Scientific Studies of Reading, 00:1–21, 2014

ISSN: 1088-8438 print/1532-799X online
DOI: 10.1080/10888438.2014.943905

Routledge
Taylor & Francis Group

Individual Differences in Fifth Graders' Literacy and Academic Language Predict Comprehension Monitoring Development: An Eye-Movement Study

Carol McDonald Connor
Arizona State University

Ralph Radach and Christian Vorstius
University of Wuppertal

Stephanie L. Day and Leigh McLean
Arizona State University

Frederick J. Morrison
University of Michigan

In this study, we investigated fifth graders' ($n = 52$) fall literacy, academic language, and motivation and how these skills predicted fall and spring comprehension monitoring on an eye movement task. Comprehension monitoring was defined as the identification and repair of misunderstandings when reading text. In the eye movement task, children read two sentences; the second included either a plausible or implausible word in the context of the first sentence. Stronger readers had shorter reading times overall suggesting faster processing of text. Generally fifth graders reacted to the implausible word (i.e., longer gaze duration on the implausible vs. the plausible word, which reflects lexical access). Students with stronger academic language, compared to those with weaker academic language, generally spent more time rereading the implausible target compared to the plausible target. This difference increased from fall to spring. Results support the centrality of academic language for meaning integration, setting standards of coherence, and utilizing comprehension repair strategies.

Over three decades of research has confirmed that children who master the alphabetic principle are better readers than those who do not, that phonological awareness and phonological processing are causally related to word reading success and failure, and that these skills can be taught effectively (National Institute of Child Health and Human Development, 2000; Torgesen, 2002). Unfortunately, skill at reading words fluently does not ensure proficient reading for understanding (Arrington, Kulesz, Francis, Fletcher, & Barnes, 2014; Oakhill & Cain, 2012; Savage,

Correspondence should be sent to Carol McDonald Connor, Learning Sciences Institute, Arizona State University, PO Box 872111, Tempe, AZ 85287. E-mail: Carol.Connor@asu.edu
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hssr.

2006; Shankweiler et al., 1999; Storch & Whitehurst, 2002). In many cases students may have ineffective comprehension monitoring skills (Cain, Oakhill, & Bryant, 2004; Kinnunen & Vauras, 1995; Rapp & van den Broek, 2005; van der Schoot, Vasbinder, Horsley, & Reijntjes, 2009). We define comprehension monitoring as the *conscious and unconscious strategies and skills used to identify and repair misunderstandings or confusion that might occur during text reading* and are interested in two potential aspects of comprehension monitoring: (a) evaluation/identification of the misunderstanding and (b) repairing the misunderstanding. Both aspects likely require at least some level of metacognition although they may be fluent and automatic, particularly for students with stronger reading and academic language skills. The purpose of this study is to examine how individual differences in fifth graders' literacy skills, academic language, and motivation might be related to key processes of comprehension monitoring from the fall of fifth grade (about 10 years of age) to the spring of fifth grade 8 months later. To the best of our knowledge, no study to date has considered reading, academic language, and motivation effects on students' comprehension monitoring using eye movement tasks.

In the context of the present study, eye movement analyses are useful because they offer an opportunity to examine how students process text as they are reading, do not rely on the sophistication of children's metacognitive skills (Garrett, Mazzocco, & Baker, 2006; Kinnunen & Vauras, 2010; Rayner, Chace, Slattery, & Ashby, 2006), and can offer clues as to why children might succeed or fail to attend to the meaning of what they are reading (Radach, Schmitten, Glover, & Huestegge, 2009; Rayner, 1998; van der Schoot et al., 2009).

COMPREHENSION MONITORING

Comprehension monitoring is described as a metacognitive act involving evaluation and the acts by which understanding is regulated by the reader (Garrett et al., 2006; Kinnunen & Vauras, 1995; Wagoner, 1983), as well as the ability to reflect on what has been read (Oakhill & Cain, 2012). Comprehension monitoring is frequently examined as a deliberate reading strategy requiring conscious reflection (i.e., metacognition) in situations where participants expect anomalous words or linguistic constructions (e.g., Kinnunen & Vauras, 2010; Meyers, Lytle, Palladino, Devenpeck, & Green, 1990; Oakhill, Hartt, & Samols, 2005). For example, Oakhill et al. (2005) studied self-corrections, repetitions, and hesitations during reading aloud as indicators for comprehension monitoring efforts. They also distinguished between monitoring for word level, sentence level, and intrasentence inconsistencies, which might all be necessary for integrated comprehension of text.

There are likely reasons for poor comprehension monitoring. For example, there is evidence that some children do not have the formal oral language skills needed to fully understand the more complex syntax and unfamiliar vocabulary that is characteristic of the academic texts they are expected to read, particularly as they begin fifth grade (Snow, 2001). Students may not have developed the metacognitive skill and knowledge required for the task of monitoring their understanding (Kinnunen & Vauras, 2010). They may have poor motivation to do the work required for making sense of more difficult or confusing text (Guthrie, Anderson, Aloa, & Rinehart, 1999). Plus they likely set lower standards for creating coherent text representations and integrating meaning and may fail to develop complete and rich situation models (Snow, 2001; van den Broek, Risden, Fletcher, & Thorlow, 1996; van der Schoot et al., 2009).

Although there is some evidence to the contrary for our age range (McConkie et al., 1991), comprehension monitoring may be developmentally sensitive—that is, aspects of identifying misunderstanding and utilizing repair strategies may differ for students of different ages, and language and metacognitive development (Garrett et al., 2006; Kinnunen & Vauras, 2010; Oakhill & Cain, 2012; Zabrucky & Moore, 1989). Hence this study explores how and for whom comprehension monitoring might change over 8 months, from the fall to spring of fifth grade in the United States, when students are between 10 and 12 years of age. This is when children are increasingly confronted with more difficult text in the classroom, they are more likely to have content-specific teachers for language arts, social studies, and science, where such texts are part of instruction, and students are expected to learn from this text (Chall, 1967).

EYE MOVEMENT AND COMPREHENSION MONITORING STUDIES

The measurement and analysis of eye movements has been successful in the study of moment-to-moment information processing in reading (Kinnunen & Vauras, 2010). Over the last three decades, a large body of work has shown that there is a close relation between the time spent viewing a linguistic unit and the mental effort need to process it at various levels ranging from early orthographic to semantic and pragmatic processing. The vast majority of this work has focused on skilled adult readers, but recently there has also been an increase in research on children (see Radach & Kennedy 2013; Rayner, 2009, for recent overviews).

Eye movement studies that have examined comprehension monitoring in children have addressed complex processing at the text level, comparing conditions in which entire statements would either fit or violate the situation model established over the course of reading a story (van der Schoot et al., 2009). For example, Kinnunen and Vauras (1995) found that students with stronger skills read faster than did students with weaker skills, their reading included significantly more regressions (i.e., rereading text), and they were better able to report the main idea of the texts. Also, longer reading times for the internal inconsistencies, compared to other inconsistencies, were most highly correlated with comprehension of the passage, which provides the rationale for using internal inconsistencies in this study. In a more recent example, van der Schoot and colleagues (van der Schoot, Reijntjes, & van Lieshout, 2012) used a narrative inconsistency task to measure students' comprehension monitoring and the extent to which they were developing and updating a more complete situation model (i.e., building a coherent mental representation). They found that students with weaker comprehension skills demonstrated more difficulty updating their situation models than did students with stronger comprehension skills.

Using a paradigm similar to the one developed for this study, Rayner, Warren, Juhasz, and Liversedge (2004) presented (among other conditions) sentences of the form

*John used **a knife/an axe** to chop the large carrots for dinner.*

In both versions of the target sentence, the instrument (*knife* vs. *axe*) could be plausibly used in conjunction with the main verb "to chop." However, even though axes are regularly used to chop things and carrots are often chopped, an axe is not regularly used to chop carrots and is, hence, implausible. Rayner et al. (2004) found that their adult participants responded to the implausible condition with longer reading times, particularly in the form of fixations made after regressions from the critical word ("carrots" in the example). Building on this work, Joseph and colleagues

(Joseph et al., 2008) compared how adults and children responded to implausible (and anomalous) sentences of the same form. They found that children were also sensitive to plausibility, but with longer reading times compared to adult readers.

For this study, we developed a task consisting of 20 sentence pairs. For each item the first sentence introduced an event or action that continued in the second. The second sentence contained either a more plausible or less plausible object or instrument in relation to the previous verb and ongoing event established in the first sentence. For example,

Last week Kyle flew to visit his family in another city.

The large ***plane/truck*** *was spacious and quickly transported them.*

The word *plane* in the second sentence is highly plausible because the child has just read that Kyle *flew* to another city. The word *truck* is implausible because trucks cannot fly. In this task, students must detect internal inconsistencies, which appear to be more difficult that other kinds of inconsistencies (e.g., lexical or external/factual inconsistencies; Kinnunen & Vauras, 1995, 2010; Zabrucky & Moore, 1989). This task has the advantage in that it allows enough time for the semantic relations in the first sentence to be processed so that developing readers should be able to develop an integrated event or mental representation before encountering the new concept.

We used two oculomotor measures in this study: *gaze duration* and *rereading time*, which are assumed to reflect different stages in the timeline of word processing. The initial fixation (when the eye first views the target word) represents orthographic and early lexical processing (Radach & Kennedy, 2004; Reingold, Reichle, Glaholt, & Sheridan, 2012), whereas *gaze duration* (the sum of all fixations before the first saccade) is closely related to later stages of word processing, including lexical access. *Rereading time,* which is the summed duration of all fixations on the target word after the first saccade leaves the word, is assumed to reflect postlexical integration of meaning at the sentence or text level (Garner & Reis, 1981; Inhoff & Radach, 1998; Radach & Kennedy, 2013; Rayner, 1998). We conjecture that longer gaze duration and rereading time for implausible words compared to plausible words, respectively, might be considered diagnostic for the two different aspects of comprehension monitoring: (a) detecting inconsistencies and (b) repairing misunderstanding or confusions. We hypothesize that these two aspects of comprehension monitoring may be sensitive to individual student differences in literacy, academic language, and motivation and to overall development from the beginning to the end of fifth grade.

ACADEMIC LANGUAGE

Accumulating research continues to highlight the significant role that children's linguistic skills play in reading for understanding (Cain et al., 2004; National Early Literacy Panel, 2008; National Institute of Child Health and Human Development, 2000; Snow, 2001). Whereas the connections between early oral language skills and later success in literacy have been well defined (Snow, Burns, & Griffin, 1998), the underlying mechanisms are less well understood (Oakhill & Cain, 2012). Established theories suggest direct links between oral language and literacy development, with early vocabulary providing a foundation for later literacy development and, along with decoding, composing an integral part of reading comprehension (Hoover & Gough, 1990; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001; Storch & Whitehurst, 2002). Building on these theories, recent research suggests that literacy may be dependent on a wide array of early oral language skills, including background and academic knowledge, which interact to impact literacy development and proficiency (Connor et al., 2014; Dickinson, Anastasopoulos, McCabe, Peisner-Feinberg, & Poe, 2003). These skills develop at different times during early literacy acquisition, in the context of instruction, and interact in complex ways to affect children's literacy achievement. Hence, the language skills students bring to the process of reading for understanding may be more complex than can be captured by any one measure of language (Kim & Phillips, 2014; Oakhill et al., 2005). We posit that students' academic knowledge, narrative language, and vocabulary form a complex oral language construct, distinct from literacy, that focuses on the language skills that would tend to be critical in the context of schooling, *academic language* (Snow, 2010).

MOTIVATION

Children's motivation and purpose for reading may particularly influence the standards of coherence they set for meaning integration motivation (Guthrie et al., 2009; van den Broek, White, Kendeou, & Carlson, 2009), which would, in turn influence the effort applied to comprehension monitoring. Hence, a critical element of successful reading achievement includes children's interest and motivation to read for a variety of purposes (Snow et al., 1998).

THE CURRENT STUDY

Today's world requires high standards for learning from complex text and content area literacy (e.g., in the United States; Common Core State Standards Initiative, 2010). Better understanding how students respond to text that is difficult to understand and thus may not make sense to them, particularly if motivation, literacy, and academic language skills influence this response, has theoretical, educational, and policy implications. This includes what might be done to improve comprehension monitoring for students with weak literacy skills.

Therefore, the purpose of this study was to examine developmental and individual differences in fifth graders' comprehension monitoring. We hypothesized that students with weaker literacy skills would have longer reading times compared to students with stronger literacy skills primarily because their reading processing would be slower overall. The key question, however, was whether the *difference* between the plausible and implausible target words for gaze duration and rereading time would vary with individual differences in children's literacy skills, academic language skills, and motivation for reading. One reasonable hypothesis is that children with weaker literacy and academic language might be less sensitive to whether the target word was plausible or implausible and so the gaze duration difference between plausible and implausible target words would be smaller for children with weaker literacy and academic language skills and larger for children with stronger skills.

Another reasonable hypothesis is that children with weaker literacy, academic language, and/or motivation might set lower standards of coherence for meaning integration (Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007). That is, they may care less about their confusion when reading and so do not employ repair strategies. In this case, the children with weaker skills might be expected to have shorter rereading time differences between plausible and implausible sentenced compared to children with stronger skills. That is, they might set lower standards of coherence for meaning integration than would children with stronger skills. If differences in comprehension monitoring are related to motivation, we would see a similar result—rereading time differences would be greater for children with greater rather than less motivation for reading (Logan, Medford, & Hughes, 2011). Because children are actively involved in instruction and continue to mature throughout the school year, we conjecture that overall reading time might decrease after 8 months of schooling in fifth grade, whereas comprehension monitoring might improve. That is, gaze duration and rereading differences between plausible and implausible words might increase over time.

METHODS

Participants

The 52 fifth-grade English-speaking students (*M* age = 10.6 years in the fall, range = 10–12 years, *SD* = .62) in this study attended four classrooms in one U.S. school in north Florida that was ethnically and economically diverse with approximately 47% of the children qualifying for the U.S. Free and Reduced Lunch Program, a widely used indicator of family poverty. Sixteen percent were White, 77% were African American, and the remaining students belonged to other ethnicities. Forty-five percent were girls. The students were part of a larger study on fifth-grade instruction and achievement conducted in six schools (*n* = 396 students). The eye movement study was conducted at one of the schools, and all the fifth graders attending that school for whom we had written parental consent participated in the eye movement study. The school was intentionally selected because it was highly diverse ethnically, racially, and economically and the children demonstrated a range of literacy skills. Except for the eye movement procedures, all of the students in the larger study followed the same protocols.

Measures

Comprehension Monitoring Assessment

Eye movement task. Items in the reading experiment consisted of 20 pairs of declarative sentences. The first sentence of each reported a simple event or action (e.g., flying to another city), whereas the second served to extend or continue this event. A noun in the second sentence represented either a typical or atypical object or instrument in the given context, creating a plausible versus implausible continuation of the ongoing action (e.g., plane vs. truck). Target words in both versions of the second sentence were matched for word length (range = 4–8 letters, *M* = 5.4 for both versions), number of syllables (implausible 1.5 vs. 1.6 plausible) and morphological complexity (implausible 1.10 vs. implausible 1.05 morphological components). The norms developed by Zeno, Ivens, Millard, and Duvvuri (1995) were used to ensure the target words were less than fifth-grade level. Mean word frequencies in the in the CELEX word corpus (Baayen, Piepenbrock, & van Rijn, 1993) amounted to 56.7 word per million for implausible and 46.7 words per million for implausible words (all differences *ns*). Students completed the eye movement task in September and October of the school year and again in May, with a mean of 8 months between assessments.

Apparatus. Using a 21-in. monitor with a display resolution of 1024 × 768 pixels and a refresh rate of 120 Hz, sentence pairs were displayed one by one in black text on a light gray background. The text font was Courier New 15 point. The distance between the reader's eyes and the monitor was 68 cm, resulting in a visual angle of .33° per letter. Eye movements were recorded using an EyeLink 1000 video-based pupil and corneal tracking system (SR Research Ltd.). Every four trials, a 3-point calibration was performed and drift checks preceded every trial. The online saccade parser was set to detect saccades (the movement of the eye between fixations) with an amplitude of 0.15 degrees or greater using an acceleration threshold of 8,000 degrees/second and a velocity threshold of 30 degrees/second. These settings have produced accurate and reliable data in a number of reading studies with adults and children (see Inhoff & Radach, 1998, for a detailed discussion of eye-tracking methodology).

Procedure. The eye movement task was one of three administered to each of the children, and so they were familiar with the calibration routines. Children were first seated in front of the display monitor and were then given a set of directions on the screen instructing them to read the presented sentence pairs in sequential order. Each child was presented with 20 sentence pair items and of those 20, 10 contained plausible words and 10 contained implausible words (see Figure 2). The lists were counterbalanced (see Appendix B) over two lists so a student who saw the more plausible version in the fall would get the less plausible version in the spring. The present experiment was actually mixed with another one that also included 40 sentences (with no implausible words) for a different research question, so that these extra sentences served as fillers. Therefore, participating children were presented with a total of 80 sentences that contained 10 contextually atypical nouns creating internal inconsistencies in the context of the first sentence.

Data analysis. EyeLink software was used to segment raw data into saccades and fixations. Gaze duration, Reread time, and Total time (ms) were computed using a custom-built software suite, EyeMap, (Tang, Reilly, & Vorstius, 2011) and SPSS (version 19). All data were visually inspected, and any fixations with durations shorter than 70 ms or longer than 2 standard deviations of the participants' mean were removed from further analyses, excluding about 4% of the fixation data.

Literacy, Language, and Motivation Assessment

Literacy measures. Students' literacy skills were assessed in the fall (August and September) of 2010 in a quiet place in their school following the standardized test protocols. We administered the Woodcock–Johnson III (WJ-3; Woodcock, McGrew, & Mather, 2001) Letter-Word Identification and Passage Comprehension subtests and the Gates–MacGinitie Reading Test, Fourth Edition (MacGinitie, MacGinitie, Maria, & Dreyer, 2002) reading comprehension and reading vocabulary tests. For this study, we used only the fall assessment data. The letter-word identification test asks children to read lists of words of increasing complexity. The passage comprehension task utilizes a cloze procedure where children read sentences or short passages of increasing difficulty and are asked to supply the missing word. Both tests have excellent psychometric properties (median reliabilities of .91 and .83 respectively) and are administered individually.

The Gates–MacGinitie Reading Test reading vocabulary and comprehension are group administered. Reading vocabulary presents the children with a written sentence containing the target word and asks the children to choose one of four definitions. The reading comprehension test asks children to read passages of varying complexity and to answer multiple-choice questions that vary in the extent to which the correct answer must be inferred. Again, the assessments have good psychometric properties.

Language measures. Language skills were assessed using the WJ-3 picture vocabulary, academic knowledge, and story-recall subtests, following the assessment protocols. The picture vocabulary task asks children to identify pictures of increasingly unfamiliar words and has a median reliability of .77. The academic knowledge task asks children to respond to questions about humanities, social studies, and science topics. For example, in the science subtest, children are asked, "What organ in the human body exchanges carbon dioxide for oxygen?" This task has a median reliability of .88. The story recall task is designed to assess aspects of oral language and memory. It requires the children to listen to passages of increasing complexity. After listening to a passage, they are asked to recall as many details of the story as they can remember. The assessment has a median reliability of .87.

Motivation. Motivation was assessed for 44 of the 52 children in January 2011 using a survey adapted from measures designed by Pintrich (2003) and Wigfield and Guthrie (1997). Missing data analyses revealed no differences in the reading and language measures between students who did and did not complete the survey. Students completed the survey independently in their classrooms or in small groups in a quiet location in their school supervised by a trained research assistant. The students read probes such as, "I learn new things when I read" or "I learn more when I work hard" and rated from 1 to 4 whether this was *very different from me* (1) to *a lot like me* (4). In general, students with higher scores are assumed to have a greater motivation to master or understand what it is they are reading (mastery goal orientation) whereas students with lower scores are assumed to read because it is a requirement or to avoid failure (performance goal orientation; Pintrich, 2003; Pintrich & Schrauben, 1992). Reliability was adequate ($\alpha = .74$).

RESULTS

Literacy and Academic Language

To better understand our language and literacy constructs and to increase the precision of our predictors, we utilized structural equation modeling confirmatory factor analysis to identify latent variables composed of multiple measures to represent our constructs (Hoyle, 1995; Kline, 1998) using the larger sample of fifth graders ($n = 396$) that included the 52 children in this study. Correlations and means for this sample of 52 students are provided in Appendix A, Table A1. We tested models systematically comparing fit and parsimony. The model that best fit the data (Figure 1) fit the data adequately (Tucker–Lewis index = .964, comparative fit index = .984, root mean square error of approximation = .077, Akaike Information Criterion =

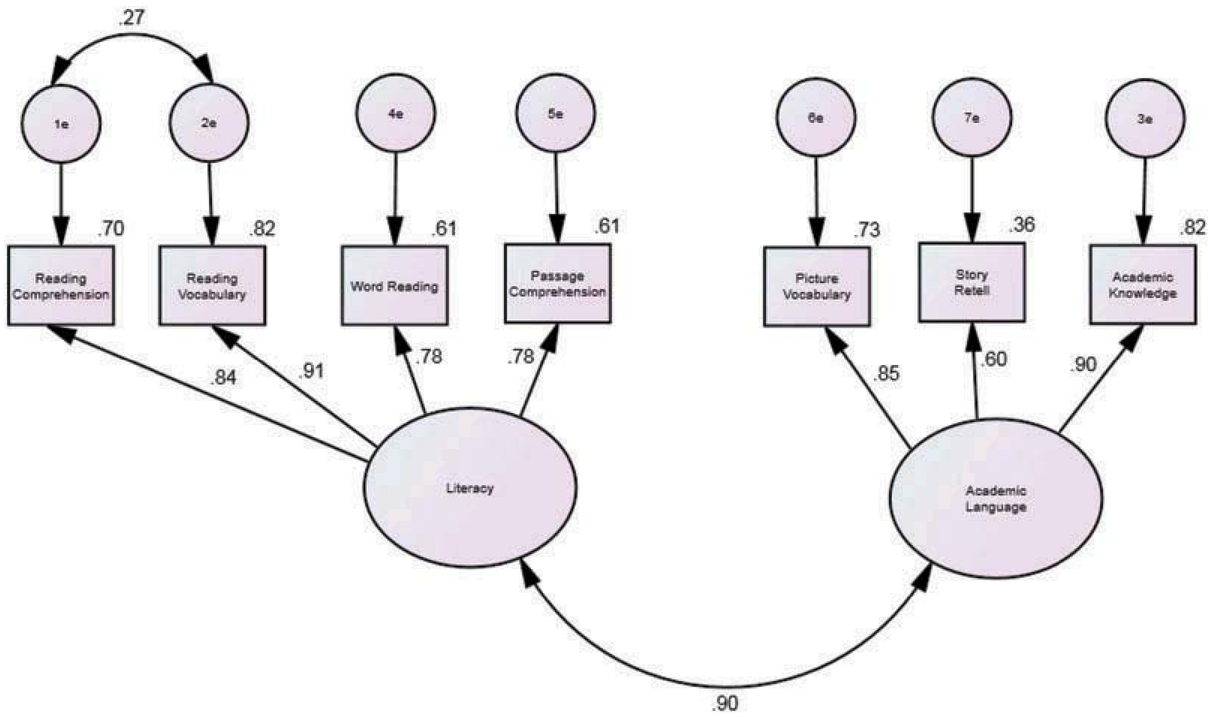

FIGURE 1 Results of structural equation modeling showing the two latent variables model. *Note.* All standardized path coefficients shown are significantly greater than 0 ($p < .05$).

86.108)[1] and supported our hypothesis that academic language represented a single complex construct, separate from literacy. Hence we used two latent variables in our analyses—Literacy and Academic Language. To create the variables, W scores (WJ-3 tests) and scale scores (Gates–MacGinitie Reading Test) were z scored, totaled, and divided by the number of assessments for the respective variables Literacy and Academic Language and used in all subsequent analyses (see Table 1).

In general, the 52 students' word reading and comprehension skills were in line with national norms (see Table A1) and similar to the sample ($n = 396$) as a whole. Fall Literacy and Academic Language z scores were significantly correlated ($r = .717$, $p < .001$), as were Literacy and Motivation ($r = .300$, $p = .034$). Academic Language and Motivation scores were not significantly correlated. All other descriptive statistics are provided in Table 1.

Comprehension Monitoring Model Results

Hierarchical Linear Modeling (Raudenbush & Bryk, 2002) revealed that, as anticipated, on average, Gaze duration and Reread time were significantly longer for implausible than for plausible target words (differences of 64 and 93 ms, respectively). However, there was no significant main effect of session (fall vs. spring) for either measure (see Table A2). The intraclass correlation,

[1]The two-variable model had superior fit compared to the single and three construct models (comprehension, word reading, and academic language).

TABLE 1
Descriptive Statistics of Variables Used in HLM

Variable Name	M	SD	*Min-Max*
Literacy (*z* score)	−.60	.98	−2.34–2.15
Academic Language (*z* score)	−.40	.80	−2.71–2.18
Motivation (raw score)[a]	7.96	4.80	−5–19
Fall Gaze Duration (ms)	376	137	167–857
Implausible Target	398	134	
Plausible Target	355	139	
Fall Reread Time (ms)	181	144	0–754
Implausible Target	211	157	
Plausible Target	152	124	
Fall Total Reading time (ms)	558	212	205–1,282
Implausible Target	609	202	
Plausible Target	507	211	
Spring Gaze Duration (ms)	368	114	186–805
Implausible Target	329	134	
Plausible Target	409	118	
Spring Reread Time (ms)	185	195	0–916
Implausible Target	243	232	
Plausible Target	129	127	
Spring Total Reading Time (ms)	554	247	200–1,319
Implausible Target	615	275	
Plausible Target	458	168	

[a]Some items were subtracted from the total.

which is the proportion of between-student variance explained, was .37 for Gaze duration and .38 for Reread time.

We next considered individual student difference effects for each outcome by entering Literacy, Academic Language and Motivation at Level 2 and testing interaction effects (by plausible/implausible target word and by session). Motivation did not significantly predict either Gaze duration or Reread time and so was trimmed from the model to preserve parsimony. Because Literacy and Academic Language were correlated, there were concerns with multicollinearity, particularly with interaction terms. Therefore, we considered each variable individually and used these results to inform the final model. Variables that did not significantly predict the outcome were trimmed from the model to preserve parsimony. Results for these models are described in Appendix A.

Final Models

Results for the final models (see Table 2 and Figure 2) reveal that overall, the stronger the children's Literacy score, the shorter were their Gaze duration (Table 2, top) and Reread times (Table 2, bottom). Students' Gaze duration and Reread Time were longer for implausible than for plausible words.

For Reread time, there was an Academic Language × Implausible Target interaction effect. Students' Reread time depended on their academic language skills: Differences were generally

TABLE 2
Final HLM Results for Gaze Duration (GAZEDUR, top) and Re-Reading Time (REREAD, bottom) (ms) as a Function of Children's Reading (READ) and Academic Language Skills (LANG)

Gaze Duration

Fixed Effect	*Coefficient*	SE	t *Ratio*	*Approx.* df	p
Fitted Mean GAZEDUR, β_{00}	303.65	15.52	19.565	49	< 0.001
Literacy, β_{01}	−88.28	12.57	−7.023	49	< 0.001
Academic Language, β_{02}	32.90	20.10	1.636	49	0.108
For IMPLAUSIBLE target word, π_1					
Intercept, β_{10}	59.94	17.13	3.498	288	< 0.001
Academic Language, β_{11}	−10.00	18.70	−0.535	288	0.593
For SPRING session, π_2					
Intercept, β_{20}	−3.36	17.39	−0.194	288	0.847
Academic Language, β_{21}	38.89	18.91	2.056	288	0.041

Random Effect	SD	*Variance Component*	df	χ^2	p
Intercept, r_0	24.31	591.16	49	53.934	0.291
level-1, *e*	100.46	10092.70			

Reread Time

Fixed Effect	*Coefficient*	SE	t *Ratio*	*Approx.* df	p
Fitted Mean REREAD, β_{00}	109.52	24.63	4.45	49	<.001
Literacy, β_{01}	−57.35	24.70	−2.32	49	.024
Academic Language, β_{02}	26.33	34.24	0.77	49	.446
For IMPLAUSIBLE target word slope, π_1					
Intercept, β_{10}	121.22	21.22	5.71	288	<.001
Academic Language, β_{11}	66.45	23.17	2.87	288	.004
For SPRING slope, π_2					
Intecept, β_{20}	19.29	22.60	0.85	288	.394
Academic Language, β_{21}	58.70	24.13	2.43	288	.016

Random Effect	SD	*Variance Component*	df	χ^2	p
INTRCPT1, r_0	96.07	9230.35	49	144.295	<.001
level-1, *e*	124.44	15485.52			

Note. Continuous variables are grand mean centered. There was no significant between classroom variance and so more parsimonious 2-level models were used—repeated measures nested in students. The final model for Reread time was

$$\textit{Reread time}_{ti} = \beta_{00} + \beta_{01}{}^{*}\textit{ZLiteracy}_i + \beta_{02}{}^{*}\textit{ZAcademic Lanuage}_i$$
$$+ \beta_{10}{}^{*}\textit{Implausible Target}_{ti} + \beta_{11}{}^{*}\textit{ZAcademic Language}{}^{*}\textit{Implausible Target}_{ti}$$
$$+ \beta_{20}{}^{*}\textit{Spring Session}_{ti} + \beta_{21}{}^{*}\textit{ZAcademic Language}_i{}^{*}\textit{Spring Session}_{ti}$$
$$+ r_{0i} + e_{ti}$$

Thus, in this model, we identified the main effects of Literacy (β_{01}) and Academic Language (β_{02}) on Reread Time for students reading the plausible targets in the fall, as well as the difference in effect for the implausible target (1 = Implausible, 0 = Plausible) and Spring Session (0 = Fall, 1 = Spring). We also tested the interaction effects for Academic Language × Implausible Target (β_{11}) and Academic Language × Spring Session (β_{21}).

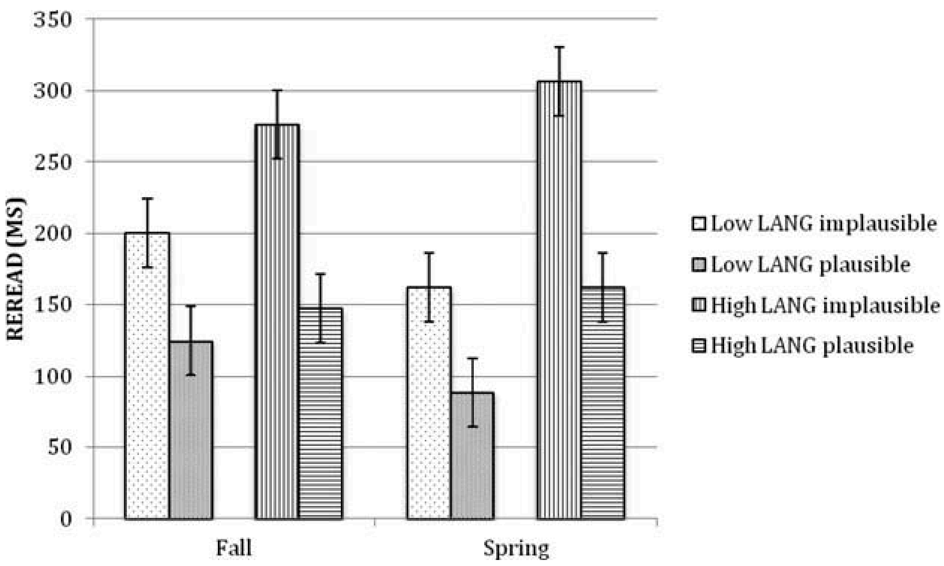

FIGURE 2 Comparing the difference in Rereading time (REREAD) for implausible versus plausible target words for children with high versus low Academic Language (LANG) scores. *Note.* Because LANG is a continuous variable, it was modeled selecting values that were well within the range of the data. Low LANG was modeled at the 25th percentile of the sample (−.775) and High LANG was modeled at the 75th percentile of the sample (.109). Error bars represent the HLM standard error for the intercept. Literacy scores are held constant at the sample mean.

greater as students' fall Academic Language scores were greater (see Figure 2). There was no interaction effect on Gaze duration.

With regard to fall to spring changes in Gaze duration and Reread times, there was an Academic Language × Session effect. The stronger were students' fall Academic Language scores, and the greater were both the Gaze duration and Reread times in the spring compared to fall. That is, Gaze duration and Reread time increased more over the school year as students' fall Academic Language increased (see Figure 2).

DISCUSSION

As predicted, students with stronger literacy skills processed text faster compared to students with weaker literacy skills. Regardless of literacy skill, academic language, or motivation, students' gaze duration was greater for words that were not plausible in context (i.e., internal inconsistencies: *truck* vs. *plane*) compared to those that were. Hence *identifying the misunderstanding* aspect of comprehension monitoring does not appear to be the culprit in weak understanding and might be largely unconscious and automatic. Rather, it was what students did after they encounter the implausible word and *the extent to which they attempted to repair their misunderstanding* (i.e., comprehension regulation) that distinguished the students—students with weaker academic language skills generally spent less time rereading and trying to repair their understanding of implausible compared to plausible words than did children with stronger academic language skills. We had hypothesized that literacy and academic language would operate in similar ways, but this was not supported by the results. Instead, greater literacy skill predicted shorter gaze duration and rereading times, whereas academic language skills predicted the extent to which students attempted to repair their comprehension represented by the difference in reread time for plausible versus implausible words.

Comprehension Monitoring

Academic Language

The results of this study point to the centrality of academic language skills in reading for understanding and indicate why comprehension may break down even when reading skills are adequate for the task. The standards of coherence for meaning integration that children set for themselves and the extent to which they reread words that are implausible in context may be based on their ability to apply their academic language skills—vocabulary, background knowledge, and understanding of text structure—to the task of understanding text. This finding offers a potentially important reason why academic language skills are critical to proficient reading comprehension (Snow, 2010; Verhoeven, van Leeuwe, & Vermeer, 2011). This stands in some contrast to previous results where literacy skills were related to rereading time (Garner & Reis, 1981; Kinnunen & Vauras, 1995; Zabrucky & Moore, 1989). It is the case, however, that academic language was not included as a potential source of individual child differences in these studies. Literacy and Academic Language are correlated; thus these earlier findings do not contradict ours. Rather, our study replicates and extends these findings by considering the independent influence of literacy and academic language ability and how they might operate differently to support comprehension monitoring identification and repair.

Automatic Versus More Conscious Processes

As we consider theories of reading, these results suggest that there are automatic as well as more controlled or conscious aspects to understanding text. Our results indicate that detection of internal incoherence (i.e., the implausible word) might be more automatic than previously conjectured (Cain et al., 2004; Connor, 2013; Oakhill et al., 2005; Rapp et al., 2007). Even children with weaker literacy and academic language skills had longer gaze durations for implausible words than for plausible words. Hence the metacognitive skill of recognizing when comprehension breaks down might be more automatic, and perhaps more unconscious, than previously thought, at least for fifth graders. Recent evidence from neurological studies supports this distinction (Grammer, Carrasco, Gehring, & Morrison, 2014; Morrison, Grammer, Kim, & Gehring, 2014). After committing an error, both adults and children exhibit a very rapid neurophysiological response (termed Error-Related Negativity) followed by a later response (Error Positivity), thought to reflect more conscious awareness or emotional reactivity to the error. The Error-Related Negativity was developmentally invariant while the Error Positivity improved with age and predicted academic achievement. These findings reinforce results from the present study in differentiating automatic versus more conscious responses to stimulus or response discrepancies.

If children automatically notice when text does not make sense, then interventions might focus on bringing this automatic reaction to their attention and helping them develop strategies to repair their understanding. This might be most effective when accompanied by learning opportunities that build general language and text-specific skills, as well as the academic knowledge they will need to resolve the confusion.

Developmental Effects

Developmentally, our results suggest that the automatic processes of comprehension monitoring (i.e., identification as represented by gaze duration), as well as the metacognitive and potentially more conscious aspects of comprehension monitoring (i.e., repairing as represented by rereading time) are improving from fall to spring of fifth grade but only for children with stronger literacy and academic language skills. In the few available longitudinal studies (e.g., McConkie et al., 1991), general oculomotor parameters changed very little from Grade 4 to 5, but this was partly due to the use of "grade-appropriate" text materials that became more difficult (on several psycholinguistic dimensions) so that existing developmental progress may be obscured. This problem has been avoided in our study by using the same sentence frames at both testing points while critical words pairs were counterbalanced over two parallel lists (Vorstius, Radach, Mayer, & Lonigan, 2013, for a similar approach).

At the same time, our results generally support findings in other longitudinal studies that included comprehension monitoring. For example, Oakhill and colleagues investigated predictors of comprehension from 7 to 12 years of age (Oakhill & Cain, 2012). They found that early comprehension monitoring independently predicted reading comprehension gains (residualized change), providing evidence of a potential causal link, which may be reciprocal. Our results also support a potential reciprocal effect inasmuch as we observed gains in comprehension monitoring (i.e., greater re-read time from fall to spring) but only for those students with stronger academic language abilities. We conjecture that stronger academic language skills should support stronger comprehension monitoring (and in turn reading for understanding), which in turn would tend to support stronger academic language gains (see also Connor et al., 2014). Children with weaker academic language skills were generally not improving their comprehension monitoring and repair skills to the same extent, which may put them at increased risk for academic underachievement.

Motivation

Motivation for reading did not predict either gaze duration or rereading times, although children with stronger literacy skills did tend to report higher levels of motivation for reading and learning. There are a number of reasons that might explain the null findings. First, we had only one measure of motivation and so could not compute a latent variable, which would have helped to reduce measurement error and is a limitation of this study. Although our alphas were acceptable (.74), there may have been too much noise, so to speak, to detect a potentially important effect given our somewhat limited sample size. Second, motivation was assessed in the middle of the school year rather than in the fall when the literacy measures were administered. In addition, our measure of more general motivation for reading might not be applicable to this kind of, arguably, artificial task. Another explanation may be that the causal direction of the association observed in this and other studies is from reading to motivation—that stronger readers are also more motivated to learn from text (Guo, Sun, Breit-Smith, Morrison, & Connor, in press). Hence, the effect of motivation would be realized through the effect of reading (and indirectly for academic language) and would not present a unique influence. This remains to be tested.

Limitations

There are a number of limitations that should be noted when interpreting these results. First, there may have been specific characteristics of this sample that amplified the effects we observed. For example, in this sample there was a fairly substantial gap between word reading scores and comprehension (see Appendix A), and we had only one measure of word reading. Second, there is evidence that by fifth grade, decoding and reading comprehension form separate constructs (Oakhill & Cain, 2012), which was not the case with this sample; we found only one multidimensional literacy factor (but see Mehta, Foorman, Branum-Martin, & Taylor, 2005). Finally, these children all attended one school where almost half of the children were living in poverty. These results may not generalize to more mainstream populations.

Implications for Instruction

This study's focus on comprehension monitoring and how students respond to implausible and confusing text suggest directions for instruction and intervention. First, early support for language and academic knowledge development is warranted. As early as preschool, oral language skills are strong predictors of later reading proficiency (National Early Literacy Panel, 2008; Storch & Whitehurst, 2002), and these results suggest an additional rationale for supporting early academic language development. As we see with the children in this study, in general, U.S. schools have been successful in teaching children how to decode text but have been less successful in teaching comprehension and academic language skills. These children were part of a longitudinal study, and first-grade classroom observations revealed that, although children generally received sufficient instruction in phonological awareness, phonics, and decoding, substantially less time was spent in more meaning-focused and content area instruction, which tend to support academic language and comprehension skills (Connor & Morrison, 2012; Connor, Morrison, et al., 2009; Connor et al., 2011; Connor, Piasta, et al., 2009). Another target for intervention might include raising students' standards of coherence for meaning integration, perhaps by teaching them to more actively detect and stop when a word or sentence does not make sense and then use strategies to repair their misunderstandings. Think aloud strategies may facilitate this (Baumann, Seifert-Kessell, & Jones, 1992). Based on these findings, we are developing an interactive e-book designed to build word knowledge and comprehension monitoring, which uses unfamiliar vocabulary, embedded comprehension questions, and a choose-your-own-adventure format. If the child is not setting high standards of coherence for meaning integration, there are consequences—choosing a particular vocabulary word (e.g., adamant vs. concede) changes the plot of the story, and they are required to reread the pages when they answer comprehension questions incorrectly.

The next great educational challenge in the United States and, we argue, for most nations will be supporting students' reading for understanding as they confront the challenges of a global and highly complex world. Eye movement studies that consider individual differences and development, in combination with other research, can enhance our understanding of individual student characteristics and how they contribute to or interfere with processes that promote proficient literacy skills. Designing and using effective instructional regimens and targeted interventions based on this knowledge may help students and their teachers meet the challenges of reading for understanding and learning.

ACKNOWLEDGMENTS

We thank the Individualizing Student Instruction (ISI) Project team and the children, parents, teachers, and school participants.

FUNDING

Funding was provided by the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development: R21 R21HD062834 and R01HD48539 and the U.S. Department of Education, Institute of Education Sciences: R305H04013, R305B070074 and R305F100027.

REFERENCES

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*, *18*, 325–346. doi:10.1080/10888438.2014.902461

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database [CD-ROM]. Philadelphia, PA: University of Pennsylvania.

Baumann, J. F., Seifert-Kessell, N., & Jones, L. A. (1992). Effect of think-aloud instruction on elementary students' comprehension monitoring abilities. *Journal of Literacy Research*, *24*, 143–172. doi:10.1080/10862969209547770

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.

Chall, J. S. (1967). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. Retrieved from http://www.corestandards.org/assets/CCSSI_Math Standards.pdf

Connor, C. M. (2013). Intervening to support reading comprehension development with diverse learners. In B. Miller & L. E. Cutting (Eds.), *Unraveling the behavioral, neurobiological and genetic components of reading comprehension: The Dyslexia Foundation and NICHD* (pp. 222–232). Baltimore, MD: Brookes.

Connor, C. M., & Morrison, F. J. (2012). Knowledge acquisition in the classroom: Literacy and content area knowledge In A. M. Pinkham, T. Kaefer, & S. B. Neuman (Eds.), *Knowledge development in early childhood: Sources of learning and classroom implications* (pp. 220–241). New York, NY: Guilford.

Connor, C. M., Morrison, F. J., Fishman, B., Ponitz, C. C., Glasney, S., Underwood, P. . . . , Schatschneider, C. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, *38*, 85–99.

Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J., Lundblom, E. G., Crowe, E., & Fishman, B. (2011). Effective classroom instruction: Implications of child characteristic by instruction interactions on first graders' word reading achievement. *Journal for Research on Educational Effectiveness*, *4*, 173–207.

Connor, C. M., Phillips, B., Kaschak, M., Apel, K., Kim, Y.-S., Otaiba, S., . . . Lonigan, C. (2014). Comprehension tools for teachers: Reading for understanding from prekindergarten through fourth grade. *Educational Psychology Review*. Advance online publication. doi:10.1007/s10648-014-9267-1

Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., . . . Morrison, F. J. (2009). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development*, *80*, 77–100.

Dickinson, D. K., Anastasopoulos, L., McCabe, A., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, *95*, 465–481.

Garner, R., & Reis, R. (1981). Monitoring and resolving comprehension obstacles: An investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders. *Reading Research Quarterly*, *16*, 569–582. doi:10.2307/747316

Garrett, A. J., Mazzocco, M. M. M., & Baker, L. (2006). Development of the metacognitive skills of prediction and evaluation in children with or without math disability. *Learning Disabilities Research & Practice*, *21*, 77–88. doi:10.1111/j.1540-5826.2006.00208.x

Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: A school-based intervention. *Developmental Cognitive Neuroscience*, *9*, 93–105.

Guo, Y., Sun, S., Breit-Smith, A., Morrison, F. J., & Connor, C. M. (in press). Behavioral engagement and reading achievement in elementary-school aged children: A longitudinal cross-lagged analysis. *Journal of Educational Psychology*.

Guthrie, J. T., Anderson, E., Aloa, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *The Elementary School Journal*, *99*, 343–366.

Guthrie, J. T., McRae, A., Coddington, C. S., Lutz Klauda, S., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities*, *42*, 195–214. doi:10.1177/0022219408331039

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

Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of congitive processes. In G. Underwood (Ed.), *Eye guidance in rading and scene perception* (pp. 29–54). Oxford, UK: Elsevier.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults processing of analomaly and implausibility during reading: evidence from eye movements. *Journal of Experimental Psychology*, *61*, 708–723.

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

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high- and low-achieving primary school children's reading. *Learning and Instruction*, *5*, 143–165. doi:10.1016/0959-4752(95)00009-R

Kinnunen, R., & Vauras, M. (2010). Tracking online metacognition: Monitoring and regulating comprehension in reading. In A. Efklides & P. Misailidi (Eds.), *Trends and prospects in metacognition research* (pp. 209–258). New York, NY: Spring.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: Guilford.

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences*, *21*, 124–128. doi:10.1016/j.lindif.2010.09.011

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2002). *Gates-MacGinitie Reading Tests*, *Fourth Edition Technical Report for Forms S & T*. Itasca, IL: Riverside.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). *Children's eye movements during reading. Vision and visual dyslexia*. London, UK: Macmillan

Mehta, P. D., Foorman, B. R., Branum-Martin, L., & Taylor, W. P. (2005). Literacy as a unidimensional multilevel construct: Validation, sources of influence, and implications in a longitudinal study in Grades 1 to 4. *Scientific Studies of Reading*, *9*, 85–116.

Meyers, J., Lytle, S., Palladino, D., Devenpeck, G., & Green, M. (1990). Think-aloud protocol analysis: An investigation of reading comprehension strategies in fourth-and fifth-grade students. *Journal of Psychoeducational Assessment*, *8*, 112–127. doi:10.1177/073428299000800201

Morrison, F. J., Grammer, J., Kim, M., & Gehring, W. (2014, February). *Executive function in the transistion to school: Integrating brain and behavior*. Paper presented at the Pacific Coast Research Conference, San Diego, CA.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington DC: National Institute for Literacy and the National Center for Family Literacy.

National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (National Institute of Child Health and Human Development, National Reading Panel report). Washington, DC: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Child Health and Human Development.

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*, 91–121. doi:10.1080/10888438.2010.529219

Oakhill, J., Hartt, J., & Samols, D. (2005). Levels of comprehension monitoring and working memory in good and poor comprehenders. *Reading and Writing: An Interdisciplinary Journal*, *18*, 657–686.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*, 667–686.

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. H. Schunk & J. L. Meece (Eds.), *Students perceptions in the classroom* (pp. 149–183). Hillsdale, NJ: Erlbaum.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading. Past controversies, current deficits and an agenda for future research. *European Journal of Cognitive Psychology*, *16*, 3–26.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology*, *66*, 4429–4452.

Radach, R., Schmitten, C., Glover, L., & Huestegge, L. (2009). How children read for comprehension: Eye movements in developing readers. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The biological and behavioral foundations of reading comprehension* (pp. 75–106). New York, NY: Guilford.

Rapp, D. N., & van den Broek, P. (2005). Dynamic text comprehension: an integrative view of reading. *Current Directions in Psychological Science*, *14*, 276–279.

Rapp, D. N., van den Broek, P., McMaster, K., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading*, *11*, 389–312.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movments as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*, 241–255.

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*, 31–74.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 1290–1301.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology*, *65*, 177–206.

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

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., . . . Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading*, *3*, 69–94. doi:10.1207/s1532799xssr0301_4

Snow, C. E. (2001). *Reading for understanding*. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, *328*, 450–452. doi:10.1126/science.1182597

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

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

Tang, S., Reilly, R. G., & Vorstius, C. (2011). EyeMap: a software system for visualizing and analyzing eye movement data in reading. *Behavioral Research Methods*, *44*, 420–438.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, *40*, 7–26.

van den Broek, P., Risden, K., Fletcher, C. R., & Thorlow, R. (1996). A "landscape view" of reading: fluctuating patterns of activation and the construction of a stable memory representation. In B. K. Brtitton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 165–187). Mahwah, NJ: Erlbaum.

van den Broek, P., White, M. J., Kendeou, P., & Carlson, S. (2009). Reading between the lines: Developmental and individual differences in cognitive processes in reading comprehension. In R. K. Wagner, C. Schatschneider, & Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 107–123). New York, NY: Guilford.

van der Schoot, M., Reijntjes, A., & van Lieshout, E. D. M. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor reading comprehenders. *Reading and Writing*, *25*, 1665–1690. doi:10.1007/s11145-011-9337-4

van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Reijntjes, A. (2009). Lexical ambiguity resolution in good and poor comprehenders: An eye fixation and self-paced reading study in primary school children. *Journal of Educational Psychology*, *101*, 21–36.

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*, 8–25.

Vorstius, C., Radach, R., Mayer, M., & Lonigan, C. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, *42*, 191–206.

Wagoner, P. (1983). Comprehension monitoring: What it is and what we know about it. *Reading Research Quarterly*, *18*, 329–346.

Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, *89*, 420–432.

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

Zabrucky, K., & Moore, D. (1989). Children's ability to use three standards to evaluate their comprehension of text. *Reading Research Quarterly*, *24*, 336–352. doi:10.2307/747773

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York, NY: Touchstone

APPENDIX A

TABLE A1

Correlations and Descriptive Statistics for Assessments Used in the Structural Equation Models for Students ($n = 52$) in This Study, Which Were Not Significantly Different for the Students ($n = 396$) in the Larger Study

	LW	*PC*	*GMC*	*GMRV*	*AK*	*PV*	*SR*
Passage Comprehension (PC)	.54***						
Gates–MacGinitie Comprehension (GMC)	.509***	.722***					
Gates–MacGinitie Reading Vocabulary (GMRV)	.797***	.583***	.649***				
Academic Knowledge (AK)	.656***	.519***	.464**	.747***			
Picture Vocabulary (PV)	.606***	.561***	.460**	.658***	.780***		
Story Recall (SR)	.347*	.496***	.307*	.392***	.553***	.619***	
M SS/PR (*SD*)	100.26 (10.03)	95.53 (10.09)	34.26 (28.19)	45.62 (23.54)	90.19 (9.45)	97.12 (9.45)	97.21 (13.64)

Note. Gates–MacGinitie means provided as percentile rank (PR); all other means are standard scores (SS). Correlations were computed using equal interval IRT scale scores (i.e., W or ESS scores). LW = Letter-Word Identification test.

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

TABLE A2
HLM Results for Gaze Duration (GAZEDUR) and Rereading Time (REREAD) in ms

Fixed Effect	*Coefficient*	SE	t *Ratio*	*Approx.* df	p
M GAZEDUR, β_{00}	342.46	16.70	20.503	51	<.001
Implausible Word, β_{10}	64.04	15.55	4.118	290	<.001
Spring, β_{20}	−14.33	16.37	−0.876	290	.382
r_0 *SD* = 67.4 ($p < .001$), e *SD* = 102.0. ICC (between child) for unconditional model = .37.					
M REREAD, β_{00}	133.96	22.95	5.836	51	<.001
Implausible Word, β_{10}	93.95	19.84	4.735	290	<.001
Spring, β_{20}	−4.44	21.13	−0.210	290	0.834
r_0 *SD* = 104.8 ($p < .001$), e *SD* = 130.1. ICC (between child) for unconditional model = .38.					

Note. Means are fitted means (i.e., intercepts) in the models. Implausible Word = 1, plausible word = 0; For Session, Spring = 1; Fall = 0. Continuous variables are grand mean centered.

TABLE A3
Description of Interim Models

Gaze Duration. When Literacy was entered into the model and the interactions tested, results showed that the greater the child's Literacy *z* score, the shorter was Gaze duration (coefficient = –79.9 ms, $p < .001$). Children's Literacy did not predict differences in Gaze duration between plausible and implausible words. Their Gaze duration time increased from fall to spring (coef = 43.88, $p = .027$). Results were similar for Academic Language. Children with stronger academic language skills had shorter fitted means (–47.0, $p = .011$) and shorter Gaze duration in the spring (coef = 42.0, $p = .008$) compared to children with weaker Academic Language. Academic Language did not predict differences between plausible and implausible words.

Reread Time. Literacy score had no significant effect on the fitted mean Reread time, the difference between plausible and implausible words, or on session. Therefore interaction terms were not included in the final model. When Academic Language was considered, there was no significant effect of Academic Language on overall Reread times. However, the difference in Reread times for plausible and implausible targets were grader as students' fall than Academic Language was greater (coef = 66.44, $p = .005$). Therefore, for the final models, we included both Literacy and Academic Language as main effects in the model with interaction effects included for Academic Language but not Literacy (see Table 2). We ran the same model for Gaze duration and Reread time. Hence, all results control for children's Literacy *z* score.

APPENDIX B

Plausible and Implausible Target Words, Preceding Target, and Sentences

Pretarget	*Implausible*	*Plausible*	*Sentences*
wore	plant	dress	For the wedding Linda wore her best outfit. The colorful plant/dress was one of her favorites.
sawed	bolt	blade	Jeremy carefully sawed through a large oak tree in his yard. The steel bolt/blade became dull after only a few minutes.
flew	truck	plane	Last week Kyle flew to visit his family in another city. The large truck/plane was spacious and quickly transported them.
tossed	book	ball	Charles and his friends tossed the new toy outside in the yard. They were upset when the book/ball went over the fence.
washed	broom	sponge	Marcus washed dishes every night to earn his allowance. The new broom/sponge was great for getting rid of the grit.
printed	pencil	paper	Before school Jack printed his report for history class. He became angry when the pencil/paper got stuck.
barked	kitten	puppy	Every day Rover barked at the passing animals on the street. He was the most alert kitten/puppy in the neighborhood.
pounded	handsaw	hammer	In the evening Nicholas pounded two boards together in his garage. His new handsaw/hammer was a really useful tool.
sipped	plate	glass	Tyler cautiously sipped his fresh sweetened tea. He dropped the plate/glass because it was so hot.
slept	boots	sheets	Last night Bobby slept very well for many hours. His new boots/sheets were cozy and comfortable.
rowed	saddle	paddle	Tim and Landon rowed along the river in the park. Tim lost his saddle/paddle in the middle of the river.
threw	jogging	football	Justin threw the ball to his friend during the game. He liked jogging/football more than any other sport.
sewed	wrench	needle	Yesterday evening Jenny sewed patches onto her jeans. Sadly she lost the wrench/needle and could not finish.
stomped	hand	foot	When she got mad Sarah stomped on the floor. It really hurt when her hand/foot hit the ground with such force.
arrested	minister	officer	Today Sean arrested an unruly criminal as everyone watched. As an experienced minister/officer he quickly took control of the situation.
read	movie	novel	Amanda sat outside and read about a man named Arthur. She loved the movie/novel about ancient times.
dug	blower	shovel	To plant a tree Shannon dug a large hole in the ground. Her sturdy blower/shovel helped make the job easy.
poured	shoes	kettle	Michelle poured a fresh cup of coffee for her friend. She dropped the shoes/kettle and got very annoyed.
rode	kayak	taxi	Today Janet rode with her friends to the mall. The yellow kayak/taxi quickly got them to their favorite places.
sailed	bike	boat	They were all happy as they sailed along the coast. The swift bike/boat raced near the beautiful beach.

Note. The Zeno Word List (Zeno et al., 1995) was used to select pretarget and target words. All words are at less than a fifth-grade level and the target words in sentence 2 (columns 2 and 3) are matched for word length, morphological complexity, and frequency.