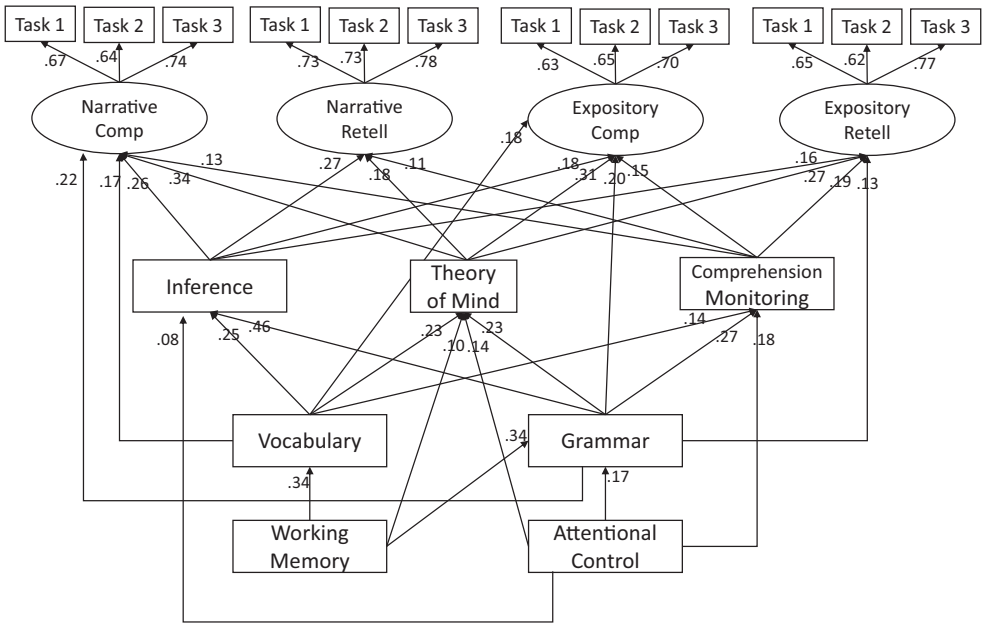


FIGURE 3 Standardized coefficients for the relations of language and cognitive skills to narrative comprehension, narrative retell, expository comprehension, and expository retell. Comp, Comprehension. Paths were allowed from all the language and cognitive skills to narrative comprehension, narrative retell, expository comprehension, and expository retell. In addition, lower order skills also predicted higher order skills. In order to reduce clutter, only statistically significant ($p < .05$) predictive paths are shown. Statistically significant correlations are as follows: .26 between inference and theory of mind; .29 between inference and comprehension monitoring; .18 between theory of mind and comprehension monitoring; .41 between vocabulary and grammatical knowledge; .28 between working memory and attentional control; .50 between narrative comprehension and narrative retell; .56 between narrative comprehension and expository comprehension; .30 between narrative comprehension and expository retell; .37 between narrative retell and expository comprehension; .59 between narrative retell and expository retell; .72 between expository comprehension and expository retell.

elicit shallow comprehension or recall of texts. Although it was argued that retell could provide insight into children's deep comprehension of texts such as their understanding of text structure and interpretation (e.g., Kida et al., 2016; Loyd & Steele, 1986), evidence indicated that retell is not strongly related with other measures of comprehension, at least in the context of reading comprehension (see a recent meta-analysis, Cao & Kim, 2021).

The hypothesis of a greater demand of working memory for retell compared to comprehension was not supported in the present study. As noted above, previous work in the context of written texts (i.e., Spencer et al., 2019) found that working memory made a unique contribution to retell but not to comprehension. In the present study, total effects of working memory were larger for comprehension than for retell. One potential explanation for the stronger relation of working memory to narrative comprehension than narrative retell might be the way working memory was measured—a listening span task where the stimuli involved *comprehension* of simple three-word sentences (e.g., Birds can fly). Although children's responses on the veracity of the statements were not scored and only their recall of final words was scored, the nature of the task (comprehension) might have influenced the present results. However, if this were the case, a stronger relation of working memory to comprehension should apply to expository texts; and therefore, this does not appear to be a strong candidate for explaining the present findings.

Different patterns of the contributions of language and cognitive skills to the four identified dimensions were also found by genre, narrative versus expository texts, which is in line with DIET (Kim, 2016). Inference had a stronger relation to comprehension and retell of narrative texts than expository texts. These findings may be attributed to differences in the degree of cohesion of texts for narrative versus expository texts. It was reported that narrative texts are less cohesive than expository texts (McNamara et al., 2012), and constructing an accurate situation model for less cohesive texts likely requires one to draw on inferences to a greater extent (Eason et al., 2012). Interestingly, all but inference—theory of mind, comprehension monitoring, vocabulary, grammatical knowledge, working memory, and attentional control—had larger total effects on the retell of *expository* texts than on narrative texts (Table 4). It is important to unpack these findings. First, unlike a previous study (Eason et al., 2012), which found a stronger relation of vocabulary to reading comprehension for narrative texts than expository texts, the total effects of vocabulary did not differ for narrative versus expository comprehension. Reasons for this discrepancy are unclear. One potential explanation is that narrative texts used in Eason et al.'s (2012) study may have had greater vocabulary demands, given that participants in that study were older (ages 10–14 years). Second, while the consistent roles of working memory and attentional control in oral discourse skills are convergent with prior empirical findings (e.g., Daneman & Merikle, 1996; Florit et al., 2011; Kendeou et al., 2008; Kim, 2015, 2016), the hypothesis of stronger relations of working memory and attentional control to expository texts than to narrative texts (Eason et al., 2012; Wu et al., 2020) was partially supported: The total effects of working memory and attentional control were larger for expository versus narrative texts in retell, but not comprehension (see Table 4). The discrepant findings may be attributed to differences between oral discourse and written discourse. Prior work (Eason et al., 2012; Spencer et al., 2019; Wu et al., 2020) was conducted involving written texts (reading comprehension) whereas the present work involved oral texts. In oral discourse, texts are ephemeral in that presented texts are no longer accessible to the listener whereas in written discourse, texts are available and the reader can go back to the text as needed. The ephemeral nature of oral texts may present similar demands of working memory and attentional control across genres of texts for successful comprehension. In contrast, the presence and accessibility of texts in written discourse contexts may allow one to differentially use mental resources as a function of genre, such that working memory and attentional control may exert a greater influence on expository texts, with the effect being particularly greater in unprobed retell/recall than comprehension. Future work is needed to further investigate whether demands of working memory and attentional control differ as a function of modality (comprehension vs. retell) and by oral discourse versus written discourse.

Third, the hypothesis about a differential role of theory of mind as a function of genre was partially supported. It was posited that theory of mind would play a greater role in comprehension of narrative texts than expository texts, and this was supported as theory of mind was more strongly related to narrative comprehension (.32) than expository comprehension (.29). However, the difference in magnitude was small (.03; this difference was statistically significant when an equality constraint was placed; results not shown). Theory of mind was also related to retell, and surprisingly, the relation was stronger for expository texts (.25) than for narrative texts (.17). Although theory of mind has been typically conceptualized and examined in the context of narrative texts, theory of mind is broadly a reasoning skill, and therefore, is important to mental representation of texts across genres, including expository texts (Kim, 2016). Importantly, text characteristics vary beyond narrativity, including linguistic features (e.g., mental state verbs such as *know*, *think*, and *believe*) and content, and these might trigger recruitment of perspective taking for encoding information and establishing a mental representation of expository texts (Kim et al., 2021). The expository texts in the present study included few mental verbs, and therefore, this might not explain the present findings. However, content of texts might be a candidate

for explaining the presenting findings. For example, the text, *Where Do People Live?*, was about differences in living in cities, the country, and the suburbs, and reasons why people live in different places; and establishing deep understanding of the latter (reasons why people live in difference places) might have induced children to consider varying perspectives in terms of what people value and associated choice of living in cities, the country, or the suburbs.

It is also interesting that the relations of theory of mind to oral discourse skills are generally stronger for comprehension of texts than for retell. We speculate that this is because retell for young children may not fully capture representation of one's understanding of perspectives. Although theoretically coherent retell is contingent on the situation model which is based on one's understanding of multiple viewpoints, unlike comprehension tasks where specific questions are designed to probe literal and inferential higher order comprehension, retell is not typically probed (i.e., free retell/recall). Therefore, unprobed retell for young children likely reflects the textbase representation which is a propositional network rather than the higher order situation model which incorporates multiple perspectives. This explanation was suggested in a recent study, which revealed that Grade 4 children's retell included a limited number of mental state talk that captures higher order cognitive skills (e.g., *believe*, *decide*, *think*; Kim et al., 2021).

An unexpected, surprising finding was that the relations of language and cognitive skills were consistently weaker for narrative retell compared to the other dimensions of oral discourse skills (see Table 4). Not surprisingly, the amount of variance explained was considerably smaller (26%) for narrative retell compared to the others (41% to 71%). The fact that the same language and cognitive skills explained different amounts of variance suggests that the four dimensions are dissociable constructs and differentially draw on language and cognitive skills. However, reasons for the drastically smaller amount of variance explained and weaker relations of cognitive skills to narrative retell are not clear, and future studies are needed.

Limitations, future directions, implications, and conclusion

Several limitations are worth noting. First, oral discourse skills include production of narrative and expository texts, multi-utterance conversations, and discussions in addition to comprehension and retell. Previous studies indicated that retell and production are likely to tap into a single construct (e.g., Gillam & Pearson, 2004; Kim et al., 2015), but future work is necessary to further investigate the dimensionality of oral discourse skills by including multiple aspects of oral discourse skills beyond comprehension and retell.

A second limitation is that although the language and cognitive skills included in the current study were relatively comprehensive, future studies should consider expanding them (e.g., topic knowledge, Best et al., 2008). In addition, future studies can expand the measurement of attentional control. Attentional control can be measured using rating scales and direct cognitive assessments, and in the present study, only the former type of measure was used. Although the validity of rating scales such as SWAN was shown (Arnett et al., 2013; Sáez et al., 2012), a future study that includes both types of attentional control measures would be useful.

A third limitation is that higher order cognitive skills, inference, theory of mind, and comprehension monitoring were measured using tasks that required some level of discourse comprehension, and this might explain their relations to oral discourse skills. This is because of the nature of these constructs. For example, making inferences using background knowledge requires inferring information in a given context (which was established by oral texts in the present study). Similarly, comprehension monitoring by definition involves some level of comprehension. In the present study, these constructs were measured using a normed measure for inference, and using well-established protocols for theory of mind (false-belief) and comprehension monitoring (inconsistency detection). Future studies can explore alternative ways to measure higher order cognitive skills and replicate and expand the present study.

Future studies can also examine the extent to which retell captures recall versus interpretation of the text. One way to achieve this is coding retell data for the extent to which retell contains literal information and inferential information. In the present study, retell was coded for overall quality following previous studies (see Gillam & Pearson, 2004 for a review), but did not specifically identify literal and inferential information. Coding literal and inferential information can illuminate whether the child's retell primarily captures recall or their deeper understanding of the text.

Lastly, the number of literal and inferential comprehension items was not equivalent in the narrative and expository comprehension tasks, which is a reflection of published assessments. Note, however, that there was sufficient variance and covariance in both types of comprehension items across narrative and expository texts, which was critical in addressing the research questions in the present study. Future work can replicate the present study with balanced items for literal and inferential questions in different genres.

The findings of this study are correlational and therefore, causal implications are limited. Notwithstanding, the results suggest the importance of taking into account modalities and genres in assessment and instruction of oral discourse skills. In terms of assessment, a reasonable implication is to include multiple genres, not a single genre, to fully capture children's oral discourse skills, though

normed and standardized assessments of oral discourse skills that include informational texts are limited. The dissociability of comprehension and retell also suggests inclusion of retell, in addition to comprehension, as part of a comprehensive assessment of children's oral discourse skills. Retell is widely assessed as a proxy for comprehension in the context of measuring reading comprehension in US elementary grades (e.g., Dynamic Indicators of Basic Early Literacy Skills and informal assessments such as QRI). The present findings, together with a recent meta-analysis (Cao & Kim, 2021), indicate caution against using retell as the sole measure of comprehension.

The findings also suggest instructional efforts in the four identified dimensions. A call for greater attention to informational texts, in addition to narrative texts, in primary grades have been made (e.g., Duke, 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). What has not received much attention is explicit teaching of retell. Although retell is widely assessed as a proxy for comprehension in the context of measuring reading comprehension, retell is rarely systematically taught in primary grades, and this is an important missed opportunity, given the role of oral retell and production in writing (e.g., Juel et al., 1986; Kim & Schatschneider, 2017). The findings of contributions of language and cognitive skills to the comprehension and retell of narrative and expository texts indicate that instruction can target these skills to improve children's oral discourse skills. Studies have shown that language skills and higher order cognitive skills are malleable with systematic and explicit teaching (e.g., Biemiller & Boote, 2006; Kendeou et al., 2020).

Oral discourse is ubiquitous, and it is an important skill in all aspects of our lives. The present study revealed that oral discourse skills are related but dissociable along the lines of genre (narrative vs. expository) and modality (comprehension vs. retell). Future efforts should shed further light on our nuanced understanding of the skills and knowledge that contribute to oral discourse skills.

ACKNOWLEDGMENTS

The author thanks study participants, including school leadership, teachers, parents, and children.

FUNDING INFORMATION

This research was supported by grants from the Institute of Education Sciences (IES), US Department of Education, R305A200312, R305A180055, and R305A130131 and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), P50HD052120. The content is solely the responsibility of the author and does not necessarily represent the official views of the funding agency.

DATA AVAILABILITY STATEMENT

The data and materials necessary to reproduce the analyses presented here are not publicly accessible. The analyses presented here were not preregistered, but the analytic code necessary to reproduce the analyses presented in this paper is available from the author upon reasonable request.

ORCID

Young-Suk Grace Kim https://orcid.org/0000-0002-4328-3843

REFERENCES

Ahmed, Y., Miciak, J., Taylor, W. P., & Francis, D. J. (2021). Structure altering effects of a multicomponent reading intervention: An application of the direct and inferential mediation (DIME) model of reading comprehension in upper elementary grades. *Journal of Learning Disabilities*, *55*(1), 58–78. https://doi.org/10.1177/0022219421995590

Anthony, J. L., Davis, C., Williams, J. M., & Anthony, T. I. (2014). Preschoolers' oral language abilities: A multilevel examination of dimensionality. *Learning and Individual Differences*, *35*, 56–61. https://doi.org/10.1016/j.lindif.2014.07.004

Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., & Olson, R. K. (2013). The SWAN captures variance at the negative and positive end of the ADHD symptom dimension. *Journal of Attention Disorders*, *17*, 152–162. https://doi.org/10.1177/1087054711427399

Baker, L. (1984). Children's effective use of multiple standards for evaluating their comprehension. *Journal of Educational Psychology*, *76*, 588–597. https://doi.org/10.1037/0022-0663.76.4.588

Barnes, A. E., Kim, Y. S., & Phillips, B. M. (2014). The relations of proper character introduction to narrative quality and listening comprehension for young children from high poverty schools. *Reading and Writing*, *27*(7), 1189–1205. https://doi.org/10.1007/s11145-013-9481-0

Bellinger, J. M., & Diperna, J. C. (2011). Is fluency-based story retell a good indicator of reading comprehension? *Psychology in the Schools*, *48*(4), 416–426. https://doi.org/10.1002/pits.20563

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, *43*(2), 79–120. https://doi.org/10.1080/01638530709336894

Best, R. M., Floyd, R. G., & McNamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*(2), 137–164. https://doi.org/10.1080/02702710801963951

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62. https://doi.org/10.1037/0022-0663.98.1.44

Cain, K., & Oakhill, J. (Eds.). (2007). *Children's comprehension problems in oral and written language: A cognitive perspective*. Guilford Press.

Cao, Y., & Kim, Y.-S. G. (2021). Is retell a valid measure of reading comprehension? *Educational Research Review*, *32*, 100375. https://doi.org/10.1016/j.edurev.2020.100375

Carrow-Woolfolk, E. (1999). *Comprehensive assessment of spoken language*. Pearson Assessment.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

Dore, R. A., Amendum, S. J., Golinkoff, R. M., & Hirsh-Pasek, K. (2018). Theory of mind: A hidden factor in reading comprehension? *Educational Psychology Review*, *30*, 1067–1089. https://doi.org/10.1007/s10648-018-9443-9

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly*, *35*(2), 202–224.

Eason, S., Goldberg, L. F., Young, K. M., Geist, M. C., & Cutting, L. E. (2012). Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. *Journal of Educational Psychology*, *104*, 515–528. https://doi.org/10.1037/a0027182

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. https://doi.org/10.1080/0163853X.2010.494244

Gillam, R. B., & Pearson, N. A. (2004). *Test of narrative language*. Pro-Ed.

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

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 193–202.

Graesser, A. C., McNamara, D. S., & VanLehn, K. (2005). Scaffolding deep comprehension strategies through Point&Query, AutoTutor, and iSTART. *Educational Psychologist*, *40*(4), 225–234. https://doi.org/10.1207/s15326985ep4004_4

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology*, *78*(4), 243–255. https://doi.org/10.1037/0022-0663.78.4.243

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. https://doi.org/10.1111/j.1467-9817.2008.00370.x

Kendeou, P., McMaster, K. L., Butterfuss, R., Kim, J., Bresina, B., & Wagner, K. (2020). The inferential language comprehension (iLC) framework: Supporting children's comprehension of visual narratives. *Topics in Cognitive Science*, *12*, 256–273. https://doi.org/10.1111/tops.12457

Kendeou, P., & van den Broek, P. (2007). The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, *35*(7), 1567–1577. https://doi.org/10.3758/BF03193491

Kida, A. S. B., de Ávila, C. R. B., & Capellini, S. A. (2016). Reading comprehension assessment through retelling: Performance profiles of children with dyslexia and language-based learning disability. *Frontiers in Psychology*, *7*, 787. https://doi.org/10.3389/fpsyg.2016.00787

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

Kim, Y.-S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development*, *86*, 128–144. https://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. https://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*(4), 310–333. https://doi.org/10.1080/10888438.2017.1291643

Kim, Y.-S.G. (2020). 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*, *112*(4), 667–684. http://dx.doi.org/10.1037/edu0000407

Kim, Y.-S. G., Dore, R. A., Cho, M., Golinkoff, R., & Amendum, S. (2021). Theory of mind, mental state talk, and discourse comprehension: Theory of mind process is more important for narrative comprehension than for informational text comprehension. *Journal of Experimental Child Psychology*, *209*, 105181. https://doi.org/10.1016/j.jecp.2021.105181

Kim, Y.-S. G., & Petscher, Y. (2021). Influences of person, text, and assessment factors on text/discourse comprehension in oral language (listening comprehension). *Annals of Dyslexia*, *71*(2), 218–237. https://doi.org/10.1007/s11881-020-00208-8

Kim, Y.-S. G., Park, C., & Park, Y. (2015). Dimensions of discourse-level oral language skills and their relations to reading comprehension and written composition: An exploratory study. *Reading and Writing: An Interdisciplinary Journal*, *28*, 633–654. https://doi.org/10.1007/s11145-015-9542-7

Kim, Y.-S. G., & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology*, *109*, 35–50. https://doi.org/10.1037/edu0000129

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction- integration model. *Psychological Review*, *95*, 163–182.

Language and Reading Research Consortium. (2015). The dimensionality of language ability in young children. *Child Development*, *86*(6), 1948–1965. https://doi.org/10.1111/cdev.12450

Language and Reading Research Consortium. (2017). Oral language and listening comprehension: The same or different constructs. *Journal of Speech, Language & Hearing Research*, *60*(5), 1273–1284. https://doi.org/10.1044/2017_JSLHR-L-16-0039

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. https://doi.org/10.1002/RRQ.020

Leslie, L., & Caldwell, J. S. (2011). *Qualitative Reading Inventory–5th edition*. Pearson.

Lonigan, C. J., & Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *Journal of Speech, Language & Hearing Research*, *60*(8), 2185–2198. https://doi.org/10.1044/2017_JSLHR-L-15-0402

Loyd, B. H., & Steele, J. L. (1986). Assessment of reading comprehension: A comparison of constructs. *Reading Psychology*, *7*(1), 1–10. https://doi.org/10.1080/0270271860070102

Mahy, C. E. V., Bernstein, D. M., Gerrard, L. D., & Atance, C. M. (2017). Testing the validity of a continuous false belief task in 3- to 7-year-old children. *Journal of Experimental Child Psychology*, *160*, 50–66. https://doi.org/10.1016/j.jecp.2017.03.010

McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across the ages and genres. In J. P. Sabatini & E. Albro (Eds.), *Assessing reading in the 21st century: Aligning and applying advances in the reading and measurement sciences* (pp. 89–116). R&L Education.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning form text. *Cognition and Instruction*, *14*(1), 1–43.

Miller, J. F., & Iglesias, A. (2006). *Systematic analysis of language transcripts (SALT)*. University of Wisconsin.

Mouzaki, A., Spyropoulou, E., Ralli, A., Antoniou, F., Diamanti, V., & Papaioannou, S. (2020). The dimensional of oral language ability: Evidence from young Greek children. *Journal of Speech, Language, and Hearing Research*, *63*(8), 2732–2751. https://doi.org/10.1044/2020_JSLHR-19-00008

Muijselaar, M. M. L., Swart, N. M., Steenbeek-Planting, E. G., Droop, M., Verhoeven, L., & de Jong, P. F. (2017). The dimensions of reading comprehension in Dutch children: Is differentiation by text and question type necessary? *Journal of Educational Psychology*, *109*(1), 70–83. https://doi.org/10.1037/edu0000120

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus 8.4*. Muthén and Muthén.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core state standards*. Author.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second- order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*, 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

Sáez, L., Folsom, J. S., Al Otaiba, S., & Schatschneider, C. (2012). Relations among student attention behaviors, teacher practices, and beginning word reading skill. *Journal of Learning Disabilities*, *45*, 418–432. https://doi.org/10.1177/0022219411431243

Shapiro, E. S., Fritschmann, N. S., Thomas, L. B., Hughes, C. L., & McDougal, J. (2014). Concurrent and predictive validity of reading retell as a brief measure of reading comprehension for narrative text. *Reading Psychology*, *35*(7), 644–665. https://doi.org/10.1080/02702711.2013.790328

Spencer, M., Gilmour, A. F., Miller, A. C., Emerson, A. M., Saha, N. M., & Cutting, L. E. (2019). Understanding the influence of text complexity and question type on reading outcomes. *Reading and Writing: An Interdisciplinary Journal*, *32*(3), 603–637. https://doi.org/10.1007/s11145-018-9883-0

Swanson, J. M., Schuck, S., Porter, M. M., Carlson, C., Hartman, K., Sergeant, J. A., Clevenger, W., Wasdell, M., McCleary, R., Lakes, K., & Wigal, T. (2006). Categorical and dimensional definitions and evaluations of symptoms of ADHD: History of SNAP and SWAN rating scales. *The International Journal of Educational and Psychological Assessment*, *10*(1), 51–70.

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

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing*, *24*(2), 203–220. https://doi.org/10.1007/s11145-010-9266-7

Wolfe, M. B., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, *80*(3), 341–362. https://doi.org/10.1348/000709910X485700

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

Wu, Y., Barquero, L. A., Pickren, S. E., Barber, A. T., & Cutting, L. E. (2020). The relationship between cognitive skills and reading comprehension of narrative and expository texts: A longitudinal study from grade 1 to grade 4. *Learning and Individual Differences*, *80*, 101848. https://doi.org/10.1016/j.lindif.2020.101848

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kim, Y.-S. (2023). Oral discourse skills: Dimensionality of comprehension and retell of narrative and expository texts, and the relations of language and cognitive skills to identified dimensions. *Child Development*, *94*, e246–e263. https://doi.org/10.1111/cdev.13935