Published in final edited form as:
Learn Disabil Res Pract. 2014 February 1; 29(1): 25–35. doi:10.1111/ldrp.12027.

Novel Approaches to Examine Passage, Student, and Question Effects on Reading Comprehension

Amanda C. Miller[1], **Nicole Davis**[2,*], **Jennifer K. Gilbert**[2,*], **Sun-Joo Cho**[2], **Jessica R. Toste**[3], **James Street**[2], and **Laurie E. Cutting**[2]

[1]Regis University

[2]Vanderbilt University

[3]The University of Texas at Austin

Abstract

Reading comprehension is influenced by sources of variance associated with the reader and the task. To gain insight into the complex interplay of multiple sources of influence, we employed crossed random-effects item response models. These models allowed us to simultaneously examine the degree to which variables related to the type of passage and student characteristics influenced students' (*n* = 94; mean age = 11.97 years) performance on two indicators of reading comprehension: different types of comprehension questions and passage fluency. We found that variables related to word recognition, language, and executive function were influential across various types of passages and comprehension questions and also predicted a reader's passage fluency. Further, an exploratory analysis of two-way interaction effects was conducted. Results suggest that understanding the relative influence of passage, question, and student variables has implications for identifying struggling readers and designing interventions to address their individual needs.

Comprehension ability is generally thought to be a stable reflection of a reader's cognitive strengths and weaknesses, but recent studies have shown that this "ability" can vary across passages and across assessments (Keenan & Meenan, in press), suggesting that additional factors are also at play. Thus there is a push to reconceptualize reading comprehension as the product of the complex interaction between the reader and the demands associated with a specific reading task (Compton, Miller, Gilbert, & Steacy, 2013; Cutting, Benedict, Broadwater, Burns, & Fan, 2013; Snow, 2002). Such task demands include characteristics of the passage (e.g., genre, topic, complexity) and outcome measures of reading comprehension, including the types of questions used to assess comprehension, as well as passage fluency, which is considered to reflect overall reading competence (Fuchs, Fuchs, Hosp, & Jenkins, 2001). Investigating the interplay of student- and task-related factors provides insight into how these sources influence reading comprehension, and perhaps most importantly, may alert practitioners to the key characteristics of both the task and student that most likely contribute to comprehension difficulties. Additionally, gaining insight into the factors that impact comprehension has important implications for identifying struggling readers and designing interventions to address their individual needs. The goal of the present study was to explore the simultaneous influence of passage-level (i.e., characteristics of the text) and student-level influences (i.e., cognitive skills) on various indices of reading comprehension, including (a) multiple-choice questions that tap different aspects of

CONTACT AUTHOR: Laurie E. Cutting, laurie.cutting@vanderbilt.edu.
*indicates that the second and third authors contributed equally to this work and are listed in alphabetical order.

comprehension, and (b) passage fluency, which serves as a proxy for online comprehension (Eason, Sabatini, Goldberg, Bruce, & Cutting, 2013; Fuchs et al., 2001).

Passage-Level Factors

One widely researched passage-level factor is text cohesion, or the extent to which ideas conveyed in the text are explicitly connected (Graesser, McNamara, Louwerse, & Cai, 2004). Typically, the more cohesive the text, the more likely it will be comprehended. By definition, cohesive texts explicitly connect the text ideas for readers, while less cohesive texts require readers to form the connections on their own (Britton, Gulgoz, & Glynn, 1993). Other passage-level factors, such as decodability (Compton, Appleton, & Hosp, 2004), syntactic complexity (Gibson & Warren, 2004), and vocabulary (Freebody & Anderson, 1983) influence comprehension question accuracy and/or passage fluency, but relative to cohesion these constructs have received less attention in the literature.

To date, no study has examined multiple types of passage-level manipulations in the same experiment; such a design may provide insight into the interaction between text features and reader characteristics. To address this gap in the literature, the present study examined passage-level factors by manipulating features of a set of experimental “baseline” passages along four dimensions: cohesion, decodability, syntactic complexity, and vocabulary, while holding constant other important passage characteristics, such as word concreteness, word frequency, and sentence length. Performance on these four manipulated passages was then compared to the “baseline” passages. The ultimate goal of this line of work is to understand whether a given student’s relative performance across different types of passages can help identify the precise nature of the student’s reading comprehension difficulties.

Student-Level Factors

Reading comprehension requires the coordination of a number of cognitive skills. Readers must (a) identify and assign meaning to the words on the page, which requires decoding (Perfetti, 1985), morphological knowledge (Carlisle, 2000), and vocabulary knowledge (Elleman, Lindo, Morphy, & Compton, 2009); (b) remember previously read information as they read and integrate new ideas, which draws upon executive function (EF; Eason et al., 2013; Kintsch & Rawson, 2005), and (c) employ inference skills (Cain & Oakhill, 1999), as well as relevant background knowledge (Kintsch & Rawson, 2005), to infer connections that are not explicitly stated. These various indices reflect three general domains: word-level, language, and EF. While each of these domains have been studied in relation to reading comprehension, few studies have examined how different types of questions, for example, may interact with a student’s functioning in the three domains. Such information could inform valuable approaches to target struggling readers’ greatest needs.

Comprehension Indices: Question Type and Passage Fluency

Comprehension questions tap various levels of understanding. Literal questions assess knowledge of information stated explicitly in the text, while other types of questions assess the reader’s use of inferences and reading strategies. Different types of questions are known to differentially draw upon various cognitive skills, even when the questions test knowledge of the same passage (Eason, Goldberg, Young, Geist, & Cutting, 2012). Many studies have compared literal and inferential questions (e.g., Cain & Oakhill, 1999), but studies that have examined knowledge of text type and structure, reading strategies, and the readers’ ability to reflect on their own understanding are less common. For example, in a previous study (Eason et al., 2012) it was found that EF contributed uniquely to a reader’s ability to answer critical analysis and strategy questions, but not literal questions. The results of this study suggest that the depth of comprehension of the text differentially draws on various cognitive skills.

Passage fluency can be considered a proxy for online reading comprehension, as gaps in comprehension have a negative impact on passage fluency. For example, readers slow their reading rate when they encounter comprehension obstacles such as ambiguous referents (Gernsbacher, 1990) or conflicting information (O'Brien, Rizzella, Albrecht, & Halleran, 1998). Slower reading provides extra processing time that enables readers to resolve the ambiguity and repair breaks in comprehension (Graesser & McNamara, 2011). Importantly, passage fluency is a robust predictor of reading comprehension (e.g., Cutting, Materek, Cole, Levine, & Mahone, 2009; Eason et al., 2013). For example, Eason et al. (2013) showed that passage fluency accounted for up to 28% of the variance in reading comprehension over and above that accounted for by isolated word fluency measures. To our knowledge, no study has experimentally manipulated passage-level characteristics and then examined the effect these manipulations have on passage fluency; such insights may be important for understanding how measures of reading fluency, which are quick and efficient measures of overall reading ability, could potentially yield insights into readers' abilities with various text types.

Cross-Level Interactions

A number of studies have examined the interaction of passage-, student-, and/or question-level characteristics (e.g., Eason et al., 2012; Kendeou & van den Broek, 2007) and have contributed to the understanding of the complexities of comprehension. However, few studies have examined multiple types of passages and student characteristics in conjunction with multiple indices of comprehension (i.e., different question types and/or passage fluency). Such investigation may be critical to understanding why some students can comprehend despite relatively poor word recognition skills, or conversely, why some students with solid word recognition skills nevertheless struggle with comprehension (Aaron, Joshi, & Williams, 1999). Evidence suggests that to fully understand reading comprehension, consideration of the interaction between text and reader is essential; not all texts are universal in the demands they place on the reader's prior knowledge, word recognition, language, and/or EF abilities (Eason et al., 2012). Teasing apart the interactions between text and student factors has implications for the assessment and identification of struggling readers. This insight may also inform instruction and intervention practices to better help students comprehend different types of texts, an issue that has become especially important to consider with the new Common Core standards in place, which emphasize various genres of text (Porter, McMaken, Hwang, & Yang, 2011).

Although important to do, examining the interplay between passage and reader using different indices of comprehension (question type and passage fluency) is confronted by the methodological challenge of simultaneously assessing the contribution of three levels of predictors—for example, individual questions that assess knowledge of individual passages read by individual readers. Typically, researchers run separate models to explore such relationships and, therefore, lose the ability to partition variance in reading comprehension ability without aggregating at the level of student or text (Compton et al., 2013). A relatively new analytic approach, crossed random-effect item response modeling, enables researchers to simultaneously explain variability in comprehension with student-, passage-, and question-level characteristics (Compton et al., 2013; De Boeck & Wilson, 2004) and also provides less-biased estimates and standard errors than traditional regression models (Baayen, Davidson, & Bates, 2008). In the present study, we employ this method of analysis.

Research Questions and Hypotheses

1. Is comprehension (assessed by questions and passage fluency) affected if passages have less cohesion, more difficult decoding, more difficult syntax, or more advanced vocabulary as compared to baseline passages? We hypothesized that relative to the baseline passages, each of the manipulated passages would result in fewer questions answered correctly and slower reading rates.
2. Can individual differences in the probability of correctly answering questions and reading fluency be explained by student-level characteristics? We expected each of seven student factors to make a significant contribution to answering comprehension questions and reading fluency.
3. Is there a difference in the probability of correctly answering a comprehension question when it is a critical analysis, comprehension monitoring, interpretation, or reading strategy question as compared to a literal question? We hypothesized that responses to the literal questions would have higher probabilities of being correct than responses to the other types of questions.
4. Can exploratory analyses examining interactions between EF, a relatively understudied area within reading comprehension, and those variables hypothesized to be most closely linked to EF, such as passage cohesion and the higher-level questions (i.e., strategy, comprehension monitoring, inferential questions) provide proof-of-concept for the need to examine interactions and guide future studies of reader-text interactions?

Method

Participants & Procedures

We advertised the study in schools, clinics, and pediatricians' offices; 94 native English-speaking students participated (44 males; M = 11.97, SD = 1.27 years). The sample represented a normal distribution of reading ability, including children who experienced reading difficulties and typically developing children. Testing occurred in two sessions. Session 1 included a battery of cognitive and academic achievement tests; session 2 included a 1-hour reading comprehension task. Due to administration procedure changes, passage fluency data were available for 91 students and passage comprehension data were available for 85 students. Table 1 contains descriptive statistics and behavioral measures for all 94 students, including standard scores when possible to provide information on the sample's performance and general functioning.

Measures of Explanatory Variables

Literacy—*Isolated Word Reading Fluency* was measured by Sight Word Efficiency from the Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999). It measures the number of words a person reads correctly in 45 seconds.

Language—*Inferencing* was measured by the Making Inferences subtest from the Test of Language Competence-Expanded Edition (Wiig & Secord, 1989). Participants listen to a scenario (also presented in print) and deduce what may have happened. *Vocabulary* was measured by the Test of Word Knowledge (Wiig & Secord, 1989) Expressive Vocabulary subtest, which taps expressive vocabulary breadth and depth by asking participants to provide names for pictures. *Morphological Knowledge* was measured by the Morphological Relatedness Test (Mahony, Singson, & Mann, 2000). It measures awareness of derivational relations in third through sixth graders and consists of 40 items (half administered orally and the other half written). Each item consists of two words and participants indicate whether the words are related (e.g., “happy—happiness” (YES); “cat—catalogue” (NO)).

Executive Functioning—*Planning/Organization* was measured by the Delis Kaplan Executive Functioning System Tower Test (Delis, Kaplan, & Kramer, 2001), which assesses planning and organizing skill. The participant moves five discs across three pegs to match a visual model in as few moves as possible while adhering to a specific set of rules. *Non-Verbal Reasoning* was measured by the Analysis-Synthesis from the Woodcock Johnson-III (Woodcock, McGrew, & Mather, 2001) and assesses inferential reasoning skill. The participant solves a series of non-verbal problems by using various “keys” or combinations of colored square boxes that make other colors. The “keys” provide the basis for solving the problem. *Working Memory* was assessed by the Sentence Span Task (Daneman & Carpenter, 1980). The examiner reads a set of sentences, each missing its last word. The participant provides a word that makes sense in the sentence. At the end of the set, the child repeats the provided words in the administration order.

Experimental Reading Comprehension Task

Each participant read 12 expository texts about science or animals, each approximately 350 words. To assure that all baseline passages were equivalent in cohesion and syntax, each passage was entered into Coh-Metrix (Graesser et al., 2004), a computational tool designed to analyze and measure text characteristics, including surface-level (e.g., part of speech and tense) and higher-level characteristics (e.g., inferences, cohesiveness, narrativity) that are not indexed by standard text difficulty metrics, such as Flesch-Kincaid (see Graesser & McNamara, 2011). A bootstrap analysis was performed on 23 Coh-Metrix indices, in which each passage was removed from the group and its Coh-Metrix scores were compared to the means of the remaining 11 passages. Equivalency was defined as a mean score within a 90% confidence interval of the mean of the remaining passages on all 23 Coh-Metrix measures. The resulting 12 baseline passages were at a Flesch-Kincaid grade-level between 4.0 and 4.9. Of these 12 baseline passages, eight passages were manipulated on one of four dimensions (cohesion, decoding, syntax, or vocabulary) to make the text more difficult on that specific dimension. Cohesion-manipulated passages required increased inferential activity because the argument overlap between propositions was decreased and information that triggered causal connections was removed (Vidal-Abarca, Martinez, & Gilabert, 2000). Decoding-manipulated passages increased decoding difficulty by replacing regular open-class (content) words with irregular alternatives (e.g., *debris* in place of *rubbish*). Syntax-manipulated passages increased syntactic complexity by replacing simple with complex syntactic devices, while maintaining the logical and inferential relationships among propositions. Vocabulary-manipulated passages replaced high-frequency words with less familiar ones, while maintaining equivalent decodability. The same procedures described above to establish the baseline passage equivalency were used to provide a statistical measurement of increased difficulty within the manipulated dimension, while all other aspects of the text were held constant. To validate the results of the bootstrap analyses and the manipulated texts, all texts were piloted with an adult sample. The baseline passage words per minute (WPM) were compared to manipulated passage WPM (Street, Davis, Benedict, Harris, & Cutting, 2011). Passages were revised until all analyses (Coh-Metrix and adult WPM) indicated equivalency. In total, there were 20 unique passages: 12 baseline, 2 cohesion, 2 decoding, 2 syntax, and 2 vocabulary. During the testing session, participants read 4 manipulated and 8 baseline texts. Passage administration order was counterbalanced.

Passage Fluency—Passage fluency was used to measure overall reading competence (Fuchs et al., 2001). Students were timed as they read the passages aloud, and the test administrators marked mispronunciations, insertions, and omissions as errors. From this information, a count of correct WPM was calculated for each passage for each participant.

Comprehension Questions—After reading the texts aloud, participants answered multiple-choice questions about each passage for a total of 108 questions. Each passage had approximately nine questions, representing five different categories: (a) *Literal*, which measured comprehension of explicitly stated ideas; (b) *Critical analysis*, which required students to determine the author's purpose, recognize types of text, and discern patterns of organization in selections; (c) *Inferential*, which measured students' ability to make inferences, predictions, draw conclusions, and understand the central ideas; (d) *Comprehension Monitoring*, which measured attention to text, particularly inconsistencies in text; and (e) *Strategy*, which measured knowledge of text type, structure, and reading strategies. Answers were scored as correct (1) or incorrect (0).

Data Analyses

To examine passage manipulation effects, four dummy codes were used in both analyses (comprehension questions and fluency) to represent the five passage types (leaving baseline passages as the referent type): cohesion-manipulated, decoding-manipulated, syntax-manipulated, and vocabulary-manipulated. For only the comprehension question outcome, another set of four dummy codes was used to represent the five question types (leaving literal questions as the referent type): critical analysis, inferential, comprehension monitoring, and strategy. Additionally, raw scores representing student characteristics were entered in both analyses to explain individual differences in passage comprehension and fluency. In both the comprehension question and fluency models, all continuous variables were mean-centered. Model comparisons determined the importance of slope random effects. Each random component of slopes was tested to account for the possibility that the effect of a given variable varied across another level of data.

Comprehension questions—Passage, question, and student effects on reading comprehension were analyzed with a crossed random-effects item response model. Responses (to the individual comprehension questions; scored 0 or 1; R = 9168) were crossed between readers (J = 85) and questions (Q = 108), and questions were crossed with passages (P = 22), and passages were nested in topics (T = 12). Twelve responses were missing at random, resulting in 9168 instead of the full 9180 response data. The linear mixed-effects models (lmer) function of the *lme4* R package (Bates, Maechler, & Bolker, 2013) was used to fit models. Before building models to answer the research questions by adding in various fixed effects, we ran a set of competing unconditional models (i.e., a model without covariate; measurement models) to determine whether student, question, passage, and topic random effects were necessary for fitting the data. Three models were estimated and compared in terms of standard benchmarks in mixed-effect modeling (AIC, BIC, and the likelihood ratio test based on a mixed Chi-square distribution), all of which supported an unconditional model with only student and question random effects; passage and topic random effects were not included because they showed small variations across passages and topics in prior models, suggesting that the procedure used to create and manipulate passages was sufficient in controlling variance related to the passage topic. Conditional models (i.e., models with covariates) were fit to answer the four research questions. A combined model was also estimated to assess each type of effect (passage, question, and student) controlling for the others. The combined model, therefore, included passage covariates (four dummy variables comparing baseline passages to cohesion-manipulated, decoding-manipulated, syntax-manipulated, and vocabulary-manipulated passages), question covariates (four dummy variables comparing critical analysis, inferential, comprehension monitoring, and strategy questions to literal questions), and student covariates (planning/organization, inferencing, isolated word reading fluency, morphology, reasoning, vocabulary, and working memory) plus the control variable of age.

Reading fluency—To assess passage and student effects on passage fluency having a continuous outcome, data were analyzed with a crossed random-effects linear model. Responses (WPM; $R = 1091$) were crossed between readers ($J = 91$) and passages ($P = 22$) and passages were nested in topics ($T = 12$). Because of the various sources of potential dependency in the outcome, random variance components were considered in the model to account for dependency at the level of student, passage, and topic. Just as in the comprehension question models, unconditional models were compared to determine whether student, passage, and topic random effects were necessary for fitting the data. Likelihood ratio tests confirmed that both student and passage random effects were necessary to adequately fit the data; as with the comprehension questions, topic random effects were not included because they showed a small variation across topics in the prior model. One conditional model was considered for the research question related to passage effects and one related to person effects. A combined model was also considered to assess both types of effects (passage and student) controlling for the other type and controlling for age. The combined model, therefore, included passage covariates (four dummy variables comparing baseline passages to cohesion-manipulated, decoding-manipulated, syntax-manipulated, and vocabulary-manipulated passages) and student covariates (planning/organization, inferencing, isolated word reading fluency, morphology, reasoning, vocabulary, and working memory) plus age as a control variable. Markov chain Monte Carlo (MCMC) simulation-based Bayesian confidence intervals (a.k.a., credibility intervals), which do not rely on any distributional assumptions, were calculated for the fixed effects in the passage fluency results using the R package *language* (Baayen, 2009). By default, 10,000 simulations based on re-sampling are run and 95% Bayesian confidence intervals are constructed from the simulated samples. One advantage of simulation-based hypothesis testing is that the validity of results does not rely on any distributional assumptions (Efron & Tibshirani, 1993).

Results

Question 1: Passage Effects

Our first research question concerned whether manipulated passages had lower probabilities of correct responses to comprehension questions and/or slower reading times (WPM), relative to baseline passages. For the comprehension questions, the *z*-values in column 1 of Table 2 indicate that there were no significant differences between correct answers from manipulated vs. baseline passages. The *z*-values in column 4 of Table 2 (combined model) show that even when controlling for question and student effects, questions from manipulated passages did not have significantly different logit (and therefore probabilities) of being answered correctly than questions from baseline passages.

Results examining WPM as an outcome measure revealed that all manipulated passages on average were read more slowly than the baseline passages, with two manipulated passage comparisons reaching the statistical significance level of $p < .05$ (see columns 1 and 3 of Table 3). In the combined model where effects of person characteristics were controlled (including age), the decoding-manipulated and vocabulary-manipulated passages showed a significant difference between WPM from decoding-manipulated vs. baseline passages and vocabulary-manipulated vs. baseline passages. Controlling for other student and passage effects, there were, on average, 13.32 fewer WPM for the decoding-manipulated passages compared to the baseline passages and 8.56 fewer WPM for the vocabulary-manipulated passages compared to the baseline passages. Neither the cohesion- nor the syntax-manipulated passages produced a statistically significant different number of WPM compared to baseline passages.

Question 2: Student Effects

Seven student characteristics were entered into the model to assess the contribution of individual differences to both indices of comprehension. Columns 3 and 4 of Table 2 show that controlling for the effects of student, question, and passage characteristics, participants with higher isolated word reading fluency, morphology skills, reasoning skills, vocabulary, and working memory had higher probabilities of correct responses to comprehension questions ($p < .05$). No differences were detected for varying levels of planning/organization or inferencing skills.

For fluency (columns 2 and 3 of Table 3), the 95% Bayesian confidence intervals show that controlling for age as well as the other effects of student and passage characteristics, students with better planning/organization and isolated word reading fluency read more WPM than students with lower skills in those areas; morphology approached significance. No differences in passage fluency were detected for varying levels of inferencing, reasoning, vocabulary, or working memory.

Question 3: Question Effects

We also examined whether questions that required more than just recalling literal information from the text would be more difficult to answer than literal questions, as indicated by columns 2 and 4 in Table 2. While all question types had lower probabilities of a correct answer than literal questions, the combined model indicates that significant effects were detected only for inferential questions and reading strategy questions, such that both were associated with lower probabilities of a correct answer than literal questions. No significant differences were found between literal and critical analysis or comprehension monitoring questions.

Question 4: Exploratory Interaction Analyses with EF

Although we recognize that our models have insufficient power to detect interactions, we ran exploratory “proof-of-concept” models that examined the potential interaction between each of the three EF variables (working memory, reasoning, and planning/organization) and (a) passage cohesion (relative to baseline), and (b) inferential, reading strategy, and comprehension monitoring questions (all relative to literal questions). Only one of these interactions, EF (planning/organization) by reading strategy questions, was statistically significant at $p < .05$. Contrary to expectation, this interaction suggested that for students with poorer planning/organization skills, there was only a minimal difference on performance between reading strategy questions and factual questions, but for students with higher planning/organization skills, there was a bigger difference in performance between those two question types, with literal questions having a higher probability of correct response than reading strategy questions. Although not reaching the traditional levels of significance, interaction analyses also revealed that those with poorer reasoning abilities had a higher probability of answering questions correctly for the cohesion manipulated text as compared to baseline passages ($p = .09$).

Discussion

A reader’s comprehension ability reflects the relative influence of multiple factors associated with both the reader and the task (Compton et al., 2013; Cutting et al., 2013; Snow, 2002). The current study reports results from an experimental reading assessment designed to provide insight into passage and reader characteristics that influence reading comprehension, as indicated by: (a) performance on five types of comprehension questions, each designed to tap different aspects of comprehension, and (b) reading fluency, a proxy for online reading comprehension. More specifically, we employed a relatively new statistical approach, crossed random-effects item response models, that enabled us to simultaneously examine the effects of (a) *passage type* (comparing baseline passages to cohesion-, decoding-, syntax-, and vocabulary-manipulated passages), (b) *student characteristics* (age, isolated word reading fluency, vocabulary, inferencing, morphological skill, and three aspects of EF: planning/organizing, working memory, and non-verbal reasoning), and (c) *question type* (comparing critical analysis, inferential, comprehension monitoring, and strategy questions to literal questions) on comprehension of expository passages.

Passage Effects

To understand the influence of passage characteristics on reading comprehension, measured by questions and passage fluency, we created a set of "baseline" passages and then manipulated specific aspects of these passages to test how passage characteristics impact comprehension processing when carefully controlling for influences related to topic. Findings revealed statistically significant effects for the decoding and vocabulary manipulations on passage fluency independent of topic effects, such that relative to baseline passages, participants' passage fluency slowed when they read passages that were more difficult to decode or included more difficult vocabulary. These findings demonstrate that word-level passage features (decoding and vocabulary) can influence a reader's passage fluency even when carefully controlling for higher-level text characteristics, such as cohesion and syntax, and for student-level differences, such as isolated word reading ability and individual differences in vocabulary knowledge. This finding also supports previous research demonstrating the influence of semantic skills (vocabulary ability) on passage fluency (Eason et al., 2013). Unlike previous studies, we did not find that cohesion and syntax impacted comprehension at a statistically significant level. One potential explanation is that our cohesion- and syntax-manipulated passages carefully controlled other text features on multiple dimensions (e.g., vocabulary and decodability). Nevertheless, although not reaching statistical significance, syntax- and cohesion-manipulated passages were read more slowly than baseline passages and were associated with a lower probability of answering comprehension questions correctly. Thus, even though not all of the passage manipulations revealed statistically significant differences from baseline, the manipulated passages were consistently associated with poorer performance. As such, we are careful not to draw conclusions from these null findings as they could reflect relatively low statistical power.

Student Effects

A number of student-level characteristics were significant predictors of both comprehension questions and passage fluency. Students with higher isolated word reading fluency, morphology skills, vocabulary, and EF (reasoning skills and working memory) had higher probabilities of correct responses on the comprehension questions. These results are largely consistent with previous literature and are important because they demonstrated that these student-level factors remained significant predictors of reading comprehension across the different types of passages, questions, and topics. These findings provide additional support for theoretical models of text processing that posit that word reading, language skills, and EF are involved in the development of a reader's mental text representation, especially with longer texts (Kintsch, 1998) and texts expository in nature (Eason et al., 2012).

With respect to student-level factors that influence passage fluency, students with higher EF (planning/organization) and isolated word reading fluency read faster than students with lower skills in those areas. One potential interpretation for the finding that EF predicted passage fluency is that EF enables the reader to form a coherent mental representation of the text by facilitating the integration of incoming and previously read text ideas (Cooke, Halleran, O'Brien, 1998), activating relevant prior knowledge (Kintsch & Rawson, 2005), and inhibiting irrelevant knowledge (Gernsbacher, 1997). Having a coherent mental representation of the text may facilitate fluent word identification through the spreading of semantic activation.

Question Effects

Consistent with previous research, inferential and reading strategy questions were more difficult than literal questions (Eason et al., 2012). These findings held when simultaneously controlling for the other student- and passage-level effects. Both of these question types require the reader to go beyond text-based information and build connections among text ideas and prior knowledge. Contrary to our hypotheses, comprehension monitoring and critical analysis questions were no more difficult than literal questions. It may be that performance on these question types is closely linked to student characteristics, and differences between the question types are not evident once student predictors are controlled. Given the existing literature's suggestion that these higher-level questions reflect deeper comprehension, interactions between student predictors and these question types will be an important area of exploration for future research.

Finally, our exploratory "proof-of-concept" interaction results suggested a potential interaction between EF and reading strategy questions and cohesion-manipulated passages. Interestingly, both findings suggested that students with *poorer* EF performed *better* on reading strategy questions and cohesion-manipulated passages. Of course, these findings could simply be spurious given our low power and/or reflect flaws in our experimental design of passages; however, given that all our main effects were very consistent with the literature, flaws in passage design do not appear to be the straightforward answer. Additionally, both interaction findings were counterintuitive in the same direction, perhaps suggesting that the results were not simply spurious. As such, we view the results of our exploratory interactions as intriguing and illustrating the need to explore not only the main effects of, but the interplay between passage, reader, and question characteristics. Ultimately, three-way interactions may explain these findings. Here we are reminded of McNamara, Kintsch, Songer, and Kintsch (1996), who found a counterintuitive interaction in which high knowledge readers showed better comprehension on low-cohesion passages than high-cohesion passages, and low-knowledge readers showed the opposite pattern. Later, O'Reilly and McNamara (2007) showed that this interaction only held for less-skilled readers and concluded from this three-way interaction that less-skilled readers are less likely to make knowledge-based inferences, but when demanded to do so by a less-cohesive text, inferencing improves.

Implications and Future Directions

Conceptualizing comprehension as the complex interaction of multiple sources of influence has implications for the assessment and identification of struggling readers. Standardized reading comprehension assessments vary widely in their passage length, genre, topic, and format and consequently capture different cognitive processes (Cutting & Scarborough, 2006; Eason et al., 2012; Keenan, Betjemann, & Olson, 2008). Practitioners should be cognizant of the fact that even though a student may not demonstrate difficulties on one comprehension assessment, he may in fact present comprehension difficulties when assessed by a different instrument (Keenan & Meenan, in press). Thus, practitioners should consider these dynamics when selecting comprehension assessments for identification purposes and be aware that multiple levels of influence may be at play; any one factor or the interaction of several factors could artificially inflate or deflate the student's "true" comprehension ability. Moreover, it is critical that educators be aware that two students who are observed as having comprehension difficulties may likely have very different learning profiles and benefit from individualized interventions that target their areas of greatest need. Furthermore, interventions should prepare students to employ different strategies depending on the text's features (Ozuru, Briner, Best, McNamara 2010). Together this insight could prove influential in both identifying struggling comprehenders and designing interventions to ameliorate their difficulties. Uncovering the complexities of comprehension, and in particular the complex interactions among passage, reader, and question profiles will help determine how best to provide an intervention and for whom the intervention will likely be most effective.

Furthermore, the present study demonstrates the use of crossed random-effects item response modeling as a viable statistical method and forward-thinking approach to measure the complexities of reading comprehension. Although statistical power to detect interactions in the present study was limited, future studies should adopt this analytic approach to tease apart interactions across multiple sources of influence and also consider student characteristics not included in the present study, such as prior knowledge of the passage topic (Miller & Keenan, 2009) or a reader's standards of coherence (van den Broek, Bohn-Gettler, Kendeou, & Carlson, 2011). Additionally, future work should consider characteristics of the activity or task demands, such as instructions (McNamara & Dempsey, 2011) or reader goals (van den Broek, Lorch, Linderholm, & Gustafson, 2001). In summary, a number of influences contribute to the manifestation of reading comprehension difficulties, and ultimately understanding these complex interactions will aid in identifying students with comprehension difficulties and designing interventions that best facilitate the development of comprehension skills.

Acknowledgments

This work was supported by NICHD RO1HD067254, NICHD R01HD044073, Vanderbilt Kennedy Center grant 5P30HD015052-31, and VICTR Resources grant.

References

Aaron PG, Joshi RM, Williams KA. Not all reading disabilities are alike. Journal of Learning Disabilities. 1999; 32:120–137. [PubMed: 15499713]

Baayen, RH. R package version 1.4. 2009. languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics".

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language. 2008; 59:390–412.

Bates, D.; Maechler, M.; Bolker, B. R package version 0.999999-0. 2013. lme4: Linear mixed-effects models using S4 classes.

Britton, BK.; Gulgoz, S.; Glynn, S. Impact of good and poor writing on learners: Research and theory. In: Britton, BK.; Woodward, A.; Binkley, MR., editors. Learning from textbooks: Theory and practice. Hillsdale, NJ: Erlbaum; 1993. p. 1-46.

Cain K, Oakhill JV. Inference making ability and its relation to comprehension failure in young children. Reading and Writing: An Interdisciplinary Journal. 1999; 11:489–503.

Carlisle JF. Awareness of the structure and meaning of morphologically complex words: Impact on reading. Reading and Writing: An Interdisciplinary Journal. 2000; 12:169–190.

Compton DL, Appleton AC, Hosp MK. Exploring the relationship between text-leveling systems and reading accuracy and fluency in second-grade students who are average and poor decoders. Learning Disabilities Research & Practice. 2004; 19(3):176–184.

Compton, DL.; Miller, AC.; Gilbert, JK.; Steacy, LM. Using item response random-effects modeling techniques to explore the relations between reader and text characteristics in subtypes of poor comprehenders. In: Cutting, LE.; Miller, B.; McCardle, P., editors. Unraveling the behavioral, neurobiological, & genetic components of reading comprehension. Baltimore, MD: Brookes Publishing; 2013.

Cooke AE, Halleran JG, O'Brien EJ. What is readily available during reading? A memory-based view. Discourse Processes. 1998; 26:109–129.

Cutting, LE.; Benedict, S.; Broadwater, A.; Burns, S.; Fan, Q. From words to text: Neurobiology and behavioral bases of reading comprehension. In: Miller, B.; Cutting, L.; McCardle, P., editors. Unraveling the Behavioral, Neurobiological, and Genetic Components of Reading Comprehension. Baltimore: Paul Brookes Publishing; 2013. Forthcoming

Cutting LE, Materek A, Cole CAS, Levine TM, Mahone EM. Effects of fluency, oral language, and executive function on reading comprehension performance. Annals of Dyslexia. 2009; 59:34–54. [PubMed: 19396550]

Cutting LE, Scarborough HS. 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. 2006; 10(3):277–299.

Daneman M, Carpenter PA. Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior. 1980; 19(4):450–466.

De Boeck, P.; Wilson, M., editors. Explanatory item response models: A generalized linear and nonlinear approach. New York: Springer; 2004.

Delis, DC.; Kaplan, E.; Kramer, JH. The Delis-Kaplan Executive Function System. San Antonio, TX: Psychological Corp; 2001.

Eason SH, Goldberg LF, Young KM, Geist MC, Cutting LE. Reader-text interactions: How differential text and question types influence cognitive skills needed for reading comprehension. Journal of Educational Psychology. 2012; 104:515–528.

Eason SH, Sabatini J, Goldberg L, Bruce K, Cutting LE. Examining the relationship between word reading efficiency and oral reading rate in predicting comprehension among different types of readers. Scientific Studies of Reading. 2013; 17:199–223. [PubMed: 23667307]

Efron, B.; Tibshirani, R. An introduction to the bootstrap. Boca Raton, FL: Chapman & Hall/CRC; 1993.

Elleman AM, Lindo EJ, Morphy P, Compton DL. The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. Journal of Research of Educational Effectiveness. 2009; 2:1–44.

Freebody P, Anderson RC. Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. Reading Research Quarterly. 1983; 18(3):277–294.

Fuchs LS, Fuchs D, Hosp MK, Jenkins JR. Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading. 2001; 5:239–256.

Gernsbacher, MA. Language comprehension as structure building. Hillsdale, NJ: Erlbaum; 1990.

Gernsbacher MA. Two decades of structure building. Discourse Processes. 1997; 23:265–304.

Gibson E, Warren T. Reading time evidence for intermediate linguistic structure in long-distance dependencies. Syntax. 2004; 7:1:55–78.

Graesser AC, McNamara DS. Computational analyses of multilevel discourse comprehension. Topics in Cognitive Science. 2011; 3(2):371–398.

Graesser AC, McNamara DS, Louwerse MM, Cai Z. Coh-Metrix: Analysis of text on cohesion and language. Behavior Research Methods, Instruments, and Computers. 2004; 36:193–202.

Keenan JM, Betjemann RS, Olson RK. Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. Scientific Studies of Reading. 2008; 12:281– 300.

Keenan JM, Meenan CE. Test differences in diagnosing reading comprehension deficits. Journal of Learning Disabilities. in press.

Kendeou P, van den Broek P. The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. Memory & Cognition. 2007; 35(7):1567–1577. [PubMed: 18062535]

Kintsch, W. Comprehension: A paradigm for cognition. New York: Cambridge University Press; 1998.

Kintsch, W.; Rawson, KA. Comprehension. In: Snowling, MJ.; Hulme, C., editors. The science of reading: A handbook. Malden MA: Blackwell; 2005. p. 209-226.

Mahony D, Singson M, Mann V. Reading ability and sensitivity to morphological relations. Reading and Writing: An Interdisciplinary Journal. 2000; 12:191–218.

McNamara, DS.; Dempsey, K. Reader expectations of question formats and difficulty: Targeting the zone. IAP Information Age Publishing; Charlotte, NC: 2011. p. 321-352.

McNamara DS, Kintsch E, Songer NB, Kintsch W. Are Good Texts Always Better? Interactions of Text Coherence, Background Knowledge, and Levels of Understanding in Learning from Text. Cognition and Instruction. 1996; 14:1–43.

Miller AC, Keenan JM. How word reading skill impacts text memory: The centrality deficit and how domain knowledge can compensate. Annals of Dyslexia. 2009; 59:99–113. [PubMed: 19475514]

O'Brien EJ, Rizzella ML, Albrecht JE, Halleran JG. Updating a situation model: A memory-based text processing view. Journal of Experimental Psychology: Learning, Memory, & Cognition. 1998; 24:1200–1210.

O'Reilly T, McNamara DS. Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. Discourse Processes. 2007; 43:121–152.

Ozuru Y, Briner S, Best R, McNamara DS. Contributions of self-explanation to comprehension of high- and low-cohesion texts. Discourse Processes. 2010; 47(8):641–667.

Perfetti, CA. Reading ability. Oxford: Oxford University Press; 1985.

Porter A, McMaken J, Hwang J, Yang R. Common core standards: The new U.S. intended curriculum. Educational Researcher. 2011; 40:103–116.

Snow, CE. Reading for understanding: Toward a research and development program in reading comprehension by the RAND Reading Study Group. 2002. MR-1465-OERI

Street, JS.; Davis, N.; Benedict, SL.; Harris, HC.; Cutting, LE. Reader-text interactions and oral reading fluency. Poster presented at the Society for the Scientific Study of Reading Annual Meeting; 2011 Jul.

Torgesen, JK.; Wagner, RK.; Rashotte, CA. Test of Word Reading Efficiency. Austin, TX: ProEd; 1999.

van den Broek, PW.; Bohn-Gettler, C.; Kendeou, P.; Carlson, S.; White, MJ. When a reader meets a text: The role of standards of coherence in reading comprehension. In: McCrudden, MT.; Magliano, JP.; Schraw, G., editors. Text relevance and learning from text. Greenwich, CT: Information Age Publishing; 2011. p. 123-140.

van den Broek P, Lorch RF Jr, Linderholm T, Gustafson M. The effects of readers' goals on inference generation and memory for texts. Memory & Cognition. 2001; 29:1081–1087. [PubMed: 11913743]

Vidal-Abarca E, Martínez G, Gilabert R. Two procedures to improve instructional text: Effects on memory and learning. Journal of Educational Psychology. 2000; 92(1):107–116.

Wiig, E.; Secord, W. Test of Language Competence–Expanded. San Antonio, TX: Psychological Corp; 1989.

Woodcock, RW.; McGrew, KS.; Mather, N. Examiner's manual. Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing; 2001.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Descriptive Statistics of the Outcome Measures and Covariates

Level	Type	Variable	Mean	SD	Min	Max	Scale
Response (R = 9168)	Outcome measure	Passage-specific comprehension questions	.80	.40	0	1	Dichotomous (1=correct/ 0=incorrect)
Response (R = 1079)	Outcome measure	Passage-specific fluency	140.22	43.96	25.06	318	Raw words per minute
Student (J = 94)	Covariate						
		Planning/organization	10.02	2.30	4	15	Scaled
		Inferencing	9.27	2.62	4	15	Scaled
		Isolated word reading fluency	100.53	13.23	55	133	Standard
		Morphology	9.35	1.24	5	11	Raw
		Reasoning	108.23	15.37	75	151	Standard
		Vocabulary	9.35	3.40	3	17	Scaled
		Working memory	6.41	2.50	0	12	Raw
		Age	11.97	1.27	9.92	14.75	Raw in years
Passage (P =22)	Covariate						
			Freq.	%age	Min	Max	Scale
		Passage Type					
		Baseline*	14	63.64	0	1	Dummy
		Cohesion manipulated	2	9.09	0	1	Dummy
		Decoding manipulated	2	9.09	0	1	Dummy
		Syntax manipulated	2	9.09	0	1	Dummy
		Vocabulary manipulated	2	9.09	0	1	Dummy
Question (Q = 108)	Covariate						
		Question Type					
		Critical analysis	9	8.33	0	1	Dummy
		Comprehension mornitoring	12	11.11	0	1	Dummy
		Inferential	36	33.33	0	1	Dummy
		Literal*	35	32.41	0	1	Dummy
		Reading strategy	16	14.81	0	1	Dummy

*Referent category; left out of model.

Table 2

Results of Conditional Crossed Random-Effects Item Response Comprehension Question Models

	Models											
	Passage Model			Question Model			Student Model			Combined Model		
Fixed Effects	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***
Grand Mean												
Intercept	**1.92**	**0.16**	**12.38**	**2.26**	**0.21**	**10.70**	**1.92**	**0.12**	**16.16**	**2.36**	**0.20**	**11.94**
Passage Covariates												
Cohesion manipulated	−0.18	0.15	−1.20							−0.17	0.15	−1.18
Decoding manipulated	0.12	0.16	0.72							0.11	0.16	0.66
Syntax manipulated	−0.26	0.15	−1.74							−0.24	0.15	−1.59
Vocabulary manipulated	−0.08	0.15	−0.53							−0.07	0.15	−0.49
Question Covariates												
Critical analysis				−0.66	0.39	−1.68				−0.66	0.41	−1.58
Comprehension monitoring				−0.19	0.36	−0.52				−0.18	0.38	−0.49
Inferential				**−0.58**	**0.25**	**−2.30**				**−0.59**	**0.26**	**−2.24**
Reading Strategy				**−0.82**	**0.32**	**−2.58**				**−0.84**	**0.33**	**−2.52**
Student Covariates												
Planning/Organization							0.00	0.02	0.03	−0.00	0.02	−0.02
Inferencing							0.01	0.01	0.48	0.01	0.01	0.50
Isolated word reading fluency							**0.01**	**0.01**	**2.05**	**0.01**	**0.01**	**2.15**
Morphology							**0.14**	**0.05**	**2.68**	**0.14**	**0.05**	**2.66**
Reasoning							**0.06**	**0.02**	**3.10**	**0.06**	**0.02**	**3.04**
Vocabulary							**0.11**	**0.02**	**5.57**	**0.11**	**0.02**	**5.60**
Working memory							**0.09**	**0.03**	**2.93**	**0.08**	**0.03**	**2.81**
Age							−0.08	0.05	−1.54	−0.08	0.05	−1.57
Random Effects: Variances												
Intercept (Question)	1.13			1.00			1.16			1.11		
Morphology slope (Question)							0.01			0.00		
Reasoning slope (Question)							0.00			0.00		
Vocabulary slope (Question)							0.00			0.00		

NIH-PA Author Manuscript

NIH-PA Author Manuscript

	Models											
	Passage Model			**Question Model**			**Student Model**			**Combined Model**		
Fixed Effects	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***	**Est.**	**SE**	***z***
Intercept (Student)	1.00			1.00			0.16			0.16		
Decoding passage slope (Student)	0.66									0.60		

Note. Coefficients for fixed effects are in logit. Parameters with corresponding *z*-values > 1.96 are significant at the 5% significance level; these values have been bolded.

Table 3

Results of Conditional Crossed Random-Effects Item Response Passage Fluency Models

	Model					
	RQ1. Passage Model		RQ2. Student Model		Combined Model	
Fixed Effects	**Est.**	**95% CI**	**Est.**	**95% CI**	**Est.**	**95% CI**
Grand Mean						
Intercept	**143.66**	**(138.64, 148.31)**	**140.59**	**(136.55, 144.76)**	**143.66**	**(139.60, 147.94)**
Passage Covariates						
Cohesion manipulated	−6.19	(−15.10, 2.44)			−6.18	(−14.54, 2.85)
Decoding manipulated	**−13.32**	**(−21.94, −4.50)**			**−13.32**	**(−21.51, −4.67)**
Syntax manipulated	−5.88	(−14.37, 3.01)			−5.87	(−14.27, 2.96)
Vocabulary manipulated	−8.56	(−16.98, 0.19)			**−8.56**	**(−16.98, −0.19)**
Student Covariates						
Planning/Organization			**1.00**	**(0.07, 1.99)**	**1.00**	**(0.02, 1.97)**
Inferencing			0.73	(−0.13, 1.56)	0.73	(−0.14, 1.54)
Isolated word reading fluency			**2.11**	**(1.74, 2.46)**	**2.11**	**(1.76, 2.46)**
Morphology			3.04	(−0.25, 6.39)	3.03	(−0.46, 6.13)
Reasoning			0.84	(−0.11, 1.81)	0.84	(−0.09, 1.81)
Vocabulary			0.56	(−0.54, 1.68)	0.56	(−0.54, 1.71)
Working memory			0.05	(−1.71, 1.77)	0.06	(−1.73, 1.80)
Age			−1.29	(−4.04, 1.70)	−1.30	(−4.34, 1.46)
Random Effects: Variances						
Intercept (Passage)	25.10		39.80		24.86	
Intercept (Student)	1584.45		564.20		563.83	
Residual	303.93		303.94		303.97	

Note. Parameters with 95% Bayesian confidence intervals that do not include 0 are significant at at the 5% significance level; these values have been bolded.