Journal of Educational Psychology
2009, Vol. 101, No. 4, 765–778

0022-0663/09/$12.00 DOI: 10.1037/a0015956

Predicting Reading Comprehension in Early Elementary School: The Independent Contributions of Oral Language and Decoding Skills

Panayiota Kendeou
McGill University

Paul van den Broek
Leiden University

Mary Jane White
University of Memphis

Julie S. Lynch
Saginaw Valley State University

The authors examined the development of oral language and decoding skills from preschool to early elementary school and their relation to beginning reading comprehension using a cross-sequential design. Four- and 6-year-old children were tested on oral language and decoding skills and were retested 2 years later. In all age groups, oral language and decoding skills formed distinct clusters. The 2 clusters were related to each other in preschool, but this relation became weaker in kindergarten and 2nd grade. Structural equation modeling showed that both sets of skills in 2nd grade independently predicted a child's reading comprehension. These findings confirm and extend the view that the 2 clusters of skills develop early in a child's life and contribute to reading comprehension activities in early elementary school, with each cluster making a considerable, unique contribution.

Keywords: reading comprehension, oral language, decoding, development

Despite intensive instruction, many children in early elementary school fail to reach functional levels of literacy. As a result, there has been considerable research on early intervention programs for fostering decoding skills—skills that support decoding, such as phonological awareness and letter and word identification (see Ehri, Nunes, Stahl, & Willows, 2001; Papadopoulos, Das, Parrila, & Kirby, 2003; Snowling & Hulme, 2005; Storch & Whitehurst, 2002, for reviews). There also is growing recognition that the development of these skills should be complemented by fostering oral language skills[1]—skills that support comprehension, such as receptive vocabulary (i.e., understanding of spoken words) and narrative comprehension (e.g., Lonigan, Burgess, & Anthony, 2000; Pressley et al., 2001; see also the influential "simple view of reading," Gough & Tunmer, 1986; Hoover & Gough, 1990; Tunmer & Hoover, 1992). However, the findings with respect to the relative contribution of oral language skills in early reading comprehension have been contradictory.

On the one hand, the results of some studies suggest that oral language skills are not central to early reading comprehension and that they become fully operative only when the child has acquired decoding skills (Bryant, McLean, & Bradley, 1990; Speece, Roth, Cooper, & de la Paz, 1999; Vellutino, Tunmer, Jaccard, & Chen, 2007). On the other hand, results of other studies suggest the opposite, highlighting the importance of such skills in early reading comprehension (Bishop & Adams, 1990; Catts, Fey, Zhang, & Tomblin, 1999; Paris & Paris, 2003).

The lack of consensus in the literature may be the result of several limitations. One issue concerns the different ways that oral language skills and reading comprehension skills have been conceptualized and measured. The far-reaching implications of such variation, in both predictor and outcome measures, have been discussed recently with respect to several widely used tests of reading comprehension (Cutting & Scarborough, 2006; Fletcher, 2006; Keenan & Betjemann, 2006; Magliano, Millis, Ozuru, & McNamara, 2007; Ozuru, Rowe, O'Reilly, & McNamara, 2008; VanderVeen et al., 2007). Indeed, there is direct evidence that commonly used tests of reading comprehension are not tapping into the same cognitive processes (Kendeou & Papadopoulos,

Panayiota Kendeou, Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada; Paul van den Broek, Department of Pedagogy, Leiden University, Leiden, the Netherlands; Mary Jane White, Department of Psychology, University of Memphis; Julie S. Lynch, Department of Psychology, Saginaw Valley State University.

This project was supported by grants from the Center for the Improvement of Early Reading Achievement (CIERA) at the University of Michigan School of Education, the Center for Cognitive Sciences at the University of Minnesota, and the U.S. National Institute of Child Health and Human Development (Grant No. HD-07151). Writing of this article was supported by a Fonds Quebecois de la Recherche sur la Societe et la Culture (FQRSC) Grant No. 2009-NP-125707 to Panayiota Kendeou and a Golestan and Lorentz fellowship from the Netherlands Institute for Advanced Study to Paul van den Broek. We are grateful to Danielle S. McNamara for her comments on the article.

Correspondence concerning this article should be addressed to Panayiota Kendeou, Department of Educational and Counselling Psychology, 3700 McTavish Avenue, Montreal, Quebec H3A 1Y2, Canada. E-mail: panayiota.kendeou@mcgill.ca

[1] There are many different terms in the literature for oral language skills (e.g., oral comprehension, language comprehension). We used the term *oral language skills* in the present study because we included receptive vocabulary in addition to narrative comprehension when estimating the construct.

2009; RAND Reading Study Group, 2002), and as a result, there is a renewed interest in the development of theory-based assessments of comprehension (Francis et al., 2006; Francis, Fletcher, Catts, & Tomblin, 2005).

Another issue concerns the age range on which existing studies have focused. With only a few exceptions, the majority of the studies assessing oral language and decoding-related skills and their usefulness as predictors of reading comprehension have investigated children in kindergarten through second grade (for a review, see Storch & Whitehurst, 2002). Many of these skills, however, develop well before kindergarten, so research focusing on younger children is needed to provide a full understanding of the developmental trajectories of certain skills. A final issue is that the majority of these studies have been cross sectional. Longitudinal investigations would more accurately reflect developmental patterns.

In the present study, we explored the relative contributions and interrelations of oral language and decoding skills in reading comprehension, while addressing the aforementioned issues by (a) using a theoretical framework to guide our assessments, (b) starting with preschool children, and (c) adopting a cross-sequential longitudinal design. Specifically, we conceptualized and measured oral language and reading comprehension skills within the framework of the causal network theory (CNT; see Goldman & Varnhagen, 1986; Graesser & Clark, 1985; Stein & Glenn, 1979; Trabasso, Secco, & van den Broek, 1984) and followed longitudinally two cohorts of children in preschool and kindergarten over the course of 2 years (i.e., a cross-sequential longitudinal design). This design allowed for longitudinal yet time-efficient data collection. More important, it allowed us to compare consecutive age groups both across the same materials (the cross-sectional component) and within the same individuals across time (the longitudinal component).

Assessing Comprehension Within the CNT Framework

Although researchers and educators use the term *comprehension* in different ways, there is considerable agreement that central to comprehension in the context of reading is the construction of a coherent mental representation of the text. In this representation, the reader successfully connects statements and ideas from the text. This mental representation is based on both the text itself and the readers' background knowledge (e.g., Applebee, 1978; Gernsbacher, 1990; Graesser, Singer, & Trabasso, 1994; Kintsch & van Dijk, 1978; Mandler & Johnson, 1977; Oakhill & Cain, 2007; Pearson & Hamm, 2005; Stein & Glenn, 1979; Trabasso et al., 1984; van den Broek, 1994). Thus, comprehension depends on knowledge that cannot always be found in a single word or sentence (Whitehurst & Lonigan, 1998) or even in the text proper. In addition, comprehension is not a unitary phenomenon but rather a family of skills that develop simultaneously (Cutting & Scarborough, 2006; Duke, 2005; van den Broek et al., 2005; Vellutino et al., 2007). This family of skills includes higher order processes—such as inference generation and reasoning—that enable readers to identify meaningful relations among text elements and between text elements and background knowledge.

Events in a text can be related in many ways, but one particularly important type of connection is causal (Graesser et al., 1994; van den Broek, 1997). When readers engage in making inferences and generate different types of causal connections to interconnect the events of the narratives they read, they form a mental network representation of the narrative. This mental network is based on the causal structure of the narrative itself. Consider, for example, the excerpt from the story *The Cat's Purr* in Figure 1. In the accompanying network, each clause in the story is represented in the network as a circle with the corresponding number. The arrows between the circles represent the causal connections between the clauses that the reader may identify. For instance, when reading the story, the reader may make the connection that the fact that Rat wanted to copy Cat (Sentence 4) causes several actions such as Rat's building a house like the house that Cat built (Sentence 5) and planting a tree (Sentence 7).

Such networks have been found to capture important aspects about adults' comprehension of stories (e.g., Goldman & Varnhagen, 1986; O'Brien & Myers, 1987; Trabasso & van den Broek, 1985). Adult readers are more likely to remember events with many causal connections than events with few connections. For example, readers are more likely to remember that Rat liked to copy Cat (Sentence 4; a statement that has five connections) than that Cat let Rat play a tune (Sentence 13; a statement that has one connection). In addition to better remembering events with many connections, readers tend to include these events in summaries more often and rate them as more important than events with only a few connections (van den Broek, 1986; van den Broek, Lorch, & Thurlow, 1996). In addition, when readers are reminded of one event in the text, they remember related events more quickly than unrelated events, even when the latter are closer together in the original text (e.g., O'Brien & Myers. 1987; van den Broek & Lorch, 1993). Finally, when readers are asked questions about why

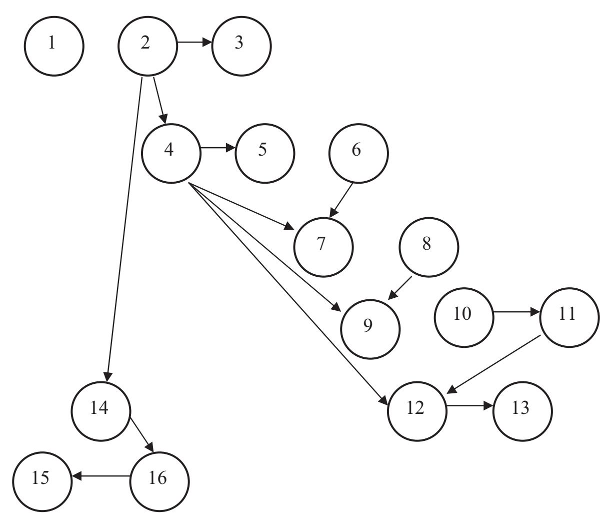

Figure 1. Excerpt from *The Cat's Purr,* the aural story used in Year 1: 1→ Once upon a time, 2→ Cat and Rat were best friends. 3→ They lived in houses right next to each other. 4→ Rat liked to copy Cat. 5→ Rat built a house that was just like Cat's. 6→ Cat planted a tree by his house. 7→ Rat planted one too. 8→ Cat made a straw mat for his house, 9→ and Rat made one too. 10→ Once, Cat made a flute 11→ and played sweet tunes. 12→ "Let me play a tune, too," said Rat. 13→ So, Cat let Rat play a tune. 14→ Cat and Rat also worked together in their vegetable garden. 15→ They planned to have a big party for their friends 16→ when all of the vegetables were ready.

an event happened, they respond with answers that are based on the causal connections in the text (Graesser & Clark, 1985; Trabasso, van den Broek, & Liu, 1988). For example, when asked why Rat built a house that was just like Cat's (Sentence 5), readers are very likely to answer that this was because Rat liked to copy Cat (Sentence 4).

With age, children's sensitivity to the causal structure of narratives as well as their ability to generate causal inferences necessary to connect ideas in the text and to form a coherent text representation develop (Ackerman, 1988; Casteel, 1993; Trabasso & Nickels, 1992; van den Broek, 1988, 1989a, 1990). At an early age, nonstructural properties of events, including superficial ones such as the event's vividness, influence young children's attention more so than do causal structure properties. At later ages, the role of structural properties increases and that of nonstructural factors decreases (Bourg, Bauer, & van den Broek, 1997; van den Broek, 1997). Young children also have a tendency to focus much of their attention on observable, concrete actions rather than on internal causes such as a character's goals. As comprehension skills improve, children's understanding of motives attributable to a character's action, event, or state improves (Goldman & Varnhagen, 1986; Mandler, 1984; Stein & Glenn, 1979; van den Broek, 1989a). Finally, children's tendency to limit their connection building to events *within* each episode gradually diminishes and building possible connections *between* episodes increases. These latter connections are most likely to be related to the overall theme or message of the text and, hence, are central to obtaining a complete picture of the meaning of the narrative as a whole (Brown & Smiley, 1978; Goldman & Varnhagen, 1986; Trabasso & Nickels, 1992; Williams, 1993).

Observations about the development of children's comprehension have primarily involved elementary school children. Research on children's comprehension processes before elementary school is more limited. Because preschool children cannot read, assessment of their comprehension necessarily involves a nonreading context in which stories are presented in media other than text. For example, stories can be presented using pictures (e.g., Paris & Paris, 2003), aurally (e.g., Stein & Glenn, 1979; Trabasso et al., 1984), or via television (e.g., Lorch, Bellack, & Augsbach, 1987; van den Broek et al., 1996). Studies on event comprehension have revealed several developmental patterns from infancy to adulthood and have shown that even young children are sensitive to the relations that exist between different events and particularly causal relations (Mandler & Johnson, 1977; Stein & Glenn, 1979). In fact, children as young as 2 years old can identify different types of causal relations in event sequences (Bauer, 1996, 1997; Wenner & Bauer, 2001). These results suggest that although there are clear developmental differences, even very young children, when comprehending events, use cognitive processes that are similar to those older children and adults use: They identify connections and generate inferences (Lorch & Sanchez, 1997; Mares, 2006; Trabasso & Nickels, 1992; van den Broek et al., 1996). Moreover, these results also suggest that similar processes contribute to comprehension of narratives across different media (also see Kendeou, van den Broek, White, & Lynch, 2007; Lynch et al., 2008; van den Broek et al., 2005). Thus, using different media provides a unique opportunity to assess comprehension independently of decoding skills and before children begin formal instruction.

The Present Study

The fact that coherence-focused comprehension skills already are present and developing in preschool children raises important questions about the developmental trajectory of oral language skills in individual children. One set of questions concerns whether oral language and decoding skills develop according to independent and unique tracks or whether they are interdependent. The results of several recent studies have shown that within an age group, the correlations between decoding and oral language skills tend to be low, suggesting that at least within age groups, there is no direct relation between the two sets of skills at these ages (Cain, Oakhill, & Bryant, 2004; Catts et al., 1999; Gough & Tunmer, 1986; Kendeou, Savage, & van den Broek, in press; Paris & Paris, 2003; Savage, 2006). There is also evidence, though, that oral language and decoding skills are highly related, especially in early years (Storch & Whitehurst, 2002). Another set of questions concerns the stability and predictive power of oral language skills as a child develops: Are individual differences relatively stable across time, and, hence, are skills at an early age predictive of those at a later age? Moreover, are oral language skills predictive of reading comprehension at a later age?

To answer these questions, we examined oral language and decoding skills of two cohorts of children, 4- and 6-year-olds, and retested them when they were 6- and 8-year-olds, respectively. We have taken three steps in our analyses. First, we examined the relation between the development of children's oral language skills and decoding skills from preschool to early elementary school. To do so, we used structural equation modeling to test two models (one for each cohort) derived from the literature (Storch & Whitehurst, 2002). In these models, oral language and decoding skills are connected during preschool (age 4) and kindergarten (age 6). Because oral language skills precede the development of decoding skills, and therefore, influence their development, the link in the structural equation model flows from oral language skills to decoding skills. Second, in these models, we examined the stability and predictive power of children's oral language skills, namely whether oral language skills at an early age predicted those at a later age and whether oral language skills were predictive of reading comprehension at a later age. Third, we tested a number of alternative, theoretically driven models to provide empirical support for the appropriateness of our hypothetical conceptual models as well as the narrative comprehension measures we used to assess oral language skills.

Method

Participants

Two hundred ninety-seven children participated at two test points, 2 years apart. Complete data were collected from 113 children in the 4- to-6-year-old cohort (mean age at the first test point = 4 years, 6 months; range = from 4 years, 0 months to 4 years, 11 months) and 108 children in the 6- to-8-year-old cohort (mean age at the first time point = 6 years, 4 months; range = from 6 years, 0 months to 6 years, 11 months). At both test points, the 6-year-old children were in or had recently completed kindergarten. At the second test point, the 8-year-old children were in or had recently completed second grade. All participants were from a large city in the upper Midwest, and their parents traditionally collaborated in research projects at the university at which the study was conducted. The large majority (96%) were White. Consequently, this sample was very homogeneous.

Materials

Oral Language Skills Measures

At each time point, two narratives were used as part of the narrative comprehension assessment (listening and television) and the Peabody Picture Vocabulary Test–III (PPVT–III; Dunn & Dunn, 1997) was used to assess vocabulary.

Listening comprehension. For the listening comprehension task, one narrative was presented aurally (on an audiotape) to the children at each time point. At Time 1, the story, *The Cat's Purr*, was a 7-min narrative. At Time 2, the story, *The Rabbit and the Moon*, was a 10-min narrative. Both narratives were based on American folk tales. Simple, line-drawn pictures were made to accompany each story. The pictures alone did not convey any major points from the plots. Both stories had a standard but complex structure in which the protagonists made several attempts to achieve their desired goals.

Television comprehension. For the television comprehension task, one narrative was presented audiovisually (on a 26-inch [66.04-cm] color television) to the children at each time point. At Time 1, the audiovisual story, *Autumn Leaves*, was a 12-min episode from an American children's television series, *The Rugrats*. At Time 2, the audiovisual story, *Granny's New Glasses*, was an 18-min episode from an Australian children's television series, *The Adventures of Blinky Bill*. Like the stories in the listening comprehension tests, both stories had a standard but complex structure in which the protagonists made several attempts to achieve their desired goals.

Vocabulary. The PPVT–III (Dunn & Dunn, 1997) was administered as a standardized measure of receptive vocabulary for standard English at both time points. The PPVT–III was selected because it allows students who are nonverbal to participate by pointing to a response. Words are orally given to the respondent by the examiner. The respondent points to a picture that best corresponds to the word. The total raw score is obtained by subtracting the number of errors from the numerical value of the ceiling item (highest word correctly identified). For the ages included in this study, internal consistency, as measured with coefficient alpha, ranged from .93 to .95. Split-half coefficients ranged from .86 to .95.

Decoding Skills Measures

Letter and word identification subtests from Woodcock Reading Mastery Test–Revised (Woodcock, 1987), and the Onset Recognition Fluency (OnRF) measure from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002a, 2002b) were used to assess decoding skills. All measures were administered at Time 1. At Time 2, only word identification was administered as ceiling effects were expected for the other two measures.

Letter identification. The Letter Identification Subtest from the Woodcock Reading Mastery Test–Revised (Woodcock, 1987) was administered. This test was selected because it provides information about a child's ability to identify different letters. In the test, the letters are presented to subjects, and each subject is asked to verbally identify the letter within 5 s. The test begins at an age-appropriate item (basal level) and ends when the child answers six or more consecutive items incorrectly or when the last page of the test has been administered. The total raw score is obtained by subtracting the number of errors from the numerical value of the ceiling item (the last letter or word correctly identified). For the ages included in this study, internal consistency, as measured with split-half coefficients, ranged from .84 to .94.

Word identification. The Word Identification Subtest from the Woodcock Reading Mastery Test–Revised (Woodcock, 1987) was administered. This test was selected because it provides information about a child's ability to read different words. In the test, the words are presented to subjects, and each subject is asked to verbally identify the word within 5 s. It is not assumed that the subject knows the meaning of any word that is correctly identified. The test begins at an age-appropriate item (basal level) and ends when the child answers six or more consecutive items incorrectly or when the last page of the test has been administered. The total raw score is obtained by subtracting the number of errors from the numerical value of the ceiling item (the last letter or word correctly identified). For the ages included in this study, internal consistency, as measured with split-half coefficients, ranged from .97 to .98.

Phonological awareness. The Onset Recognition Fluency (OnRF) measure from the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002a, 2002b) was used to assess children's phonological awareness. The examiner presents four pictures to the child, names each picture, and then asks the child to identify (i.e., point to or say) the picture that begins with the sound produced orally by the examiner. For example, the examiner says, "This is sink, cat, gloves, and hat. Which picture begins with /s/?" and the student points to the correct picture. The child is also asked to orally produce the beginning sound for an orally presented word that matches one of the given pictures. The examiner calculates the total number of initial sounds produced correctly in a minute. For the ages included in this study, alternate-form reliability of the OnRF measure is .91.

Reading Comprehension Measure

A written narrative was used to assess children's reading comprehension at Time 2 (8-year-olds only). The story, *The Barber's Wife,* was based on an Indian folk tale. The story had age-appropriate vocabulary and had a standard but complex structure in which the protagonists made several attempts to achieve their desired goal.

Procedure

Time 1

The children were tested individually in a single session by one of three female experimenters. The entire session was videotaped, and the comprehension tests were also audiotaped. After the children had been given a period of time to become comfortable with the experimental setting; they completed the phonological awareness and vocabulary assessments. The children then listened to the aural story, *The Cat's Purr*. The children were instructed to listen closely so they could answer questions after the story was over. While listening to the story, the children had the pictures that accompanied the story available. Immediately after the story was completed, the experimenters asked the children to "tell everything you remember about the story from the beginning." If a child did not recall any narrative events spontaneously, the experimenter asked a more specific question, "What happened at the beginning of the story?" The children were prompted to continue to recall the story (i.e., "What else do you remember?"), until they indicated that they could not recall anything else. Because young children often have difficulty spontaneously reporting their memory for narratives, slightly more specific questions were then asked. For each episode in the narrative (determined by procedures described by Mandler & Johnson, 1977; Stein & Glenn 1979) that a child remembered spontaneously, the experimenter asked a question in the following form: "You remembered X. What happened before X?" and "What happened after X?" These questions were asked only if the child did not recall the episodes that were the focus of the questions.

After a short break, the children completed the letter and word identification tests. The children then viewed the audiovisual narrative, *Autumn Leaves*. The procedure for assessing comprehension of the audiovisual narrative was identical to that of the aural narrative.

Time 2

The children were tested individually in a single session by one of three experimenters (two women and one man). All procedures were identical to those at Time 1, except for the exclusion of the phonological awareness and letter identification tests and the inclusion of the reading comprehension task for the 8-year-old children. The reading comprehension assessment took place after the children completed the word identification task. The procedure for assessing reading comprehension was parallel to that of comprehension of the aural and audiovisual narratives. To allow for longitudinal comparisons, we kept the order of the different tests and stories identical across both time points.

Coding

Prior to data collection in each year, three researchers analyzed the narratives and parsed them into individual events (generally subject–verb phrases). At Time 1, the aurally presented story had 167 events and the audiovisual narrative had 231 events. At Time 2, the aurally presented story had 158 events, the audiovisual narrative had 357 events, and the written text had 99 events. We determined the causal structure of each narrative by identifying the causal relations between all events in the story according to principles of causality (Mackie, 1980; Trabasso, van den Broek, & Suh, 1989).

The selection of stories ensured that they did not differ greatly in length across Times 1 and 2. Furthermore, the causal analysis of all stories showed that there were approximately an equal proportion of highly connected events (i.e., events with four or more connections) in each one. Specifically, the aurally presented story had 31% highly connected events in Time 1 and 30% in Time 2. The audiovisual narrative had 40% highly connected events in Time 1 and 38% in Time 2. The written text had 32% highly connected events.

The children's responses during the free recall after each story presentation (aural, audiovisual, and written) were transcribed verbatim from the videotapes and audiotapes of the experimental sessions. The children's responses were parsed into events, analogously to the parsing of the original narratives. Each recalled event was coded according to a gist criterion to the event that it most closely matched in the corresponding narrative. Recalled events that did not match an event in the story were coded separately and were not included in the following analyses. These events ranged between 11% and 15% of total events recalled across stories and age groups.

Two raters coded the transcripts. For results at Time 1, 20% of the transcripts were coded by both raters to establish and practice the coding scheme. An additional 20% of the transcripts were coded by both raters to determine interrater reliability. Interrater agreement was .62, $p < .01$ (calculated on the basis of each recalled events as an agreement or disagreement in coding). The same coding procedures were followed for the results at Time 2. Here, interrater agreement also was .62, $p < .01$.

For each narrative, a measure of children's sensitivity to the causal structure was calculated. This measure consisted of the total number of unique, highly connected story events (i.e., events with four or more causal connections to other events in the story) the children included in their free recall protocol (Graesser & Clark, 1985; Trabasso et al., 1989; van den Broek et al., 1996; van den Broek, 1989b). This measure comprised the listening, television, and reading comprehension variables.

Results

Preliminary Analyses

Preliminary inspection of the data with skewness and kurtosis indices did not suggest major deviations from normality. The means and standard deviations of all measures used are presented in Table 1.

Structural Equation Modeling

We used structural equation modeling (SEM) to explore the relations between oral language and decoding skills across development. Separate models were fitted for each cohort. In the evaluation of the goodness of fit of each model to the data, we report the model chi-square statistic associated with the *p* value, the comparative fit index (CFI), the nonnormed fit index (NNFI), the Akaike's information criterion (AIC), and the root-mean-square error of approximation (RMSEA). A nonsignificant value of the chi-square statistic indicates a good fit; however, the test is sensitive to sample size and should be considered in relation to its degrees of freedom (i.e., dividing chi-square value by its degrees of freedom should result in a value below 2, indicating a good model fit; Maruyama, 1998). CFI and NNFI indices equal to or superior to .95 are considered to indicate a good fit (Hu & Bentler, 1999). The AIC measure indicates a better fit when it is smaller (Browne & Cudeck, 1992). Finally, the RMSEA is an absolute fit

Table 1
Descriptive Statistics of Assessment Measures

Measure	Preschool (4 years old) *M*	*SD*	*Min*	*Max*	Kindergarten (6 years old) *M*	*SD*	*Min*	*Max*	Second grade (8 years old) *M*	*SD*	*Min*	*Max*
Cohort 1												
Oral language skills												
Listening comp	4.08	2.83	0	16	5.43	3.75	0	20				
Television comp	6.54	4.17	0	18	7.44	4.44	0	28				
Decoding skills												
Letter ID	17.01	10.82	0	38	36.01	4.17	21	46				
Word ID	0.82	5.07	0	40	24.68	19.16	0	81				
DIBELS	8.76	3.97	0	16	15.32	1.20	9	16				
PPVT–III	73.81	16.09	9	114	105.23	13.74	75	147				
Cohort 2												
Oral language skills												
Listening comp					9.62	5.78	0	28	10.26	5.08	1	30
Television comp					10.79	6.26	1	36	13.13	7.01	2	37
Decoding skills												
Letter ID					34.59	5.48	4	43	—	—	—	—
Word ID					18.51	20.28	0	76	65.10	10.53	36	93
DIBELS					14.71	1.88	4	16	—	—	—	—
PPVT–III					105.01	12.95	77	143	128.76	14.91	93	167
Reading comp					—	—	—	—	13.29	5.99	0	24

Note. Raw scores are included. Comp = comprehension; ID = identification; DIBELS = Dynamic Indicators of Basic Early Literacy Skills; PPVT–III = Picture Vocabulary Test (3rd ed.).

index in which the complexity of the model is considered; values less than .05 are considered a good fit (Cudeck & Browne, 1992).

Fitting the Model for the First Cohort (Ages 4–6)

For the first cohort of children at both testing times (at ages 4 and 6), indicators for the decoding skills latent variable were phonological awareness, letter identification, and word identification, whereas indicators for the oral language skills were listening comprehension, television comprehension, and vocabulary. Figure 2 depicts the fitted model for the first cohort of children with standardized parameter estimates. We hypothesized that oral language and decoding skills relate to each other within a year. Because the development of oral language skills precedes that of decoding skills, we hypothesized that in preschool and in kindergarten, oral language skills predict decoding skills. We also hypothesized longitudinal continuity, in that preschool oral language skills predict kindergarten oral language skills and preschool decoding skills predict kindergarten decoding skills. As indicated by the fit indices, $\chi^2(49, N = 113) = 56.44$, $p = .22$ ($\chi^2/df = 1.15$); CFI = .98; NNFI = .97; AIC = 114.44; RMSEA = .04 (90% confidence interval [CI]: 0.0, 0.07), the model yielded a good fit to the data. The relative magnitude and significance of the standardized coefficients show that the relation between decoding and oral language skills is strong in preschool (with oral language skills predicting 28% of the variance in de-

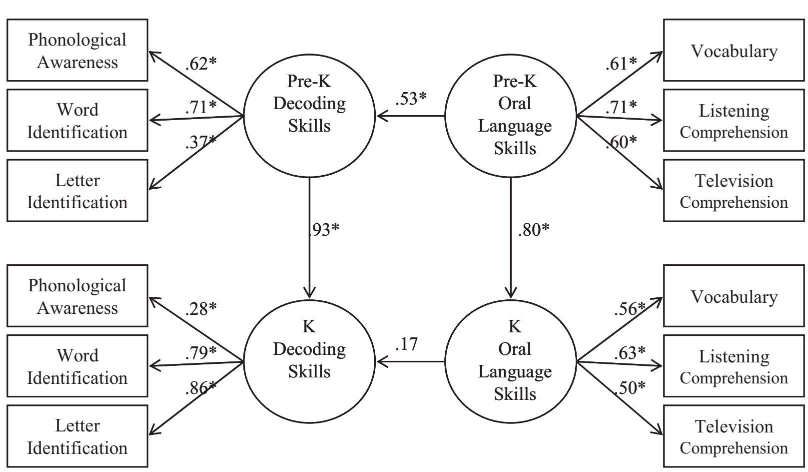

Figure 2. Model for Cohort 1 (ages 4–6) depicting relations between decoding and oral language skills.

coding skills) and very weak (not significant) in kindergarten. Furthermore, there is longitudinal continuity within both oral language and decoding skills. Specifically, for oral language skills, approximately 64% of the variance in kindergarten children's ability was explained by their ability 2 years earlier in preschool. Likewise, for decoding skills, approximately 75% of the variance in kindergarten children's ability was explained by their ability 2 years earlier in preschool.

Fitting the Model for the Second Cohort (Ages 6–8)

For the second cohort at Time 1 (age 6), indicators for the decoding skills latent variable were letter identification, word identification, and phonological awareness. Indicators for oral language skills were listening comprehension, television comprehension, and vocabulary. At Time 2 (age 8), decoding skills were estimated by a single variable, word identification; indicators for the oral language latent variable were listening comprehension, television comprehension, and vocabulary. Reading comprehension also was estimated by a single variable. Figure 3 depicts the fitted model for the second cohort of children with standardized parameter estimates. Consistent with the model fitted for the first cohort, we hypothesized that kindergarten and second grade oral language skills precede and, therefore, predict decoding skills. We also hypothesized longitudinal continuity in that kindergarten oral language skills predict second grade oral language skills and kindergarten decoding skills predict second grade decoding skills. In addition, we hypothesized that both oral language and decoding skills in second grade independently predict reading comprehension in second grade. As indicated by the fit indices, $\chi^2(36, N = 108) = 38.03$, $p = .38$ ($\chi^2/df = 1.06$); CFI = .99; NNFI = .99; AIC = 98.03; RMSEA = .02 (90% CI: 0.0, 0.07), the model yielded a good fit to the data.

The relative magnitude and significance of the standardized coefficients show that the relation between decoding and oral language skills is very weak and not significant in either kindergarten or second grade. Furthermore, there is longitudinal continuity within decoding skills; approximately 63% of the variance in a child's ability in second grade was accounted for by his or her ability in kindergarten. Longitudinal continuity within the oral language skills is rather weak; approximately only 3% of the variance in a second grade child's ability was significantly explained by his or her ability in kindergarten. Of note, reading comprehension in second grade (age 8) is influenced by both decoding (i.e., word identification) and oral language skills. Specifically, approximately 47% of the variance in reading comprehension is explained by oral language and decoding skills.

Testing Alternative Models

The hypothesized models fit the data well, as suggested by the goodness-of-fit measures. However, it is possible that our data support other models that are also theoretically meaningful. To explore this hypothesis, we tested a series of alternative models for each cohort and compared them with the models in Figures 2 and 3. When the alternative and original models were nested (i.e., they included the same number of parameters), we evaluated their overall fits based on traditional fit indices and compared them on the basis of overall fit, the discrepancy/degrees of freedom ratio (χ^2/df), the AIC, and a chi-square test. The discrepancy/degrees of freedom ratio of 2 or less indicates a close fit, whereas the AIC measure indicates a better fit when it is smaller (Browne & Cudeck, 1992). With the chi-square test, we tested the null hypothesis of no significant difference in fit by evaluating whether the chi-square difference between the two models is significant for the given degrees of freedom and a chosen significance level. If the

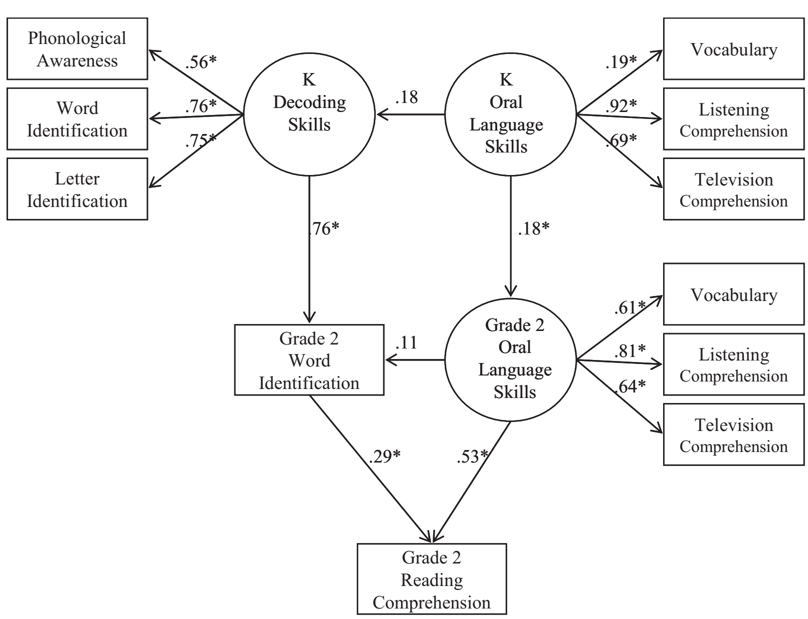

Figure 3. Model for Cohort 2 (ages 6–8) depicting relations between decoding skills, oral language skills, and reading comprehension. *Note.* Values are standardized coefficients. $^* p < .05$.

difference is significant, then the null hypothesis is rejected, and significant differences between the models are indicated. When the alternative and original models were not nested, we evaluated overall fits based on traditional fit indices and compared them on the basis of overall fit, the discrepancy/degrees of freedom ratio (χ^2/df), and the AIC.

The first alternative model that we tested for the first cohort differed from the original model in Figure 2 in that it eliminates the within-age-group relations between decoding and oral language skills in each year. Thus, this model assumes complete independence between decoding and oral language skills (Figure 4, Model A). The overall fit of this model was not acceptable, $\chi^2(51, N = 113) = 70.67$, $p = .03$ ($\chi^2/df = 1.38$); CFI = .94; NNFI = .92; AIC = 124.67; RMSEA = .06 (90% CI: 0.02, 0.09). This fit is significantly weaker than that of the original model, $\Delta\chi^2(2) = 14.24$, $p < .05$. Following the same rationale, we tested the same alternative model (i.e., without the within-age-group relations between decoding and oral language skills) for the second cohort (Figure 4, Model B). The overall fit of this model was not acceptable, $\chi^2(40, N = 108) = 66.34$, $p = .004$ ($\chi^2/df = 1.66$); CFI = .92; NNFI = .89; AIC = 128.34; RMSEA = .08 (90% CI: 0.0, 0.08). This fit is significantly weaker than that of the original model, $\Delta\chi^2(4) = 28.31$, $p < .05$. In summary, this alternative model significantly decreased in fit for both cohorts when compared with the original models depicted in Figures 2 and 3. On this ground, we accepted the original models.

The second alternative model that we tested for the first cohort differed from the original model in that it includes longitudinal relations between decoding and oral language skills. In this model, we assumed relations between decoding and oral language skills within and across years (Figure 4, Model C). The overall fit of this model was good, $\chi^2(47, N = 113) = 55.77$, $p = .18$ ($\chi^2/df = 1.18$); CFI = .97; NNFI = .96; AIC = 117.77; RMSEA = .04 (90% CI: 0.0, 0.08). In this model, the added path from prekindergarten oral language to kindergarten decoding skills was significant (.22, $p < .05$), whereas the path from prekindergarten decoding skills to kindergarten oral skills was not significant (.10, $p > .05$). The chi-square test suggested that the two models were not significantly different, $\Delta\chi^2(2) = .66$, $p > .05$; however, the original model had smaller RMSEA, AIC, and χ^2/df. Following the same rationale, we tested this alternative model—that is, with the addition of the longitudinal relations between decoding and oral language skills—for the second cohort (Figure 4, Model D). The overall fit of this model also was good, $\chi^2(34, N = 108) = 35.63$, $p = .39$($\chi^2/df = 1.05$); CFI = .99; NNFI = .99; AIC = 99.63; RMSEA = .02 (90% CI: 0.0, 0.07). Both paths added in this model, from kindergarten oral language to second grade decoding skills (.14, $p > .05$) and from kindergarten decoding skills to second grade oral skills (.08, $p > .05$), were not significant. The chi-square test suggested that the two models were not significantly different, $\Delta\chi^2(2) = 2.40$, $p > .05$; however, the original model had smaller AIC and χ^2/df. In summary, for both cohorts this alternative model and the original model had a good fit to the data. However, the original model is less complex than the alternative model. The principle of parsimony suggests adopting the model with fewer estimated parameters (Marsh & Hau, 1996), namely the original model for each cohort depicted in Figures 2 and 3.

The third alternative model that we tested for the first cohort eliminated the within-age-group relations between decoding and oral language skills and added longitudinal relations. In this model, we assumed independence between decoding and oral skills within years but not across years (Figure 4, Model E). The overall fit of this model was not acceptable, $\chi^2(49, N = 113) = 68.29$, $p = .03$ (ratio $\chi^2/df = 1.40$); CFI = .94; NNFI = .92; AIC = 126.29; RMSEA = .06 (90% CIs: 0.10, 0.09). This fit is weaker than that of the original model. Following the same rationale, we tested this alternative model for the second cohort (Figure 4, Model F). The overall fit of this model was good, $\chi^2(36, N = 108) = 41.28$, $p = .25$ (ratio $\chi^2/df = 1.15$); CFI = .97; NNFI = .98; AIC = 101.28; RMSEA = .04(90% CI: 0.0, 0.08); however, the original model had smaller RMSEA, AIC, and χ^2/df. In summary, this alternative model significantly decreased in fit for both cohorts when compared with the original model. On this ground, we accepted the original model for each cohort depicted in Figures 2 and 3.

Testing Sensitivity to the Causal Structure as a Comprehension Measure

On the basis of strong evidence in the literature, we used sensitivity to causal structure as the primary, theoretically derived way of assessing children's comprehension of stories and of tracking their development. Accordingly, in assessing children's comprehension, we used recall of highly connected events in each story and not simply overall recall. The assumption here was that recall of highly connected events is a better indicator of children's comprehension than overall recall, even though the two measures often are highly interrelated. Indeed, correlations in both cohorts ranged from .89 to .97. To test this hypothesis, for each cohort we formulated an alternative model that included overall recall as indicators for listening and for television and reading comprehension and compared it with the original model in Figures 2 and 3. The alternative and original models were evaluated on the basis of traditional fit indices and compared on the basis of overall fit, the discrepancy/degrees of freedom ratio (χ^2/df), and the AIC.

The alternative model to the original model in Figure 2 for the first cohort included overall recall of the listening and the television comprehension in addition to vocabulary as indicators for the oral skills in prekindergarten and kindergarten (instead of the highly connected recall). The overall fit of this model was good, $\chi^2(49, N = 113) = 62.12$, $p = .09$ ($\chi^2/df = 1.27$); CFI = .96; NNFI = .95; AIC = 120.12; RMSEA = .05 (90% CI: 0.0, 0.08), but weaker than that of the original model; the original model had smaller RMSEA, .04, $\chi^2/df = 1.15$, and AIC = 68.44. The alternative model to the original model in Figure 3 for the second cohort included overall recall of the listening and television and reading comprehension as indicators (instead of the highly connected recall). The overall fit of this model was good, $\chi^2(36, N = 108) = 46.45$, $p = .11$ (ratio $\chi^2/df = 1.29$); CFI = .97; NNFI = .96; AIC = 106.45; RMSEA = .05 (90% CI: 0.0, 0.09), but weaker than that of the original model; the original model had smaller RMSEA, .02, $\chi^2/df = 1.06$, and AIC = 76.03. In summary, the alternative model for each cohort slightly decreased in fit when compared with the original models in Figures 2 and 3. On this ground, we accepted the original model in Figures 2 and 3.

Alternative Models Tested for Cohort 1 (Left) and Cohort 2 (Right)

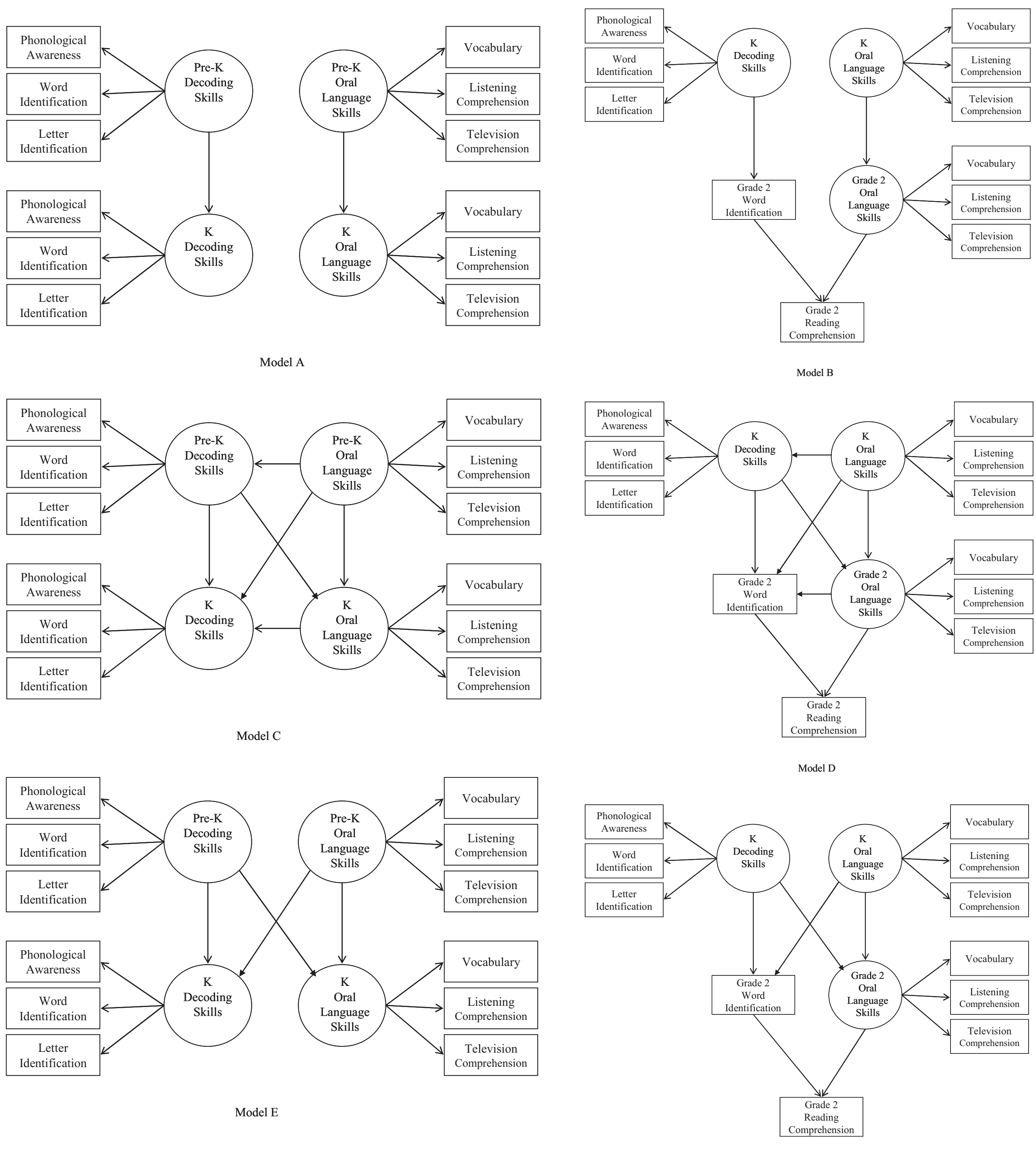

Figure 4. Alternative models tested for Cohort 1 (left) and Cohort 2 (right).

Cohort Comparison

In this study, data were collected twice from 6-year-old children. The 6-year-old children in the first cohort (at Time 2) and the 6-year-old children in the second cohort (at Time 1) were administered the same decoding and vocabulary measures but different aural and television narratives. To evaluate the degree to which the two cohorts of children differed in their performance, we directly compared their scores on the three decoding and three oral skill measures. To do so, we conducted a multivariate analysis of variance with cohort as the independent variable and phonological awareness, letter identification, word identification, listening comprehension, television comprehension, and vocabulary as the dependent variables. This analysis revealed significant cohort differences, $F(6, 214) = 10.99$, $p = .0001$, $\eta^2 = .24$. First, with respect to decoding skills, the first cohort (6 years old at Time 2) performed significantly higher than the second cohort (6 years old at Time 1) in phonological awareness, $F(1, 219) = 8.89$, $p = .003$, $\eta^2 = .04$; letter identification, $F(1, 219) = 4.74$, $p = .03$, $\eta^2 = .03$; and word identification $F(1, 219) = 5.39$, $p = .02$, $\eta^2 = .02$ (all means are provided in Table 1). Second, with respect to oral language skills, the first cohort (6 years old at Time 2) performed significantly lower than the second cohort (6 years old at Time 1) in listening comprehension, $F(1, 219) = 41.12$, $p = .0001$, $\eta^2 = .16$, and television comprehension, $F(1, 219) = 21.21$, $p = .0001$, $\eta^2 = .09$. There were no differences between cohorts in vocabulary, $F(1, 219) < 1$, $p > .05$.

These cohort differences suggest that the first cohort performed higher on decoding skills than the second cohort, whereas the second cohort performed higher on oral language skills than the first cohort. Although these cohort differences emerged, the models for the two cohorts showed remarkable consistency with respect to the relations between decoding and oral language skills. In both cohorts at age 6, the standardized coefficients linking oral language and decoding skills were not significant and were of comparable magnitude (.17 and .18 for Cohorts 1 and 2, respectively).

Discussion

The present study had two goals. The first goal was to examine the relation between the development of children's oral language skills and decoding skills from preschool to early elementary school. The second goal was to examine the stability and predictive power of children's oral language skills for later reading comprehension to determine whether oral language skills at an early age predicted those at a later age and whether oral language skills were predictive of reading comprehension at a later age in early elementary school. To achieve these goals, we examined oral language and decoding skills of two cohorts of children, 4 and 6 years old, and retested them when they were 6 and 8 years old, respectively.

The findings show, first, that within each age group, oral language skills and decoding skills formed distinct clusters and that these clusters showed longitudinal continuity. With regard to the latter finding, SEM showed that oral language skills at one age uniquely predicted oral language skills 2 years later for both cohorts of children (i.e., from age 4 to age 6, and from age 6 to age 8) and, likewise, that decoding skills at an early age uniquely predicted decoding skills 2 years later. Second, the findings showed that the two sets of skills were strongly interrelated in preschool and that with development this relation became weaker, both from preschool to kindergarten and from kindergarten to second grade. Specifically, SEM showed that in preschool, oral language skills predicted decoding skills but that this pattern was weaker in kindergarten and second grade. Third, oral language and decoding skills each independently predicted a child's reading comprehension in second grade.

These conclusions are based on the original model fitted for each cohort hypothesizing that oral language skills predict decoding skills within each year. To ascertain the validity of this model, we considered several plausible alternative models. One set of models hypothesized complete independence of decoding and oral language skills within and across years (Models A and B for the respective cohorts). Another set of models hypothesized that oral language and decoding skills predict each other within and across years (Models C and D). A third set of models hypothesized that oral language and decoding skills predict each other only across years (Models E and F). These alternative models either had poorer fit to the data or, in the case of Models C and D, had a good fit to the data but were less parsimonious than the original models.

Given that the principle of parsimony is a methodological criterion, it is worthwhile to consider the possible implications of adopting a less parsimonious, theoretically viable model that has good fit to the data. In this model (Models C and D for the respective cohorts), oral language skills not only predict decoding skills *within* a year but oral language skills at a younger age also predict decoding skills at a later age, that is, *across* years. Although these two sets of skills follow their own developmental trajectories across time as indicated by the dissociation between oral language and decoding skills that has been reported for older children (Oakhill, Cain, & Bryant, 2003; Paris & Paris, 2001, 2003; van den Broek, et al., 1996; Vellutino et al., 2007; Whitehurst & Lonigan, 1998), the current findings extend that observation by showing that distinct clusters of oral language skills and decoding skills are present at a much earlier age than previously shown and that they have a reciprocal relation: The development of one influences directly the development of the other and vice versa.

With respect to reading comprehension, the current results confirm a time-honored view showing that the two clusters of skills contribute to a child's reading comprehension activities in early elementary school, with each cluster making a sizable, unique contribution. Thus, the particular requirements of comprehending a text involve both the ability to "break the code" by translating written symbols into meaningful words (Adams, 1990; Ehri, 1999, 2005; Perfetti, 1985; Stanovich, 1986) and the ability to extract meaning about events and facts and identify semantic relations between those events and facts (cf. Ehri, 1998; Graesser et al., 1994; Kintsch, 1988; Ruddell & Ruddell, 1994; van den Broek & Kremer, 1999; Vellutino et al., 2007; Whitehurst & Lonigan, 1998). Our results show that both types of skills begin developing during the preschool years and that these early skills are predictive of reading comprehension in second grade (see also Cain & Oakhill, 2007; Paris & Paris, 2003).

In light of the finding that both oral language and decoding skills predict reading comprehension, one would expect the impact of decoding skills to gradually diminish as elementary school children become more proficient, with increasing amounts of variance in successful reading comprehension being contributed by oral language skills. Indeed, that is what we observed in our study. Oral language skills accounted for more variance in reading comprehension than did decoding skills. Likewise, when the contributions of skills such as these to reading comprehension were assessed in second/third grade (i.e., an age similar to that of the oldest group in our study) and in sixth/seventh grade, the contribution of oral language skills was found to increase across age (Vellutino et al., 2007).

In this study, comprehension was operationalized as children's sensitivity to the causal structure of the narrative. Such operationalization has considerable psychological validity (Graesser & Clark, 1985; Trabasso et al., 1989; van den Broek et al., 1996; van den Broek, 1989a) and reflects the fact that creating coherence is at the center of successful oral and reading comprehension (Graesser et al, 1994; Kintsch, 1988; Oakhill & Cain, 2007; Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). This measure accounts for qualitative as well as quantitative developmental aspects of children's comprehension skills, and our findings suggest that it estimates children's comprehension skills better than measures that are simply focused on overall recall. We suggest that sensitivity to the causal structure provides an alternative to the passage comprehension subsections of standardized tests frequently used in prior studies. Such tests have been designed for students who have mastered decoding skills and are widely criticized as invalid measures of comprehension (e.g., Cutting & Scarborough, 2006; Fletcher, 2006; Hannon & Daneman, 2001; Keenan & Betjemann, 2006).

These findings have important theoretical and practical implications. With respect to theoretical implications, they show that successful reading comprehension depends on decoding skills, such as phonological awareness and letter and word identification, but also on oral language skills such as vocabulary and discourse comprehension. Both sets of skills make unique contributions to reading comprehension when children become conventional readers. These findings extend to the earlier preschool-level findings on school-age children (Catts, Hogan, & Fey, 2003; Nation & Snowling, 1997, 1998; Oakhill et al., 2003; Paris & Paris, 2001, 2003; Savage, 2006; Tunmer & Hoover, 1992), and moreover, they do so by following individual children in a cross-sequential longitudinal design rather than by means of cross-sectional designs. The two sets of skills begin to develop early in children's lives, at first intertwined but gradually developing relatively independently, with each having considerable stability over development. In the early elementary grades, these sets of skills combine to support children's reading comprehension. Thus, instead of a single causal sequence of skill development, the reality is one of multiple causality, with the two different sets of skills both being necessary and neither alone sufficient for later reading comprehension success (van den Broek et al., 2005).

From a practical point of view, these findings have implications for comprehension assessment, intervention, and direct instruction. First, they show that early assessment of comprehension skills is not only possible but also useful because the results predict future comprehension performance in reading contexts, namely when children are engaged in reading activities (Kendeou et al., 2005, 2007; Paris & Stahl, 2005; van den Broek et al., 2005). The overlap between comprehension performances in different media indicates that comprehension assessment in nonreading contexts (e.g., listening and television comprehension) may be used for early identification of students who are likely to experience later difficulties in reading comprehension.

It is important to note that the assessment of comprehension skills focusing on children's sensitivity to the causal structure not only is supported by our data but also accounts for qualitative as well as quantitative developmental aspects of children's comprehension skills. In this assessment, we focused not only on the *number* of the connections in individuals' representations as a function of the story structure but also of the *types* of connections included. Whereas one sign of improved comprehension is that a child identifies more relations in a story, another—and perhaps more telling—sign is that a child has advanced to include complex types of relations (such as causal connections).

Second, with regard to comprehension interventions, the findings suggest that improvement and development of comprehension skills in preschool children may lead to improved comprehension once those children reach reading age. As with assessment, such interventions could rely on nonreading (e.g., televised or listening) contexts. For example, activities around television viewing or listening may provide the opportunity for developing and fostering comprehension strategies even before the beginning of formal instruction or for older, struggling readers who experience decoding difficulties.

Third, with regard to instruction, the dissociation of oral language and decoding skills has direct educational implications as it provides a conceptual framework for designing appropriate teaching practices that target both decoding and comprehension skills (e.g., Aaron, 1991; Kendeou et al., 2005, 2007; McNamara, 2007; Oakhill et al., 2003; Savage, 2001, 2006). For instance, the separation of these skills enables teachers to understand what they need to teach about decoding and comprehension within a broad curriculum. For example, in the United Kingdom, the simple view of reading has been adopted as the theoretical basis of the revised national curricular advice to all schools in England (Kendeou et al., in press; Rose, 2006; Stuart, Stainthorp, & Snowling, 2008; U.K. Department for Education and Skills, 2006).

Successful reading comprehension is the result of a confluence of elemental skills, each of which has its own developmental trajectory. The trajectories may intertwine and influence each other at early stages (e.g., in the 4-year-old children in our study), but they also remain independent to a considerable degree. The risk of a child developing reading comprehension difficulties is smallest when he or she progresses appropriately along each trajectory. The more educational practice can help each child move forward along each dimension, the more it ensures the child against failure. This means that decoding skill development should be part of the curriculum—as it traditionally has—but so should oral language skill development including narrative comprehension.

References

Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? *Journal of Learning Disabilities, 24,* 178–186.

Ackerman, B. P. (1988). Thematic influences on children's judgments about story adequacy. *Child Development, 59,* 918–938.

Adams, M. J. (1990). *Beginning to read.* Cambridge, MA: MIT Press.

Applebee, A. N. (1978). *The child's concept of a story: Ages two to seventeen.* Chicago: University of Chicago Press.

Bauer, P. J. (1996). What do infants recall of their lives? Memory for specific events by 1- to 2-year-olds. *American Psychologist, 51,* 29–41.

Bauer, P. J. (1997). Development of memory in early childhood. In N. Cowan (Ed.), *The Development of Memory in Childhood* (pp. 83–111). Sussex, UK: Psychology Press.

Bishop, D. V. M., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry, 31,* 1027–1050.

Bourg, T., Bauer, P., & van den Broek, P. (1997). Building the bridges: The development of event comprehension and representation. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation: Bridging fictional and actual events* (pp. 385–407). Hillsdale, NJ: Erlbaum.

Brown, A. L., & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development, 49,* 1076–1088.

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

Bryant, P., MacLean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. *Applied Psycholinguistics, 11,* 237–252.

Cain, K., & Oakhill, J. (2007). Reading comprehension difficulties: Correlates, causes, and consequences. In Cain, K. & Oakhill, J. (Eds). *Children's comprehension problems in oral and written language: A cognitive perspective* (pp. 41–75). New York: Guilford Press.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96,* 31–42.

Casteel, M. (1993). Effects of inferences necessity and reading goal on children's inferential integration. *Journal of Educational Psychology, 88,* 484–507.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, B. J. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3,* 331–361.

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *Journal of Learning Disabilities, 36,* 151–164.

Cudeck, R., & Browne, M. W. (1992). Constructing a covariance matrix that yields a specified minimizer and a specified minimum discrepancy function value. *Psychometrika, 57,* 357–369.

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

Duke, N. K. (2005). Comprehension of what for what: Comprehension as a non-unitary construct. In S. Paris & S. Stahl (Eds.), *Current issues in reading comprehension and assessment* (pp. 93–104). Mahwah, NJ: Erlbaum.

Dunn, L. M., & Dunn, L. M. (1997*). Peabody picture vocabulary test (3rd ed.) (PPVT–III).* Circle Pines, MN: American Guidance Services.

Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme & R. M. Yoshi (Eds.), *Reading and spelling: Development and disorders* (pp. 87–111). London: Erlbaum.

Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 79–108). Oxford, UK: Blackwell.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9,* 167–188.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71,* 393–447.

Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading, 10,* 323–330.

Francis, D. J., Fletcher, J. M., Catts, H., & Tomblin, B. (2005). Dimensions affecting the assessment of reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369–394). Mahwah, NJ: Erlbaum.

Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading, 10,* 301–322.

Gernsbacher, M. A. (1990). *Language comprehension as a structure building.* Hillsdale, NJ: Erlbaum.

Goldman, S. R., & Varnhagen, C. K. (1986). Memory for embedded and sequential story structures. *Journal of Memory and Language, 25,* 401–418.

Good, R. H., & Kaminski, R. A. (2002a). *DIBELS oral reading fluency passages for first through third grades.* (Technical Report No. 10). Eugene, OR: University of Oregon.

Good, R. H., & Kaminski, R. A. (2002b). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement. Retrieved April 12, 2006, from http://dibels.uoregon.edu/

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

Graesser, A., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative comprehension. *Psychological Review, 101,* 371–395.

Graesser, A. C., & Clark, L. F. (1985). *The Structures and Procedures of Implicit Knowledge.* Norwood, NJ: Ablex.

Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component processes of reading comprehension. *Journal of Educational Psychology, 93,* 103–128.

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

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the gray oral reading test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading, 10,* 363–380.

Kendeou, P., Lynch, J. S., van den Broek, P., Espin, C. A., White, M. J., & Kremer, K. E. (2005). Developing successful readers: building early comprehension skills through television viewing and listening. *Early Childhood Education Journal, 33,* 91–98.

Kendeou, P., & Papadopoulos, T. C. (2009). *Reading comprehension tests: What skills do they assess?* Manuscript in preparation.

Kendeou, P., Savage, R. S., & van den Broek, P. (in press). Revisiting the simple view of reading. *British Journal of Educational Psychology.*

Kendeou, P., van den Broek, P., White, M., & Lynch, J. (2007). Preschool and early elementary comprehension: Skill development and strategy interventions. In D. S. McNamara (Ed.) *Reading comprehension strategies: Theories, interventions, and technologies,* (pp. 27–45). Mahwah, NJ: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95,* 163–183.

Kintsch, W., & van Dijk, T. A. (1978). Towards a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology, 36,* 596–613.

Lorch, E., Bellack, D., & Augsbach, L. (1987). Young children's memory for televised stories: Effects of importance. *Child Development, 58,* 453–463.

Lorch, E. P., & Sanchez, R. P. (1997). Children's memory for televised events. In P. van den Broek, P. J. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension and representation* (pp. 271–291). Mahwah, NJ: Erlbaum.

Lynch, J. S., van den Broek, P., Kremer, K. E., Kendeou, P., White, M., & Lorch, E. P. (2008). The development of narrative comprehension in its relation to other early reading skills. *Reading Psychology, 29,* 327–365.

Mackie, J. L. (1980). *The cement of the universe: A study of causation.* Oxford, UK: Clarendon.

Magliano, J. P., Millis, K. K., Ozuru, Y., & McNamara, D. S. (2007). A multidimensional framework to evaluate reading assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107–136). Mahwah, NJ: Erlbaum.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory.* Hillsdale, NJ: Erlbaum.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111–151.

Mares, M. L. (2006). Repetition increases children's comprehension of television content—Up to a point. *Communication Monographs, 73,* 216–241.

Marsh, H. W., & Hau, K. T. (1996). Assessing goodness of fit: Is parsimony always desirable? *The Journal of Experimental Education, 64,* 364–390.

Maruyama, G. M. (1998). *Basics of structural equation modeling.* Thousand Oaks, CA: Sage Publications.

McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theory, interventions, and technologies.* Mahwah, NJ: Erlbaum.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67,* 359–370.

Nation, K., & Snowling, M. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language, 39,* 85–101.

O'Brien, E. J., & Myers, J. L. (1987). The role of causal connections in the retrieval of text. *Memory and Cognition, 15,* 419–427.

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In McNamara, D. (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47–72). New York: Erlbaum.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18,* 443–468.

Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. S. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods, 40,* 1001–1015.

Papadopoulos, T. C., Das, J. P., Parrila, R. K., & Kirby, J. R. (2003). Children at-risk for developing reading difficulties: A remediation study. *School Psychology International, 24,* 340–366.

Paris, A. H., & Paris, S. G. (2003). Assessing narrative comprehension in young children, *Reading Research Quarterly, 38,* 36–76.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist, 36,* 89–101.

Paris, S. G., & Stahl, S. A. (Eds.). (2005). *Children's reading comprehension and assessment.* Mahwah, NJ: Erlbaum.

Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices-past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 13–69). Mahwah, NJ: Erlbaum.

Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C. C., Morrow, L., Tracey, D., et al. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5,* 35–58.

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

Rapp, D. N., van den Broek, P., McMaster, K. L., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11,* 289–312.

Rose, J. (2006). *Independent review of the teaching of early reading: Final report.* London: Department for Education and Skills/TSO.

Ruddell, R. B., & Ruddell, M. R. (1994). Language acquisition and literacy process. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading,* (4th ed., pp. 83–103). Newark, DE: International Reading Association.

Savage, R. S. (2001). The "simple view" of reading: Some evidence and possible implications. *Educational Psychology in Practice, 17,* 17–33.

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

Snowling, M. J., & Hulme, C. (2005). *The science of reading: A handbook.* Malden, MA: Blackwell.

Speece, D. L., Roth, F. P., Cooper, D. H., & de la Paz, S. (1999). The relevance of oral language skills to early literacy: A multivariate analysis. *Applied Psycholinguistics, 20,* 167–190.

Stanovich, K. E. (1986). Cognitive processes and the reading problems of learning-disabled children: Evaluating the assumption of specificity. In B. Y. L. Wong & J. K. Torgeson (Eds.), *Psychological and educational perspectives on learning disabilities* (pp. 87–131). Orlando, FL: Academic Press.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. 0. Freedle (Ed.), *New directions in discourse processing*: *Vol. 2. Advances in discourse processes*. (pp. 53–120). Hillsdale, NJ: Erlbaum.

Storch, S., & Whitehurst, G. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934–947.

Stuart, M., Stainthorp, R., & Snowling, M. (2008). Literacy as a complex activity: Deconstructing the simple view of reading. *Literacy, 42,* 59–66.

Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15,* 249–275.

Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. L. Stein, & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 83–111). Hillsdale, NJ: Erlbaum.

Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language, 24,* 612–630.

Trabasso, T., van den Broek, P., & Liu, L. (1988). A model for generating questions that assess and promote comprehension. *Questioning Exchange, 2,* 25–38.

Trabasso, T., van den Broek, P. W., & Suh, S. Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes, 12,* 1–25.

Tunmer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 175–214). Hillsdale, NJ: Erlbaum.

U.K. Department for Education and Skills. (2006). *Five-year strategy for children and learners: Maintaining the excellent progress. Strategy paper.* London: TSO.

van den Broek, P. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 423–446). Hillsdale, NJ: Erlbaum.

van den Broek, P. (1994). Comprehension and memory of narrative texts.

In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539–588). London: Academic Press.

van den Broek, P. (1997). Discovering the cements of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. Bauer, & T. Bourg (Eds.), *Developmental spans in event comprehension: Bridging fictional and actual events* (pp. 321–342). Mahwah, NJ: Erlbaum.

van den Broek, P., Kendeou, P., Kremer, K., Lynch, J. S., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. Stahl & S. Paris (Eds.), *Children's reading comprehension and assessment,* (pp. 107–130). Mahwah, NJ: Erlbaum.

van den Broek, P., & Kremer, K. E. (1999). The mind in action: What it means to comprehend during reading. In B. Taylor, M. Graves, & P. van den Broek (Eds.), *Reading for meaning* (pp. 1–31). New York: Teacher's College Press.

van den Broek, P., Lorch, E. P., & Thurlow, R. (1996). Children's and adults' memory for television stories: The role of causal factors, story-grammar categories, and hierarchical level. *Child Development, 67,* 3010–3028.

van den Broek, P., & Lorch, R. F. (1993). Network representations of causal relations in memory for narrative texts: Evidence from primed recognition. *Discourse Processes, 16*(1–2), 75–98.

van den Broek, P. W. (1986). Judging the importance of events in stories: The influence of structural variation and grade level. *Dissertation Abstracts International, 46*(9-B), 3243.

van den Broek, P. W. (1988). The effects of causal relations and hierarchical position on the importance of story statements. *Journal of Memory and Language, 27,* 1–22.

van den Broek, P. W. (1989a). Causal reasoning and inference making in judging the importance of story statements. *Child Development, 60,* 286–297.

van den Broek, P. W. (1989b). The effects of causal structure on the comprehension of narratives: Implications for education. *Reading Psychology: An International Quarterly, 10,* 19–44.

VanderVeen, A., Huff, K., Gierl, M., McNamara, D. S., Louwerse, M., & Graesser, A. C. (2007). Developing and validating instructionally relevant reading competency profiles measured by the critical reading sections of the SAT. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 137–172). Mahwah, NJ: Erlbaum.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skill model of reading development. *Scientific Studies of Reading, 11,* 3–32.

Wenner, J., & Bauer, P. J. (2001). Bringing order to the arbitrary: One- to two-year-olds' recall of event sequences. *Infant Behavior and Development, 22,* 585–590.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69,* 848–872.

Williams, J. P. (1993). Comprehension of students with and without learning disabilities: Identification of narrative themes and idiosyncratic text representations. *Journal of Educational Psychology, 85,* 631–642.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Test (WRMT).* Circle Pines, MN: American Guidance Services.

Received April 8, 2008
Revision received March 11, 2009
Accepted March 27, 2009 ■