Published in final edited form as:
J Educ Psychol. 2014 February ; 106(1): 162–180. doi:10.1037/a0033826.

The Effects of Student and Text Characteristics on the Oral Reading Fluency of Middle-Grade Students

Amy E. Barth, **Tammy D. Tolar**, **Jack M. Fletcher**, and **David Francis**
Department of Psychology and the Texas Center for Learning Disabilities, University of Houston.

Abstract

We evaluated the effects of student characteristics (sight word reading efficiency, phonological decoding, verbal knowledge, level of reading ability, grade, gender) and text features (passage difficulty, length, genre, and language and discourse attributes) on the oral reading fluency of a sample of middle-school students in Grades 6–8 (N = 1,794). Students who were struggling (n = 704) and typically developing readers (n = 1,028) were randomly assigned to read five 1-min passages from each of 5 Lexile bands (within student range of 550 Lexiles). A series of multilevel analyses showed that student and text characteristics contributed uniquely to oral reading fluency rates. Student characteristics involving sight word reading efficiency and level of decoding ability accounted for more variability than reader type and verbal knowledge, with small, but statistically significant effects of grade and gender. The most significant text feature was passage difficulty level. Interactions involving student text characteristics, especially attributes involving overall ability level and difficulty of the text, were also apparent. These results support views of the development of oral reading fluency that involve interactions of student and text characteristics and highlight the importance of scaling for passage difficulty level in assessing individual differences in oral reading fluency.

Keywords

oral reading fluency; middle school; text effects; struggling readers

Among elementary- and middle-grade readers, oral reading fluency plays a fundamental role in the comprehension of connected text (Catts, Adlof, & Hogan, 2005; Gough, Hoover, & Peterson, 1996; Juel, 1988). Oral reading fluency is moderately predictive of performance on reading comprehension measures (Hintze & Silberglitt, 2005; Reschley, Busch, Betts, Deno, & Long, 2009), with individual differences in oral reading fluency helping to account for individual differences in children who struggle to comprehend text (Catts et al., 2005; Cirino et al., in press; Hock et al., 2009; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003b). In addition, in latent variable studies, oral reading fluency is a construct that shows some independence from untimed reading accuracy measures and from comprehension measures (Cirino et al., in press).

Fluency and Comprehension: Theoretical Perspectives

Cognitive models of discourse processing suggest that multiple levels of language and discourse influence the rate at which text is processed (Gernsbacher, 1990; Graesser, Millis,

Correspondence concerning this article should be addressed to Amy E. Barth, University of Houston, Department of Psychology, 2151 West Holcombe, 222 Texas Medical Center Annex, Houston, TX 77204-5053. aebarth@uh.edu.

Supplemental materials: http://dx.doi.org/10.1037/a0033826.supp

& Zwaan, 1997; Kintsch, 1988). For example, the lexical quality hypothesis (Perfetti, 2007) and landscape model (van den Broek, Risden, Fletcher, & Thurlow, 1996) suggest that efficient text processing results when high-quality representations of words are retrieved easily, with their context-specific meanings reliably integrated into the existing mental model of text. The lexical quality hypothesis suggests that a word or its lexical representation is of high quality when it is semantic (i.e., meaning), orthographic (i.e., spelling), phonological (i.e., pronunciation), and grammar (grammatical class and morphosyntactic inflections) constituents are tightly coupled and available synchronously at the point of retrieval during reading (Perfetti, 2007; Perfetti & Hart, 2001). The landscape model builds on the lexical quality hypothesis and suggests that if lexical representations retrieved during reading are of low quality and jeopardize the construction of a coherent mental model of text, then comprehension monitoring processes are actively initiated in an attempt to maintain or repair coherence. Comprehension monitoring processes tap information from higher levels of cognitive processing, prior knowledge of the topic, general processing skill, and knowledge of text features in order to resolve comprehension misfires (Graesser & Clark, 1985; Graesser, Millis, & Zwaan, 1997; Kintsch, 1988; Trabasso & Magliano, 1996; van den Broek, Rapp, & Kendeou, 2005).

Taken together, cognitive models of discourse processing suggest that reading fluency is an interactive process and represents the ability to rapidly retrieve lexical representations of words and engage higher level cognitive processes at the appropriate level and speed for the *specific text being read*, which in turn facilitates the reader's construction of a coherent mental model of the text. The cognitive processes that support reading fluency in individual readers likely vary with text demands (passage difficulty, passage length, language and discourse features, and type of text being read). Thus, an important next step is to understand sources of individual differences (i.e., text, reader, and interactions of text and reader) in text processing to help to weight the contributions of these factors to the development of oral reading fluency. In the next sections, we review characteristics of the reader and features of the text that influence oral reading fluency.

Characteristics of the Reader

Sight word reading—Past research suggests that "sight" word reading efficiency accounted for 58%–82% of the variance in oral reading fluency (ORF) among elementary-grade readers (Torgesen, Rashotte, & Alexander, 2001). Among a representative sample of 527 students in Grade 8, Barth, Catts, and Anthony (2009) reported that the standardized factor loading of sight word reading efficiency on a latent ORF factor was 0.91. Recent research suggests that among middle-grade readers, the magnitude of relation between sight word efficiency and reading fluency might vary by reading ability and socioeconomic status, with the relation significantly higher among struggling readers (0.87) than typically developing readers (0.60) (Cirino et al., in press) and lower among poor comprehenders living in urban neighborhoods (0.41–0.55) (Braesseur-Hock, Hock, Kieffer, Biancarosa, & Deshler, 2011). Although the magnitude of the relation between ORF and reading comprehension may vary by reader ability and reader attributes, sight word efficiency plays a prominent role in ORF among middle-grade readers.

Phonological decoding—Recent research has also evaluated the unique contribution of phonological decoding to ORF among elementary-grade readers, with phonological decoding accounting for 2%–10% of the variance in reading fluency after controlling for sight word reading efficiency (Torgesen et al., 2001). For older readers, Barth et al. (2009) reported that phonological decoding uniquely accounted for 10% of the variance in ORF after controlling for working memory and nonverbal cognition. Cirino et al. (in press) reported that the latent correlations between phonological decoding and reading fluency was

0.73 among struggling readers and 0.57 among typically developing readers. However, Adlof, Catts, and Little (2006) reported a latent correlation of phonological decoding and reading fluency of 0.93. Although phonological decoding and reading fluency were highly correlated constructs, Adlof et al. indicated that the two constructs showed unique and reliable, nonoverlapping variance, thereby permitting them to be tested as separate constructs among middle-grade readers.

Verbal knowledge—When students read for understanding, word and world knowledge likely play a role in limiting or facilitating ORF rates. Among elementary-grade samples, verbal knowledge accounted for approximately 6%–9% of the variance in reading fluency (Torgesen et al., 2001). Among middle-grade readers, Barth et al. (2009) reported that a latent language comprehension variable uniquely accounted for 8.5% of the variance in reading fluency after controlling for working memory and nonverbal cognition. At the manifest level, the correlation between verbal knowledge, as measured by the Peabody Picture Vocabulary Test– Revised (Wiederholt & Bryant, 1982), and reading fluency, as measured by the Gray Oral Reading Test–3 (Dunn & Dunn, 1981), was .65. Adlof et al. (2006) reported that the latent correlation between reading fluency and a general language factor was .71.

Reader ability, gender, and grade level—General reading ability (e.g., skilled vs. struggling), gender, and grade level may also account for individual differences in ORF with skilled readers and older readers benefiting more from context and features of text than less skilled and younger readers (Hiebert, 2005; Myers & Paris, 1978; O'Connor et al., 2002; Torgesen et al., 2001). For example, Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003a) examined the facilitative effect of context-based reading versus context-free reading. They demonstrated that skilled students benefited significantly from context, reading connected text 3 times faster than struggling students. Daane, Campbell, Grigg, Goodman, and Oranje (2005) examined the substantive influence of student characteristics (gender, race, and reading ability) on reading fluency rates among 1,779 students in Grade 4. They reported that females read, on average, faster, more accurately, and with greater prosody than males. Differences among average reading rates were also observed among students categorized by race/ ethnicity, with 45% of White students reading at an average rate of 130 words per minute or more, compared with 18% of African American students, and 24% of Hispanic students. Lastly, the relation between reading rate and reading comprehension was positive.

Features of the Text

Passage difficulty—Variability in ORF performance has been linked to variations in text difficulty (Betts, Pickart, & Heistad, 2009; Hintze, Daly, & Shaprio, 1998). Christ & Silberglitt (2007) estimated the magnitude of standard error of measurement (*SEM*) for ORF scores across passage sets among 8,200 students in Grades 1–5. Results revealed that a major contributor to the observed magnitude of the *SEM* was variability in passage difficulty within passage sets. The *SEM* averaged 10 words read correct per minute across Grades 1–5. Francis et al. (2008) examined the effect of passage difficulty and presentation order on ORF rates among 134 students in Grade 2. Passage effects significantly altered the shape of students' growth trajectories and affected estimates of linear growth rates. Poncy, Skinner, and Axtell (2005) used generalizability theory to assess variability in ORF scores attributable to students, passages, and error among 37 students in Grade 3. Results revealed that 81% of the variance in ORF rates was due to student reading proficiency, 10% due to passage variability, and 9% due to unaccounted sources of measurement error.

Passage length—Passage length has also been reported to influence ORF rates. Biancarosa (2005) compared the predictive utility of ORF calculated for sentence reading and passage reading. Text length significantly affected the magnitude of correlations of fluency and comprehension, with passage reading rates explaining more variance (about 20%) in reading comprehension than sentence reading rates. Daane et al. (2005) examined the substantive influence of reading duration on reading fluency rates calculated for the first 60 s and the full passage among students in Grade 4. They reported that for skilled students, reading fluency estimates obtained for 1 min of reading were comparable to estimates obtained for the full passage. However, among struggling students, reading for shorter periods of time (i.e., 1 min) was associated with higher comprehension performance than reading for longer periods of time (i.e., full passage).

Genre—Several studies have demonstrated that children's reading performance differs by genre, defined as both text structure and format (Stamboltzis & Pumfrey, 2000). Knowledge of text structure may signal students about information relevant to the text's topic or structure that in turn cues germane background knowledge (Perfetti, 1994). Among older students, narrative prose is easier to read (Graesser, Hoffman, & Clark, 1980) and better understood (Best, Floyd, & McNamara, 2008; Diakidoy, Stylianou, Karefillidou, & Papageorgiou, 2005) than expository prose. Cervetti, Bravo, Hiebert, Pearson, and Jaynes (2009) reported that direct manipulations of genre where students read both a fictional narrative and expository text on the same topic did not differentially impact reading accuracy or reading rate. In contrast, Hiebert (2005) reported that students who read expository texts from science and social studies made greater gains than students who read narrative texts from basal readers.

Language and discourse features—Language and discourse features influence reading comprehension and could potentially influence reading fluency (Graesser, McNamara, & Kulikowich, 2011). Language and discourse features influence the activation of information during reading because ideas in text that are consistent with semantic and conceptual long-term memory are activated more quickly during reading (e.g., Collins & Loftus, 1975; Smith, Shoben, & Rips, 1974). Information in the online processing cycle during reading, which includes the current sentence and the highly activated information prior to the sentence being read, may reactivate information in working memory from previous processing cycles (Albrecht & O'Brien, 1993; Myers & O'Brien, 1998). Semantic overlap across processing cycles (e.g., redundancy) leads to faster processing times (Kintsch, 1988). Moreover, readers are more likely to reactivate information from previous processing cycles when the information in the online processing cycle contributes to comprehension and the evolving mental model of the text (O'Brien, Albrecht, Hakala, & Rizzella, 1995; Suh & Trabasso, 1993). Whereas consistent information across cycles speeds processing times, inconsistent information results in slower reading times because it violates a reader's goals or standards of coherence. Under such circumstances, the reader must first identify the inconsistency and then engage strategic, comprehension monitoring processes in an attempt to maintain or repair coherence (Albrecht & O'Brien, 1993). Thus, language and discourse features likely influence the rate at which complex written texts can be reliably processed and a mental model of text can be built.

Recent research has helped consolidate frameworks for assessing language and discourse factors (Graesser et al., 2011). For example, Graesser et al. (2011) reviewed the automated text analyses systems that are currently used to scale texts on multiple characteristics and identified 53 measures that characterize the words, sentences, and connections between sentences associated with deep levels of comprehension. In an attempt to reduce the 53 measures to a small number of functional dimensions, Graesser et al. performed a principal component analysis (PCA) on a corpus of 37,520 texts and validated the PCA by examining the extent to which the *z*-scores of each factor varied as a function of genre (i.e., language arts, social studies, and science) and grade (grades were based on Degrees of Reading Power scores that were transformed into grade levels that aligned with the Common Core literacy standards of the Common Core State Standards Initiative, 2010; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). Results revealed that five functional dimensions of text aligned with the multilevel theoretical frameworks of reading comprehension and explained a significant proportion of the variability among texts.

The five dimensions include (a) *Narrativity:* "captures the extent to which the text conveys a story, a procedure, or a sequence of episodes of actions and events with animate beings"; (b) *Referential cohesion:* "the extent to which explicit content words and ideas in the text are connected with each other as the text unfolds"; (c) *Syntactic simplicity* is "higher when sentences have fewer words and simpler, more familiar syntactic structures"; (d) *Word concreteness* is "higher when a higher percentage of content words are concrete, are meaningful, and evoke mental images—as opposed to being abstract"; and (e) *Deep cohesion* is "higher to the extent that clauses and sentences in the text are linked with causal and intentional (goal-oriented) connectives" (Graesser et al., 2011). Recent research has shown that language and discourse factors can reliably identify differences between written text and spoken discourse (Graesser, Jeon, Yang & Cai, 2007), among various sources, purposes, and authors of written texts (Crossley, McCarthy, Louwerse, & McNamara, 2007; Graesser, Jeon, Cai, & McNamara, 2008), and between texts of high and low cohesion (McNamara, Louwerse, McCarthy, & Graesser, 2010). However, little research has examined the extent to which these language and discourse features interact with characteristics of the reader to influence the rate in which passages are read.

Study Purpose

Previous research has suggested that characteristics of the reader and features of the text significantly influence oral reading fluency performance. However, the extent to which reader characteristics and text features *interact* to affect the *measurement* of reading fluency has not been systematically investigated, especially among middle-grade readers. Therefore, the first purpose of the present study was to evaluate the extent to which reader characteristics (i.e., reader abilities, gender, and grade level) affect ORF performance among middle-grade readers. We predict that sight word reading efficiency will have the largest effect on ORF performance given its high correlation with ORF among middle-grade readers (Adlof et al., 2006).

The second purpose was to examine how text-level features, such as passage length, genre, difficulty level, and language and discourse features, affect ORF performance. We predict that difficulty, as measured by Lexile, will have the largest effect on ORF given that the set of other measures of text ease/difficulty are strongly related to Lexiles and also how long it takes to read a passage (Graesser & McNamara, 2011). Lexile is a function of passage word frequency and syntactic complexity, with longer sentence lengths and words of lower frequency leading to higher Lexile values (Lennon & Burdick, 2004). Existing research has shown that genre (i.e., narrativity) and syntax had the most robust correlation with Lexile scores as well as reading times (Graesser & McNamara, 2011). Although there are many alternative ways of measuring text difficulty, Lexile was selected because no other measure of text difficulty yields an interval scale that can be used to scale both text difficulty and reader ability.

Our final purpose was to investigate the whether characteristics of the reader and features of text interact in the assessment of ORF. We predict that the impact of passage difficulty level on ORF rates will depend on the ability level and grade level of the student, with skilled readers modulating their reading rates more effectively than struggling students and older readers modulating their reading rates more successfully than younger students.

Method

Participants

The participants were students in Grades 6 through 8 (*N* = 1,794) who were first-year participants in a multiyear reading intervention study conducted in two large urban cities in the southwestern United States (Vaughn et al., 2010). The study participants were from seven middle schools and were selected into the study as either adequate or struggling readers (see below). Three of the seven schools were from a large urban district in one city. Four schools were from two medium-sized districts that drew both urban students from a nearby city and rural students from the surrounding areas. The percentage of students qualifying for reduced or free lunch ranged from 56% to 86% in the first site and from 40% to 85% in the second site. Sixty-two students were excluded from the analyses because of incomplete data (final sample, *N* = 1,732). The excluded students were comparably distributed across grade, gender, and reader group (i.e., typical vs. struggling readers).

Although the present study was specifically designed to evaluate the three research questions, because the parent study focused on intervention, struggling readers were overrepresented in the sample (1,028 Struggling, 704 Typical). *Struggling readers* were defined as students who scored at or below a scaled score of 2,150 on their first attempt of the state reading comprehension assessment taken in the spring prior to the study year. The scale score of 2,150 is one-half of one *SEM* above the pass–fail cutpoint (2,100) for the test. It was selected to ensure that all potential struggling readers (students who failed the test and those hovering around the cutpoint who may not meet the threshold if tested again) were included in the intervention study. In addition, students in special education who did not take the state accountability test because of an exemption and extremely poor reading skills were also defined as struggling readers. Adequate readers obtained scale scores above 2,150 on their first attempt in the spring prior to the study year. Students were excluded if (a) they were enrolled in a special education life skills class; (b) their reading performance levels were below a first-grade reading level; (c) they presented a significant sensory disability (e.g., blind, deaf); or (d) were classified as Limited English Proficient by their district. Because more than 80% of students pass the test, we randomly selected adequate readers within school (and grade) in proportion (2:3) to the number of struggling readers.

Measures

ORF was assessed with different passages (both within and between students) to evaluate text effects on student fluency.

The ORF-passage fluency (ORF-PF)—The passages used for the assessment of ORF were designed as progress monitoring assessments in Grades 6 through 8 specifically for the purposes of the present study. The ORF-PF assessment consists of graded passages administered as short 1-min probes to assess fluency of text reading. The passages were derived from other passages to which the authors had access or were written to fill in gaps in the needed levels of difficulty. For this study, there were thirty-five, 108–591 word passages. The outcome measure is words correct per minute (WCPM). Mean fluency correlations among the passages for Grades 6 – 8 were .89–.90.

Passage features—For this study, students read five passages consecutively at a single time point. The passages varied in multiple features, some of which were evaluated in this study as potential sources of within-student variation in fluency. The eight text features evaluated included (a) difficulty as measured by Lexile, (b) text type (narrative vs. expository), (c) page length, (d) narrativity, (e) syntactic simplicity, (f) word concreteness, (g) referential cohesion, and (h) deep cohesion.

1. *Lexile* (Stenner, Burdick, Sanford, & Burdick, 2007) is a measure of test difficulty and a function of word frequency and sentence length. The scale is based on item response theory and the typical text measure range is 0 –1,800 Lexiles. Lexile was a factor in the study design including the randomization of students to passage sets (see the Procedure section). The passages in this study ranged in difficulty from 390 to 1,110 Lexiles.

2. *Text type*. Passages were categorized as narrative or expository prior to the study, although this feature was not a factor in the study design. A passage was categorized as narrative if the purpose was to tell a story, to entertain, or to provide an aesthetic literacy experience for the reader. Narrative passages followed a story grammar or structure composed of (a) beginning, (b) middle, and (c) end (Tonjes, Wolpow, & Zintz, 1999). A passage was categorized as expository if its primary purpose is to convey information, to explain, to describe, or to persuade (Heller, 1995). Expository passages followed a macrostructure such as (a) description, (b) enumeration, (c) sequence/procedure, (d) compare/contrast, (e) problem/solution, or (f) argumentation/persuasion (Culatta, Horn, & Merritt, 1998; Westby, 1994). Twelve of the passages were categorized as narrative and 23 as expository. Passages were reviewed by two Language Arts teachers and rated as either narrative or expository. Disagreements were resolved by one of the primary investigators associated with the research project.

3. *Page length*. After data collection, the passages were categorized according to page length (as presented to the students). Six, 10, and 19 passages were .5, 1, and 1.5 pages in length, respectively.

   The following five passage features were calculated after study implementation using Coh-Metrix (Graesser et al., 2011). These features represent five dimensions resulting from the PCA of 53 Coh-Metric measures.

4. *Narrativity* is the degree to which text is storylike and includes animate beings as opposed to information about topics. It is a function of multiple Coh-Metrix measures including firstperson and third-person pronouns and intentional actions, events, and particles. Narrativity scores for the passages used in this study range from .10 to .98.

5. *Syntactic simplicity*. Texts with higher syntactic simplicity scores include sentences with fewer words and simpler structures (e.g., fewer words before main verb, sentences with similar syntactic structure throughout passage). Syntactic simplicity scores for the passages used in this study range from .54 to .99.

6. *Word concreteness*. Higher scores for word concreteness are associated with texts that contain “a higher percentage of words that are concrete, meaningful, and evoke mental images” (Graesser et al., 2011, p. 230). Word concreteness scores for the passages used in this study range from .24 to .97.

7. *Referential cohesion* is the extent to which words overlap across sentences in the text. Referential cohesion scores for the passages used in this study range from .07 to .68.

8. *Deep cohesion* represents the extent to which clauses and sentences in text are linked with goal-oriented connectives. Deep cohesion scores for the passages used in this study range from .10 to .98.

Verbal and reading achievement was assessed with the same measures for all students to evaluate the effects of student characteristics on ORF.

Kaufman Brief Intelligence Test-2 Verbal Knowledge subtest (KBIT; Kaufman & Kaufman, 2004)—The KBIT Verbal Knowledge subtest assesses receptive vocabulary and general information. Students are required to choose one of six illustrations that best corresponds to an examiner question. Internal consistency values range from .89 to .94, and test–retest reliabilities range from .88 to .93 in the age range of the students in this study.

Woodcock-Johnson III Letter-Word Identification (LWID; Woodcock, McGrew, & Mather, 2001)—LWID assesses the ability to read real words. Split-half and test–retest reliabilities range from .88 to .95 in the age range of the students in this study.

Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999)—The TOWRE consists of two subtests: Sight Word Efficiency (real words) and Phonemic Decoding (nonwords). For both subtests, the students were given a list of 104 words and asked to read them as accurately and as quickly as possible. The raw score is the number of words read correctly within 45 s. Alternate forms reliability and test–retest reliability coefficients are above .89 for students in Grades 6–8 (Torgesen et al., 1999). Standard scores were used.

Procedure

Students were assessed at the beginning of the school year, prior to intervention, with ORF-PF, KBIT, LWID, and TOWRE. For ORF-PF, students read five passages consecutively for 1 min each, one from each of five Lexile bands having an overall range within student of 550 Lexiles. The order in which the students read the passages was easiest to most difficult as measured in Lexiles. For the sake of assigning students to passages, passages were first organized into sets of five passages. Subsequently, within grade and reader group (typical vs. struggling), students were randomly assigned to one of five sets of five passages. Table 1 shows the number of students by grade and reader group assigned to each passage and passage set.

All participants were assessed by examiners who had completed an extensive training program conducted by the investigators focused on test administration, scoring, and verification procedures for each measure. Prior to testing study participants, each examiner demonstrated at least 95% accuracy in test administration during practice assessments. All assessments were completed at the students' school in quiet locations (e.g., library, unused classrooms).

Analytical Approach

The unique multivariate, multilevel design of this study permitted evaluation of (a) how reader characteristics affect ORF (i.e., how differences in fluency *between* students are related to reader characteristics); (b) how text characteristics affect ORF *within* students (i.e., how a student's fluency changes when the same student reads different texts); and (c) the extent to which these two sources of variability *interact* in their effects on ORF (i.e., whether the effects of text characteristics differ systematically across students on the basis of reader characteristics). To investigate these questions, we used multilevel models that differed from one another in their fixed and random effects, as well as in their specifications regarding the unexplained variance in ORF scores. We evaluated four sets of models: random-effects models, student effects models, text effects models, and full models, including student and text effects. The randomeffects models without student or text effects were designed simply to estimate the magnitude and structure of the variance and covariance in ORF scores available to be explained by the other three classes of models. All effects were estimated in SAS using PROC MIXED (SAS Institute, 2009).

We first evaluated 20 random effects models to estimate the magnitude of the variance and covariance in ORF scores and to understand the structure of these variance components. For each of these 20 models, a single fixed effect was estimated for the intercept (i.e., the model included only a grand mean estimated across all students and passages). The 20 models can be classified into four sets that differ from one another with respect to the structure of the variance components.

The four sets of models differed on the basis of whether intercept variances and/or residual variances and covariances were allowed to differ across grades. We distinguished between these models using a simple two-part pneumonic, II-RR, where the letters to the left of the hyphen specify whether intercept variances were forced equal (NG) or allowed to differ (G) across grades, and the letters to the right of the hyphen, which specify whether residual variances were forced equal (NG) or allowed to differ (G) across grades. Thus, the four classes of models are designated NG-NG, G-NG, NG-G, G-G. For example, NG-G means that the model specified no grade difference in intercept variances, but allowed for grade differences in residual variances and covariances.

Within each of these four sets of models, we evaluated five different residual covariance structures that differ in their implications for residual variances and covariances and can be ordered in terms of the number of estimated parameters as follows: variance components (1), compound symmetry (2), autoregressive (2), banded main diagonal (5), and unstructured (15). The variance components models estimated a single residual variance for all five passages with covariance between residuals set to zero. The compound symmetry models estimated a single residual variance for all five passages and a single covariance between passages. This model implies that passages are equally correlated with one another. The autoregressive model included a single residual variance estimate as well as a multiplicative factor (ρ). In this model, the residual covariance is a function of the variance and ρ, such that any covariance is equal to the variance times ρ^j, where *j* describes the distance between passage reads (e.g., adjacent reads such as Reads 1 and 2 give $j = 1$, Reads 1 and 3 give $j = 2$. Given this structure, the correlation between residuals is simply a function of the distance between measures and is equal to ρ, ρ^2, ρ^3, and so on, for measures that are one, two, three, etc., steps apart, respectively. The banded main diagonal models included separate residual variance estimates for each of the five passages with covariances between residuals set to zero.

Finally, unstructured models include different estimates for all residual variances and covariance across the five passages. Although the models differ in terms of the number of parameters required to describe the variances and covariances, they are not nested models. Consequently, they were evaluated using Akaike's (AIC) and the Bayesian (BIC) information criteria. The standardized root-meansquare residual (SRMR) was also used because it is an index of how well models reproduce sample variances and covariances and because we were interested in estimating the relative amounts of between- and within-student variance. To determine SRMR, the random effects models were estimated in Mplus (Muthén & Muthén, 2007) in addition to SAS. The criterion we used to evaluate model fit was SRMR ≤ .08 (Hu & Bentler, 1999). All subsequent analyses were based on the best random-effects model.

Student effects models were evaluated to determine the degree to which student characteristics predict between student performances without accounting for text effects. Text effects models were evaluated to determine the degree to which text-level parameters affect within-student performance. Finally, a full model that included student- and text-level effects as well as their interactions was evaluated.

Results

Passage Set Group Comparisons

Table 2 shows demographic and achievement information for each of the five groups of students randomized to a particular set of passages in each grade. Chi-square analyses and analyses of variance were conducted to determine whether groups differed on demographic and achievement measures. As expected, due to randomized assignment, the five groups within each grade did not differ in demographic composition or performance on achievement measures, with two exceptions likely due to the highly powered study design because the differences are small. Among Grade 8 students, there was an overall effect of group on age, $F(4, 645) = 2.87$, $p = .02$, and on LWID, $F(4, 645) = 2.48$, $p = .04$. On the basis of Tukey's post hoc analyses, Grade 8 students assigned to Passage Set A were significantly older than students assigned to Passage Set B (13.4 vs. 13.2 years, $p < .05$). Grade 8 students assigned to Passage Set D scored significantly higher on LWID than students assigned to Passage Set A (99.9 vs. 94.8, $p < .05$). We did not adjust for these differences because the differences are neither large nor pervasive, but simply note them here as possible challenges to model interpretation in Grade 8.

Passage Fluency Performance

Some inferences can be made from examination of mean WCPM (see Table 3) and correlations among passage WCPM within grade and passage set (see the supplemental material, Appendix A). For example, controlling for passage, mean WCPM consistently increased with grade level, with mean increases as much as 20 WCPM. This pattern suggests variability between students that can be explained by at least one reader characteristic, grade level. There are similar ranges in mean performance across passages within grade and passage sets, suggesting that there may be some variability in ORF within students. Correlations among passage WCPM are consistently high (.83–.93), suggesting that students' ORF relative to other students does not change substantially across passages. However, it is impossible to discern the mean and covariance structure of reading fluency from the descriptive data presented in Table 3 and Appendix A in the supplemental material.

Random-Effects Model Analyses

Appendix B (see the supplemental material) shows fit statistics for 12 of the 20 random-effects models evaluated to determine the model that best represented the covariance structure for ORF across middle-school grades when students were asked to read multiple passages that differed in text characteristics. The eight compound symmetry, corresponding to a uniform correlation structure, and unstructured models across all four sets of models were not included because of problems with estimation in both SAS and Mplus (e.g., nonpositive definite matrices, inability to estimate standard errors).

Within the four sets, autoregressive models fit better than variance component and banded main diagonal models (which has its nonzero elements within a "band" about the diagonal), when evaluated with AIC and BIC. The autoregressive model for residuals specified that the residual variance was equal across the five passages read by a student and that the correlation between passages was a function of the distance between the passages as indexed by the order in which the passages were read. More precisely, the correlation was largest for adjacent passages (e.g., first passage and second passage, second passage and third passage, etc.) and became progressively smaller as the distance between passages increased (e.g., first passage with third passage, second passage with fourth passage, etc.). All correlations between passages at a given distance were equal, and the correlation decreased proportionately as the distance increased. Specifically, if the correlation between adjacent passages was ρ, then the correlation between passages at a distance of two (e.g., first read passage and third read passage) is ρ^2, and the correlation at a distance of three (e.g., first read passage and fourth read passage) was ρ^3. This structure implies that the correlation between passages was not simply a function of different passages measuring a single common construct, in so far as a simple common factor model with uncorrelated residuals cannot produce this structure. This structure implied a common factor model with equal validity coefficients and a stable, autoregressive error structure. Such a structure could arise from factors like motivation, inattention, and/or fatigue that carried over from one passage to another. The greater the distance between passages, the weaker the carryover effect and the weaker the effect on the correlation between passages.

On the basis of AIC and BIC, autoregressive models that allowed either residual covariances or intercept variances and residual covariances to differ across grades were the best fitting models. However, these models did a poor job of reproducing sample variances and covariances as indicated by SRMR (NG-G, SRMR = .19; G-G, SRMR = .15). These models produced disproportionate estimates of between- and within-student variances, primarily for Grade 8 (64% and 48% model estimated between student variances vs. approximately 80% sample between student variance, see Appendix B in the supplemental material). Therefore, we based all subsequent analyses on the NG-NG autoregressive random-effects model in which intercept variances and residual covariances were forced to be equal across grades. This model reasonably reproduced sample means and covariances (SRMR = .06).

Appendix B (see the supplemental material) shows the model estimated intercept and residual variances and covariances from the autoregressive, NG-NG model. Residual variances were equal across the five reads (376.97), and covariances were equal among pairs of reads that were the same distance apart (e.g., 174.39 between adjacent reads, e.g., Reads 1 and 2; 80.68 between reads with a distance of 2, e.g., Reads 1 and 3). Covariances declined proportionally the farther apart the reads.

On the basis of the autoregressive, NG-NG, random-effects model both intercept and residual variances were significant (intercept variance = 1178.60, $z = 25.37$, $p < .001$; residual variance = 376.97, $z = 26.98$, $p < .001$), with between-student variance accounting for 76% of the variance in ORF. The model estimated mean ORF across middle-school students and passages within students at 125.47 WCPM.

Student-Level Analyses

We first evaluated the effects of reader characteristics (grade, gender, reader type – struggling vs. typical, verbal knowledge, decoding skill, and sight word reading fluency) on ORF. Evaluated individually, each reader characteristic had a significant effect on ORF (see Table 4). Not surprisingly because it assesses fluency for reading word lists (i.e., sight word efficiency), TOWRE word reading fluency performance explained the most variance in average (across the five reads) student performance (70% of variance explained; intercept variance from unconditional model = 1178.60, see above; intercept variance after accounting for TOWRE performance = 248.83, see Table 6). TOWRE performance predicted an increase of 27.5 WCPM in ORF for a 15 standard score point increase (i.e., approximately 1 *SD*) in word reading fluency. The LWID measure had an effect on ORF similar to TOWRE (i.e., 26.7 WCPM increase for 1 *SD* increase in decoding performance), although it explained less variance in student performance (52%). The remaining characteristics explained substantially less variance in student performance: reader type (24%), verbal knowledge (21%), grade (4%), and gender (2%). Typical readers read 34 WCPM faster than struggling readers; females read 11 WCPM faster than males; and Grade 7 and 8 students read eight and 16 WCPM faster than Grade 6 students. When all reader characteristics are combined in one model, each still had significant effects on ORF when controlling for the others; however, there were no significant interactions (see Table 6 for models, including only main effects and including interactions between student characteristics). Combined, reader characteristics explained approximately 80% of the between-student variance in mean ORF (i.e., intercept variance after accounting for all reader characteristics was 234.21, see Table 6).

Text-Level Analyses

We next evaluated passage effects to determine the proportion of within-student variance due to passage. Passage accounted for approximately 55% of the within-student variance in ORF (residual variance from unconditional model = 376.97, see above; residual variance after accounting for passage = 168.40). The model estimated range in mean ORF is approximately 72 WCPM (model estimated low M = 86 WCPM, high M = 158 WCPM depending on the passage being read).

Before evaluating the effect of text features on ORF, we examined the relations among the text features for the 35 passages. Text difficulty as measured by Lexile is largely a function sentence length and word frequency, and not surprisingly is most highly correlated with syntactic complexity ($r = -.82$, $p > .05$) and narrativity ($r = -.48$, $p < .05$). The correlation of Lexile and narrativity and syntax replicate what was reported in the Graesser and McNamara (2011) and a study by Nelson, Perfetti, Liben, and Liben (2012), who examined the correlation of Coh-Metrix dimensions and student performance-based difficulty measures. Together, these measures explain 75% of the variance in Lexile scores across the 35 passages. Correlations between Lexile and the other three Coh-Metrix measures (word concreteness, referential cohesion, and deep cohesion) were not statistically significant (r = .18, −.08, and .02, $p > .05$). The correlations among the five Coh-Metrix measures range from −.32 to .26, and none are statistically significant ($p > .05$). The low correlations among these measures are expected because they were derived from PCA (Graesser et al., 2011) and therefore represent statistically orthogonal factors.

We also evaluated the concordance between the categorization of the passages as narrative or expository (text type) by expert teachers and the objective (and continuous) measure of narrativity by computerized text analysis. The two measures are relatively highly correlated ($p = .60$, $p < .05$). The narrativity scores range from .10 to .98. All passages categorized as “narrative” have scores greater than or equal to .64. The majority of passages categorized as “expository” have scores less than .64. However, eight of the 23 “expository” passages have relatively high narrativity scores. For example, the passage “Suni” has a narrativity score of .86. The topic is the endangered Chinese White Dolphin, but it is written in first person from the perspective of “Suni,” a White Dolphin. It is possible that although highly similar, these two variables have unique effects on ORF. Students may read passages with relatively high fluency if they are highly narrative in style, but given two highly narrative passages, perhaps they read an informational passage more slowly than one simply telling a story. Similarly, though highly correlated, Lexile and syntactic simplicity may have unique effects on ORF. Therefore, we evaluated all eight text features as possible determinants of mean differences in ORF across passages.

First, we evaluated the effect of each text feature on ORF individually (see Table 5). Ignoring all other text features, each individual text feature except word concreteness significantly affected ORF (see Table 7). Reading fluency decreased by 8.6 WCPM per standard deviation (~212) increase in text difficulty measured on the Lexile scale and increased by 6.4 WCPM per standard deviation increase in narrativity. These effects varied across students as reflected in the statistically significant random-effects estimates (see

Table 7) in the single predictor models based on these text features. Reading fluency also increased by 1.9 and 1.6 WCPM per standard deviation increase in syntactic simplicity and referential cohesion, but decreased by 0.5 WCPM per standard deviation increase in deep cohesion. Student-level ORF was approximately 7.6 WCPM higher for narrative text than expository text and 2.5 WCPM higher for a half page of text than for more than one page of text. Other than test difficulty measured on the Lexile scale and narrativity, none of the other effects of text features on ORF varied across students (i.e., nonsignificant random effects, see Table 7) in the single-predictor models. In the single-predictor models, Lexile level accounted for 52% of the variance in ORF across texts (explaining almost all of the variance due to passage). In contrast, narrativity explained approximately 34% of the variance due to passage, whereas text type explained 16% of the passage variance. It is important to keep in mind that these estimates are not independent and that some of the variance accounted for by the different text features is shared across features due to their intercorrelations.

Because of the strong relations between Lexile and syntactic simplicity, we evaluated them together in a single model to determine whether each has a unique effect on ORF when controlling for the other. As shown in the lower half of Table 7, syntactic simplicity appears to act as a suppressor when combined with Lexile in the same model. Although ORF is higher for syntactically simpler text (i.e., positive effect when syntactic simplicity is evaluated as a single predictor), the effect becomes negative when combined with Lexile (see Table 7) level in the same model. Lexile level appears to have greater explanatory power for changes in ORF than syntactic simplicity (i.e., greater practical effect on ORF −8.6 vs. 1.9 per *SD* change and significant random effects for Lexile but not syntactic simplicity). In addition, the effect of syntactic simplicity is misleading when included in the same model as Lexile (i.e., negative and opposite its effect when evaluated as a single predictor). Therefore, for all subsequent models, we excluded syntactic simplicity.

For similar reasons as above, we evaluated narrativity and text type in a single model. As Table 7 shows, when evaluated together, narrativity has a significant effect on ORF (6.1 difference in WCPM per *SD* change in narrativity, $p < .05$), but text type does not (0.85 decrease in WCPM from narrative to expository text, $p > .05$). Although there is a theoretical distinction between these constructs, it does not appear to be statistically detectable as operationalized in this study. Because narrativity has more explanatory power than the simple dichotomy of narrative versus expository text type, for all subsequent models, we excluded text type.

The final text effects model is shown in Table 7. This model includes Lexile level, narrativity, referential cohesion, deep cohesion, and page length. Syntactic simplicity, text type, and word concreteness were not included in the final model for the reasons stated above. Each of the five text features included in the model had a significant effect on ORF when controlling for the others (see Table 7). In addition, both the Lexile and narrativity effects varied across students when included in the same model (i.e., significant random effects, $p < .05$, see Table 7). Combined, the three text features account for 52% of the variance in student-level performance. These findings are similar to that of Graesser and McNamara (2011), who reported that 54% of the variance in self-paced word reading times was accounted for by text complexity.

In the final text effects model, Lexile level and narrativity have the strongest and comparable effects on ORF (five WCPM per *SD* change in Lexile level or narrativity, see Table 7), followed by page length (four WCPM for .5 vs. 1.5 pages of text), and referential cohesion (1.5 WCPM per *SD* change). Although statistically significant, deep cohesion has a negligible practical effect on ORF (0.6 WCPM per *SD* change). Due to the complexity of combining student and text effects and potential interactions, and the minor effect of deep cohesion on ORF, we excluded this text feature from subsequent models.

Student- and Text-Level Effects on ORF

For the overall model of student- and text-level effects on ORF, we evaluated main effects for reader characteristics: grade, gender, reader type (struggling vs. typical), verbal ability (KBIT Verbal), letter word identification, and word reading efficiency (TOWRE) and text-level features: Lexile, narrativity, referential cohesion, and page length. In addition, because Lexile and narrativity effects varied across students, we evaluated specific two-way student level by text-level interactions to determine whether grade, gender, and/or reader type moderate the effects of Lexile and narrativity on ORF. Table 6 shows all estimated fixed and random effects for the final model.

As the information in Table 8 demonstrates, each text feature and reader characteristic had a significant effect on ORF when controlling the others. However, grade, gender, and reader type moderated the effects of Lexile and narrativity on ORF (i.e., all interactions were statistically significant, see Table 8). In general, and not surprisingly, students in higher grades read faster than students in lower grades; females read faster than males; typical students read faster than struggling students; and students with higher verbal abilities, word reading skills, and word reading fluency read faster than those students with lower abilities and skills. In addition, students read slower when reading was more difficult, less narrative, less cohesive, and longer texts (also not surprising), all other things being equal. However, as text becomes more difficult, students in higher grades slowed down more than students in lower grades (Lexile × Grade interaction), females slowed down more than males (Lexile × Gender interaction), and typical readers slowed down more than struggling readers (Lexile × Reader Type interaction). The moderating effect of gender on narrativity is similar to its effect on Lexile (i.e., females slowed down more than males the less narrative the text, Narrativity × Gender interaction), but students in different grades and at different reading levels (typical vs. struggling) react differently to narrativity than they do to overall text difficulty (Lexile). Grade 6 students slowed down more than Grade 8 students, but Grade 7 students slowed down less than Grade 8 students when text was less narrative. In addition, struggling readers slowed down less than typical readers when text was less narrative.

Figure 1 illustrates the practical effects of student characteristics and text features on ORF. Figure 1 (a) demonstrates the additive effects of overall text difficulty (Lexile) and narrativity on withinstudent ORF. A typical Grade 8 male reads a difficult text (942 Lexiles, + 1 *SD*) that is low in narrativity (.35, − 1 *SD*) at a rate of 126 WCPM. As the text increases in narrativity (to .87, + 1 *SD*), the student's rate goes up to 138 WCPM. If the text also becomes less difficult overall (to 518 Lexiles, −1 *SD*), the student's rate increases to 154 WCPM (an overall change of 28 WCPM for 2 *SD* change in both Lexile and narrativity).

Figure 1 (b) demonstrates that a Grade 6 male student increases ORF by six WCPM (111 to 117), a Grade 7 male student by 10 WCPM (118 to 128), and a Grade 8 male student by 16 WCPM (132 to148) as text difficulty decreases by two standard deviations (942–518 Lexiles). However, these within-student changes in fluency due to text difficulty are not as great as between-student differences when two students differ in word reading fluency by just one standard deviation (15 standard score [SS] different in TOWRE scores). For example, although a Grade 8 male may increase by 16 WCPM for a two standard deviations drop in Lexile, his counterpart whose word reading fluency is one standard deviation higher reads 22 WCPM faster regardless of Lexile (e.g., 110 vs. 132 WCPM for 85 vs. 100 SS on TOWRE).

Finally, Figure 1 (c) demonstrates the additive effects of gender and reader type on ORF. A typical Grade 8 female reads five (132 vs. 137), six (131 vs. 137), and 10 (127 vs. 137) WCPM faster than a typical male, struggling female, and struggling male if the text is relatively difficult (942 Lexiles), but will read seven (148 vs. 155), eight (147 vs. 155), and 15 (140 vs. 155) WCPM faster if the text is relatively easy (518 Lexiles). These between-student effects of gender and reader level are not as large as within-student effects of Lexile. For example, a Grade 8 typical female increases 17 WCPM (137 to 155) when the text difficulty decreases from 942 to 518 Lexiles (2 *SD*) and a Grade 8 struggling male increases 13 WCPM (127 to 140) for a similar decrease in text difficulty. However, it must be kept in mind that these differences between typical and struggling readers are somewhat understated as they are estimated, holding constant the reading rate as measured on the TOWRE; that is, the two students are presumed to be reading at comparable rates as measured on the TOWRE.

Discussion

Characteristics of the Reader

The first purpose of this study was to evaluate whether reader characteristics (i.e., reader abilities, gender, grade level, and reader type) affected ORF performance among middle-grade readers. We predicted that sight word reading efficiency would have the largest effect on ORF performance given its high correlation with ORF among middle-grade readers (Adlof et al., 2006). Collectively, these student attributes accounted for 80% of the variance in ORF among middle-grade readers. On average, students performing higher on reading-related skills (i.e., sight word reading efficiency, phonological decoding, and verbal knowledge) read texts faster than students who performed lower on these skills. Of the three reading-related skills, sight word reading efficiency had the strongest influence on fluency rates, which is not surprising given that sight word reading efficiency and passage reading fluency are highly correlated among middle-grade readers (Barth et al., 2009). As expected, typically developing students read texts faster than struggling readers, and older students (Grades 7 and 8) read faster than younger students. Interestingly, females read connected text faster than males. However, the effects of grade and gender on ORF rates were small compared with the strong influence of sight word reading and verbal knowledge.

Features of Text

The second purpose was to examine how text-level features, including passage length, genre, difficulty level, and language and discourse features (i.e., word concreteness, referential cohesion, deep cohesion, narrativity, and syntactic simplicity) affected ORF performance. We were interested in understanding how ORF rates would change as a function of reading passages that varied in text-level features. We predicted that difficulty, as measured by Lexile, would have the largest effect on ORF given that other measures of text ease/difficulty are strongly related to how long it takes to read a passage (Graesser et al., 2011). Results indicated that these general text-level features collectively accounted for approximately 55% of the variance in ORF performance, with each text-level feature except word concreteness independently influencing ORF rates after controlling for the other features. When considered independently, Lexile and narrativity accounted for greatest amount of variance in ORF abilities.

Results also revealed that the following pairs of text-level features were largely redundant: narrativity and text type as well as Lexile and syntactic simplicity. After removing redundant textlevel features (i.e., text type and syntactic simplicity) that did not vary across students, results revealed that Lexile, narrativity, referential cohesion, deep cohesion, and page length independently influenced ORF rates among middle-grade readers. Lexile and narrativity had the strongest effect on ORF and accounted for 52% of the variance in reading rates. These findings align with previous work by Graesser and McNamara (2011) in which 54% of the variance in word reading times was accounted by four groups of variables that included material layout, surface code, textbase/ situation model, and genre. Reading fluency rates tended to increase (i.e., slow) as Lexile level increased and conversely decreased (i.e., quickened) as the structure of passages became more narrativelike. The specific ordering of these effects and their currently estimated magnitudes must be interpreted somewhat cautiously for text characteristics other than text difficulty as measured by Lexile level. This cautionary note stems from the fact that within each Grade 6–8, students read a set of five passages that spanned from easy to difficult as measured by the Lexile level of the text. Other text features were not independently controlled or manipulated. Consequently, they may vary to a lesser degree across the set of passages than what might be expected. Restriction of range is known to reduce the magnitude of regression coefficients. Although the range of these text features was not explicitly restricted, the possibility exists that the full range of any particular feature was not present in the set of texts because we did not explicitly sample so as to ensure representation.

One possible explanation for why ORF rates decreased at a greater rate for narrative texts, as difficulty increased, is the higher frequency of knowledge-based inferences constructed when reading complex narrative texts. Inferences include the goals and plans that motivate the actions of characters, the knowledge and beliefs held by the character, traits, emotions, motivations of the character that cause events, spatial relationships among entities, predictions about episodes that may occur in the future, and so on (Graesser, Singer, & Trabasso, 1994). These types of inferences are generated by the reader in an attempt to construct a meaningful, referential situation model that not only addresses their goals as a reader but also explains why actions, events, and states are revealed in the narrative text (Graesser et al., 1994). Additionally, these types of inferences are generated quickly and reliably by students when reading narrative texts (Graesser & McNamara, 2011; Haberlandt & Graesser, 1985).

Interaction of Characteristics of the Reader and Features of Text

The final purpose of the study was to investigate which characteristics of the reader and features of text interact in the assessment of ORF. We predicted that the impact of passage difficulty level (as measured by Lexile) on ORF rates would depend on the ability level and grade level of the student, with skilled readers modulating their reading rates more effectively than struggling readers, and older readers modulating their reading rates more successfully than younger readers. This hypothesis was supported. Specifically, as the difficulty level of text increased, skilled readers and older readers (Grades 7 and 8) tended to slow their reading rates to a greater degree than less skilled readers and younger readers. Interestingly, among younger readers (Grade 6), a significant slowing of reading rates was observed among expository texts (e.g., less narrative in story grammar structure) but not narrative texts.

Our results also suggested that each reader and text characteristic independently influenced ORF when controlling for all other characteristics. Several reader characteristics generally lead to faster ORF. Texts were read faster when the reader was older, female, proficient (e.g., typically developing), and possessed adequate basic reading skills (i.e., verbal knowledge, word reading accuracy, and sight word reading abilities). In contrast, several features of text generally lead to slower ORF rates. For example, texts were read at slower ORF when texts were longer, less narrative in structure, less cohesive, and more difficult as measured by Lexile. These results suggest that when selecting texts appropriate for individual readers, the process should include an analysis of gender and whether the topics are age appropriate.

The fact that skilled readers and older readers modulated their reading rates to a greater degree particularly for those longer, more expository in structure, more difficult, and less cohesive texts, is not surprising. Successful comprehension of expository passages requires the use of higher level cognitive skills such as problem solving, planning, and organization to a greater degree than narrative passages (Eason, Goldberg, Young, Geist, & Cutting, 2012). More often than not, readers have specific goals when learning new information (McCrudden, Magliano, & Schraw, 2012). To achieve these goals, readers often engage strategic processes to combine information from explicit statements of the text and relevant information from their general knowledge (McNamara, 2007; van den Broek, 2010). Deployment of strategic processes helps to achieve specific goals but may slow the speed of processing text. In this instance, the slower reading rates observed among skilled readers and older readers when reading expository texts do not reflect comprehension misfires or breakdown. Rather, slower reading rates reflect a greater focus on the maintenance of cohesion when problem solving and learning new information from text and a greater engagement of higher level cognitive skills to achieve specific reader goals.

Connections to Theoretical Models

Skilled readers appear to make adjustments in their reading rates in response to various features of text, their level of understanding, and goals for understanding. In the landscape model, these adjustments reflect the engagement of constructionist processes to help repair or maintain cohesion when lexical inputs are not sufficiently reliable or coherent to establish the mental model or when the goals of the reader have not been met (van den Broek et al., 2005). This hypothesis is supported by research demonstrating that skilled readers more frequently and more strongly activate relevant information from past processing cycles to aid comprehension of the text (O’Reilly & McNamara, 2007; Suh & Trabasso, 1993; van den Broek, 2010; van den Broek, Rohleder, & Naravaex, 1996) and key background knowledge (Lucas, Tanenhaus, & Carlson, 1990; van den Broek, Rohleder, & Naravaex, 1996). Thus, in the landscape model, fluctuations in ORF rates may be observed because skilled readers tap multiple sources of information (i.e., text-based and prior knowledge), constantly update their evolving mental model of text, and deploy different strategies to achieve specific goals when reading.

In addition to modulating reading rates, skilled readers generally read connected text faster than struggling readers. The faster reading rates among skilled readers suggest that the phonological, orthographic, and semantic information of words is tightly coupled, precisely organized, and easily accessed when reading (Ehri, 2002; Ehri & Snowling, 2004). Consistent with the lexical quality hypothesis (Perfetti, 2007), high-quality lexical representations of words drive rapid processing of text and more importantly are responsible for the automaticity of word recognition, which allows cognitive resources to be allocated to the integration of contextspecific word meanings with the evolving representation of the text (Perfetti & Hart, 2001). Skilled readers more efficiently and precisely execute this process, which increases their ability to integrate word meanings within sentences and across sentences (Perfetti, 2007).

In contrast, struggling readers did not appear to modulate their reading rates across various features of text to the same degree as skilled readers. This may not necessarily reflect a strategic choice made by the student but may instead be a concomitant of decoding skills that have not reached automatized levels (Perfetti, 1985; Stanovich, 1987). The lexical quality hypothesis (Perfetti, 2007) specifically suggests that less skilled readers lack detailed knowledge of word form and meanings. In addition, less skilled readers present a reduced ability to learn the meaning of new words while reading, retrieve meaning of learned words, and integrate the meaning of a given word with the evolving mental model of the text. The consequence of generally poor lexical quality is seen both in lexical processing and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

comprehension processing, with less skilled readers unable to effectively and efficiently retrieve the meaning of words required for the specific context.

An alternate explanation for the lack of modulated fluency rates among struggling readers is that their standards of coherence or criteria for defining what is “good comprehension” may be more relaxed relative to skilled readers. According to the landscape model, because struggling readers’ beliefs about what constitutes “good comprehension” are less stringent, they are less likely to initiate strategic processes that repair or maintain cohesion while reading. Consequently, the mental model generated by struggling readers lacks specificity. In sum, because struggling readers may present significant lexical deficits and/or possess relaxed standards of coherence, fluctuations in reading rates are less likely to be observed as text features change and, as demonstrated in this study, have significant impact among middle-grade struggling readers and predictable consequences for subsequent reading development (Adams, 1990; Pearson & Hiebert, 2010; Rayner, Foorman, Perfetti, Pesetsky, & Siedenburg, 2001).

Connections to Past Research on ORF

Little research has examined student and text influences on secondary readers. Results suggested that the influence of text effects is greater among middle-school students than younger elementary students. This finding parallels findings by Kim, Petscher, Schatschneider, and Foorman (2010), who evaluated the variability in observed ORF rates across passages among students in first through third grades. They reported that among first-grade students, 2%–4% of the variance in fluency rates was attributable to text effects, with 81% of the variability estimated at the student level. By third grade, 3%–9% was attributed to text effects, and 85%–91% was attributed to student characteristics. Similarly, Christ and Ardoin (2009) reported that the proportion of variance attributed to text features ranged from 1% to 10% among students in second and third grades, with the proportion of variance attributed to student characteristics ranging from 82% to 93%. Graesser and McNamara (2011) reported that 54% of the variance in word reading times was accounted by four groups of variables that included material layout, surface code, textbase/situation model, and genre. Furthermore, Christ and Ardoin (2009) demonstrated that the amount of variance attributable to text features can be minimized within passage sets when greater control is used to select passages that are equivalent. However, this reduction in the effects of text characteristics is illusory, resulting in large measure from restriction of range on text features.

Our study was designed to examine the effects of student and text characteristics and their interactions on ORF rates, rather than to minimize the influence of texts by selecting them to be maximally comparable. These prior studies suggest that across grades, variability in ORF rates can be attributed to both text features and reader characteristics, with effects of text features possibly increasing with older students, and clearly depending on the reading level of the student, as seen in the present study. Whether the effects of text features increases with older students requires additional research where the range of text features experienced by older and younger readers are comparable. At present, inferences about the influence of text features at different age ranges come from examining the findings of different studies in which features have not been comparably controlled across studies.

Limitations of the Study

A limitation of this study is that the effects of passage features, reader characteristics, and the interaction of these two sources of variability were evaluated at a single time point (beginning of the school year). Because ORF measures are frequently used to measure reading progress across the school year, an important next step would be to evaluate the relative impact of passage and student attributes on growth in reading fluency across time. We set out to accomplish this task but found that the amount of time necessary for struggling readers to read multiple passages, especially of such length, was overwhelming, and focused instead on comparisons at a single time point. A follow-up study that uses fewer passages, in more restricted grade and Lexile levels and shorter than 1.5 pages, would permit this longitudinal assessment of ORF abilities. It would also be useful to compare elementary and secondary students of different ability levels on the same measures to more fully evaluate developmental effects.

A second limitation relates to the effect of syntactic simplicity on ORF. Because of the strong relation between Lexile and syntactic simplicity plus the suppressor effect of syntactic simplicity when combined with Lexile in the same model, syntactic simplicity was excluded from final models that examined the interaction of features of the text and characteristics of the reader. However, this effect should not be so quickly dismissed. It suggests that when Lexile level is held constant, texts that are syntactically simpler tend to be read more slowly. This effect could be observed among types of texts that are syntactically simpler once Lexile is controlled such as poems, stories with a significant amount of dialogue among characters, and expository texts with a high frequency of short, descriptive sentences. Future research might include structural equation modeling to determine whether separate Lexile scales may be necessary for expository and narrative text.

Also noteworthy was the finding of the narrative effect over and above and separable from Lexile text difficulty. However, part of the text difficulty is the commonality in variance with narrativity. Because this study was designed to maximize variability in Lexile values across passages with other text features allowed to vary, it is possible that in a larger corpus of passages where there is greater variance in language and discourse features, their effects could be quite different. Future research should more thoroughly examine the interactions of features of texts and characteristics of the reader with passages that maximize variability in language and discourse features as well as Lexile values. This would allow for a more systematic separation of the influence of language and discourse features on reading times and their interaction with reader characteristics.

Implications for Practice

Currently, measures of ORF are frequently used to monitor reading progress and predict performance on future reading outcomes, which are often used to make decisions about intervention placement and retention (O’Connor, Swanson, & Geraghty, 2010). One of the challenges in measuring ORF ability across time is the selection of passages that are equivalent or parallel within the same grade level (Betts et al., 2009; Christ & Ardoin, 2009; Francis et al., 2008). Historically, ORF passages were selected directly from classroom curriculum materials to ensure high content overlap between assessment and instruction (Deno, 1985; Good & Jefferson, 1998). However, random sampling of curriculum materials or selection of materials based on readability formulae does not adequately control for features of text (i.e., difficulty, length, and genre) previously found to impact the measurement of fluency ability across time (Betts et al., 2009; Christ & Ardoin, 2009; Francis et al., 2008). Such inconsistencies in passage features across alternate passages and assessment time points make it difficult to know whether a student’s ORF is improving, declining, or holding steady when solely examining changes in observed reading fluency scores over time (Francis et al., 2008). Consequently, when measuring ORF, the effects of passage features that extend beyond noninterval assessments of difficulty level must be statistically addressed using procedures that equate alternate forms within grade (see Betts et al., 2009; Francis et al., 2008). Furthermore, results of this study suggest that equating procedures must move beyond simply controlling for text difficulty, genre, and administration order and should include language and discourse features.

Finally, text characteristics that affect the assessment of ORF have been typically underestimated in efforts to use these assessments to understand individual differences in student progress, need for intervention, and related issues. Although the assessment of ORF is sensitive to these individual differences, the present study shows that reading is an interaction of reader and text characteristics. More attention to text characteristics in assessing ORF would lead to more rigorous psychometric measures that would improve the precision by which student reading progress is assessed when using measures of ORF.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Grants P50 HD052117, K08 HD068545-01A1, and K99HD061689 as well as Reading for Understanding from the Institute of Education Sciences, U.S. Department of Education Grant R305F100013 that was awarded to Sharon Vaughn at The University of Texas at Austin. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD, the National Institutes of Health, or the Institute of Education Sciences, U.S. Department of Education.

References

Adams, MJ. Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press; 1990.

Adlof SM, Catts HW, Little TD. Should the simple view of reading include a fluency component? Reading and Writing. 2006; 19:933–958.

Albrecht J, O'Brien E. Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1993; 19:1061–1070.

Barth AE, Catts HW, Anthony JA. The component skills underlying reading fluency among adolescent readers: A latent variable approach. Reading and Writing. 2009; 22:567–590.

Best R, Floyd R, McNamara D. Differential competencies contributing to children's comprehension of narrative and expository texts. Reading Psychology. 2008; 29:137–164.

Betts J, Pickart M, Heistad D. An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating for alternate forms. Journal of School Psychology. 2009; 47:1–17.

Biancarosa G. Speed and time, test and sentences. Written Language and Literacy. 2005; 8:79–100.

Brasseur-Hock I, Hock M, Kieffer M, Biancarosa G, Deshler D. Adolescent struggling readers in urban schools: Results of a latent class analysis. Learning and Individual differences. 2011; 21:438–452.

Catts, H.; Adlof, S.; Hogan, T. Developmental changes in reading and reading disabilities. In: Catts, HW.; Kamhi, AG., editors. The connections between language and reading disabilities. Mahwah, NJ: Erlbaum; 2005. p. 25-40.

Cervetti G, Bravo M, Hiebert E, Pearson PD, Jaynes C. Text genre and science content: Ease of reading, comprehension, and reader preference. Reading Psychology. 2009; 30:487–511.

Christ TJ, Ardoin SP. Curriculum-based measurement of oral reading: Passage equivalence and probe-set development. Journal of School Psychology. 2009; 47:55–75.

Christ TJ, Silberglitt B. Estimates of the standard error of measurement for curriculum-based measures of oral reading fluency. School Psychology Review. 2007; 36:130–146.

Cirino P, Romain M, Barth AE, Tolar T, Fletcher JM, Vaughn S. Reading skill components and impairments in middle school struggling readers. Reading and Writing. (in press).

Collins A, Loftus E. A spreading-activation theory of semantic processing. Psychological Review. 1975; 82:407–428.

Crossley S, McCarthy P, Louwerse M, McNamara D. A linguistic analysis of simplified and authentic texts. Modern Language Journal. 2007; 91:15–30.

Culatta, B.; Horn, DG.; Merritt, DD. Expository text: Facilitating comprehension. In: Merritt, D.; Culatta, B., editors. Language intervention in the classroom. San Diego, CA: Singular; 1998. p. 215-276.

Daane, MC.; Campbell, JR.; Grigg, WS.; Goodman, MJ.; Oranje, A. Fourth-grade students reading aloud: NAEP 2002 Special Study of Oral Reading (NCES 2006–469). Washington, DC: National Center for Education Statistics; 2005.

Deno SL. Curriculum-based measurement: The emerging alternative. Exceptional Children. 1985; 52:219–232. [PubMed: 2934262]

Diakidoy I, Stylianou P, Karefillidou C, Papageorgiou C. The relationship between listening and reading comprehension of different types of text at increasing grade levels. Reading Psychology. 2005; 26:55–80.

Dunn, L.; Dunn, L. Peabody Picture Vocabulary Test–Revised. Circle Pines, MN: American Guidance Service; 1981.

Eason SH, Goldgerg 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.

Ehri, LC. Phases of acquisition in learning to read words and implications for teaching. In: Stainthorp, R.; Tomlinson, P., editors. Learning and teaching reading. London, England: British Journal of Educational Psychology Monograph Series II; 2002.

Ehri, LC.; Snowling, MJ. Developmental variation in word recognition. In: Stone, CA.; Silliman, ER.; Ehren, BJ.; Appel, K., editors. Handbook of language and literacy: Development and disorders. New York, NY: Guilford Press; 2004. p. 433-460.

Francis DJ, Santi KL, Barr C, Fletcher JM, Varisco A, Foorman BR. Form effects on the estimation of students' oral reading fluency using DIBELS. Journal of School Psychology. 2008; 46:315–342. [PubMed: 19083362]

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

Good, RH.; Jefferson, G. Contemporary perspectives on curriculum-based measurement validity. In: Shinn, M., editor. Advanced applications of curriculum-based measurement. New York, NY: Guilford Press; 1998. p. 61-88.

Gough, PB.; Hoover, W.; Peterson, CL. Some observations on the simple view of reading. In: Cornoldi, C.; Oakhill, J., editors. Reading comprehension difficulties. Mahwah, NJ: Erlbaum; 1996. p. 1-13.

Graesser, AC.; Clark, LF. Structures and procedures of implicit knowledge. Norwood, NJ: Ablex; 1985.

Graesser AC, Hoffman NL, Clark L. Structural components of reading time. Journal of Verbal Learning and Verbal Behavior. 1980; 19:135–151.

Graesser, AC.; Jeon, M.; Cai, Z.; McNamara, DS. Automatic analyses of language, discourse, and situation models. In: Auracher, J.; van Peer, W., editors. New beginnings in literary studies. Cambridge, England: Cambridge Scholars; 2008. p. 72-88.

Graesser A, Jeon M, Yang Y, Cai Z. Discourse cohesion in text and tutorial dialogue. Information Design Journal. 2007; 15:199–213.

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

Graesser AC, McNamara DS, Kulikowich JM. CohMetrix: Providing multilevel analysis of text characteristics. Educational Researcher. 2011; 40:223–234.

Graesser AC, Millis KK, Zwaan RA. Discourse comprehension. Annual Review of Psychology. 1997; 48:163–189.

Graesser A, Singer M, Trabasso T. Constructing inferences during narrative text comprehension. Psychological Review. 1994; 101:371–395. [PubMed: 7938337]

Haberlandt K, Graesser AC. Component processes in text comprehension and some of their interactions. Journal of Experimental Psychology: General. 1985; 114:357–374.

Heller, M. Reading-writing: Connections from theory to practice. White Plains, NY: Longman; 1995.

Hiebert EH. The effects of text difficulty on second graders' fluency development. Reading Psychology. 2005; 26:183–209.

Hintze JM, Daly EJ, Shaprio ES. An investigation of the effects of passage difficulty level on outcomes of oral reading fluency progress monitoring. School Psychology Review. 1998; 27:433–445.

Hintze JM, Silberglitt B. A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. School Psychology Review. 2005; 34:372–386.

Hock MF, Brassuer IF, Deshler DD, Catts HW, Marquis JG, Mark CA, Wu Stribling J. What is the reading component skill profile of adolescent struggling readers in urban schools? Learning Disability Quarterly. 2009; 32:21–38.

Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999; 6:1–55.

Jenkins JR, Fuchs LS, van den Broek P, Espin C, Deno SL. Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. Learning Disabilities: Research and Practice. 2003a; 18:237–245.

Jenkins JR, Fuchs LS, van den Broek P, Espin CA, Deno S. Sources of individual differences in reading comprehension and reading fluency. Journal of Educational Psychology. 2003b; 95:719–729.

Juel C. Retention and nonretention of at-risk readers in first grade and their subsequent reading achievement. Journal of Learning Disabilities. 1988; 21:571–580. [PubMed: 3199030]

Kaufman, AS.; Kaufman, NL. Kaufman Brief Intelligence Test Manual. 2nd ed.. San Antonio TX: Pearson; 2004.

Kim Y, Petscher Y, Schatschneider C, Foorman B. Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? Journal of Educational Psychology. 2010; 102:652–667.

Kintsch W. The role of knowledge in discourse comprehension: A constructive integration model. Psychological Review. 1988; 95:163–182. [PubMed: 3375398]

Lennon, C.; Burdick, H. The Lexile Framework as an approach for reading measurement and success. Durham, NC: MetaMetrics; 2004.

Lucas M, Tanenhaus MK, Carlson GN. Levels of representation in the interpretation of anaphoric reference and instrument inference. Memory & Cognition. 1990; 18:611–631. [PubMed: 2266863]

McCrudden, MT.; Magliano, J.; Schraw, G., editors. Text relevance and learning from text. Greenwich, CT: Information Age Publishing; 2012.

McNamara, AD., editor. Theories of text comprehension: The importance of reading strategies to theoretical foundations of reading comprehension. Mahwah, NJ: Erlbaum; 2007.

McNamara D, Louwerse M, McCarthy P, Graesser A. Coh-Metrix: Capturing linguistic features of cohesion. Discourse Processes. 2010; 47:292–330.

Muthén, LK.; Muthén, BO. Mplus user's guide. 5th ed.. Los Angeles, CA: Author; 2007.

Myers J, O'Brien E. Accessing the discourse representation during reading. Memory-Based Text Processing. 1998; 26:131–157.

Myers M, Paris SG. Children's metacognitive knowledge about reading. Journal of Educational Psychology. 1978; 70:680–690.

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

Nelson, J.; Perfetti, C.; Liben, D.; Liben, M. Measures of text difficulty: Testing their predictive value in grade levels and student performance. New York, NY: Student Achievement Partners; 2012.

O'Brien E, Albrecht E, Hakala C, Rizzella M. Activation and suppression of antecedents during reinstatement. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:626–634.

O'Connor R, Bell K, Harty K, Larkin L, Sackor S, Zigmond N. Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. Journal of Educational Psychology. 2002; 94:474–485.

O'Connor R, Swanson H, Geraghty C. Improvement in reading rate under independent and difficult text levels: Influences on word and comprehension skills. Journal of Educational Psychology. 2010; 102:1–19.

O'Reilly T, McNamara DS. The impact of science knowledge, reading skills, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. American Educational Research Journal. 2007; 44:161–196.

Pearons PD, Hiebert E. National reports in literacy: Building a scientific base for practice and policy. Educational Researcher. 2010; 39:286–294.

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

Perfetti, CA. Psycholinguistics and reading ability. In: Gernsbacher, MA., editor. Handbook of psycholinguistics. San Diego, CA: Academic Press; 1994. p. 849-894.

Perfetti CA. Reading ability: Lexical quality to comprehension. Scientific Studies of Reading. 2007; 11:357–383.

Perfetti, CA.; Hart, L. The lexical basis of comprehension skill. In: Gorfein, D., editor. On the consequences of meaning selection: Perspectives on resolving lexical ambiguity. Washington, DC: American Psychological Association; 2001.

Poncy BC, Skinner CH, Axtell PK. An investigation of the reliability and standard error of measurement of words read correctly per minute using curriculum-based measurement. Journal of Psychoeducational Assessment. 2005; 23:326–338.

Rayner K, Foorman BR, Perfetti CA, Pesetsky D, Seidenberg MS. How should reading be taught? Scientific American. 2002; 286:84–91. [PubMed: 11857904]

Reschley AL, Busch TW, Betts J, Deno ST, Long JD. Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. Journal of School Psychology. 2009; 47:427–469. [PubMed: 19808123]

SAS Institute. SAS/STAT(R) user's guide. 2nd ed.. Cary, NC: Author; 2009.

Smith E, Shoben E, Rips L. Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review. 1974; 81:214–241.

Stamboltzis A, Pumfrey P. Reading across genres: A review of the literature. Support for Learning. 2000; 15:58–61.

Stanovich KE. The impact of automaticity theory. Journal of Learning Disability. 1987; 20:167–168.

Stenner, AJ.; Burdick, H.; Sanford, EE.; Burdick, DS. The Lexile framework for reading technical report. Durham, NC: MetaMetrics; 2007.

Suh S, Trabasso T. Inferences during reading: Converging evidence from discourse anlysis, talk-aloud protocols, and recognition priming. Journal of Memory and Language. 1993; 32:279–300.

Tonjes, M.; Wolpow, R.; Zintz, M. Integrated content literacy. Boston, MA: McGraw-Hill; 1999.

Torgesen, JK.; Rashotte, CA.; Alexander, AW. Principles of fluency instruction in reading: Relationships with established empirical outcomes. In: Wolf, M., editor. Dyslexia, fluency, and the brain. Parkton, MD: York Press; 2001. p. 333-356.

Torgesen, JK.; Wagner, R.; Raschote, C. Test of Word Reading Efficiency. Austin, TX: PRO-ED; 1999.

Trabasso T, Magliano JP. Conscious understanding during comprehension. Discourse Process. 1996; 21:255–287.

van den Broek P. Using texts in science education: Cognitive processes and knowledge representation. Science. 2010; 23:453–456. [PubMed: 20413489]

van den Broek P, Rapp DN, Kendeou P. Integrating memory-based and constructionist approaches in accounts of reading comprehension. Discourse Processes. 2005; 39:299–316.

van den Broek, P.; Risden, K.; Fletcher, C.; Thurlow, R. A "landscape" view of reading: Fluctuating patterns of activation and the construction of a stable memory representation. In: Bruce, B.; Graesser, A., editors. Models of understanding text. Hillsdale, NJ: Erlbaum; 1996. p. 165-187.

van den Broek, PW.; Rohleder, L.; Narváez, D. Causal inferences in the comprehension of literary texts. In: Kreuz, RJ.; MacNealy, MS., editors. Empirical approaches to literature and aesthetics. Norwood, NJ: Ablex; 1996. p. 179-200.

Vaughn S, Wanzek J, Wexler J, Barth A, Cirino PT, Fletcher J, Francis D. The relative effects of group size on reading progress of older students with reading difficulties. Reading and Writing. 2010; 23:931–956. [PubMed: 21072131]

Westby, CE. The effects of culture on genre, structure and style of oral and written texts. In: Wallach, GP.; Butler, KG., editors. Language learning disabilities in school-age children and adolescents. New York, NY: Macmillan; 1994. p. 181-218.

Wiederholt, J.; Bryant, B. Gray Oral Reading Test–3. Austin, TX: PRO-ED; 1982.

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III Tests of Cognitive Abilities. Itasca, IL: Riverside; 2001.

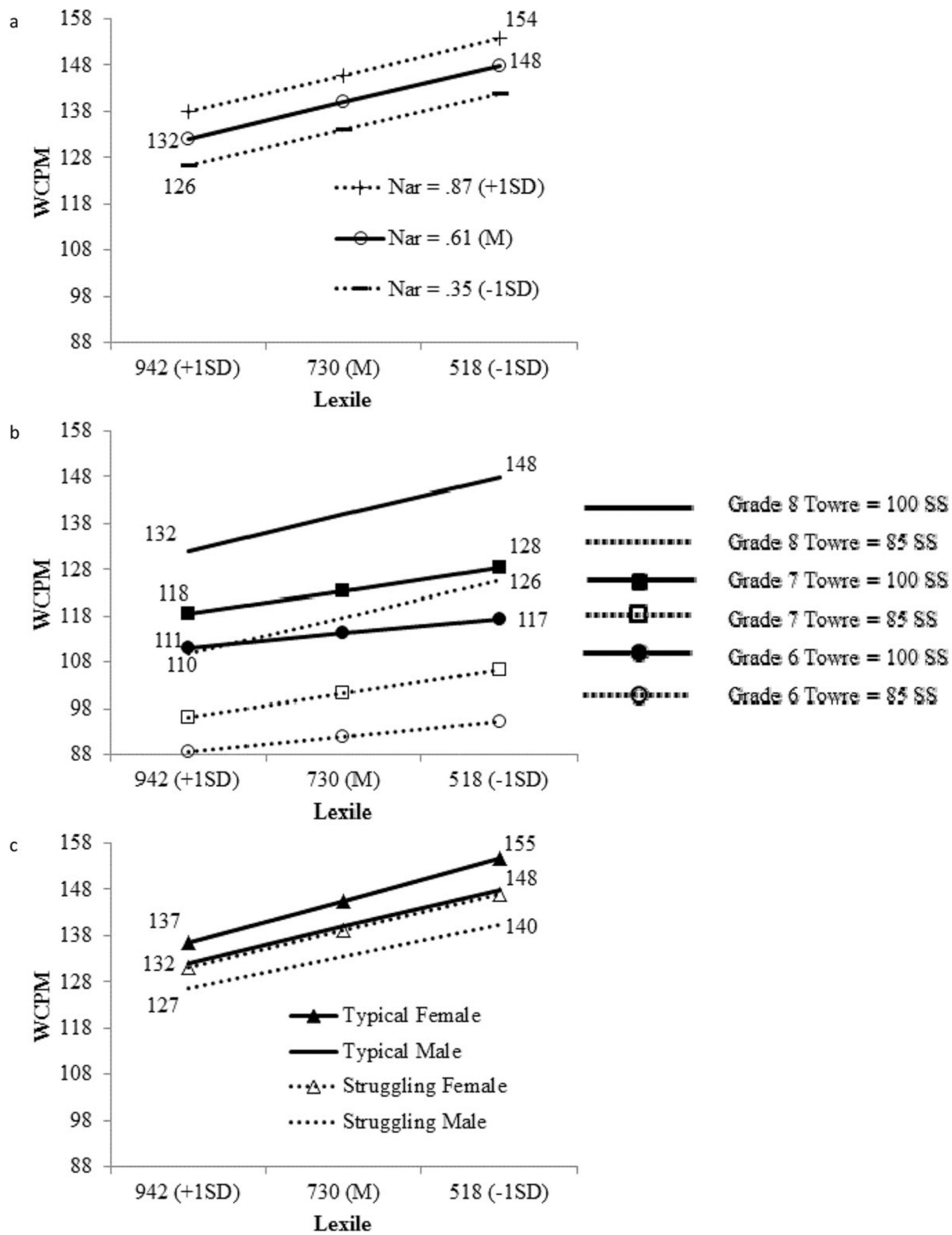

Figure 1.

a: Additive effects of Lexile and narrativity. b: Moderating effects of grade on text difficulty effects compared with between-student effects of word reading fluency on oral reading fluency (ORF). c: Moderating and additive effects of gender and reader type on ORF. WCPM = words correct per minute; Nar = Narrative; Towre = Test of Word Reading Efficiency; SS = standard score.

Table 1

Distribution (*N*) of ORF-PF Passage Assignments by Grade and Reader Group

Passage number	Passage set	Lexile band	Lexile	Text type	Page length	*N* Grade 6 S	Grade 6 T	Grade 7 S	Grade 7 T	Grade 8 S	Grade 8 T	Total
1	A	350–460	390	N	1.5	88	53					141
2	B		410	E	0.5	78	53					131
3	C		440	E	1.0	90	51					141
4	D		450	E	0.5	81	50					131
5	E		440	E	0.5	81	50					131
6	A	461–570	500	E	1.0	88	53	43	39			223
7	B		550	N	1.5	78	53	48	36			215
8	C		560	E	1.0	90	51	46	36			223
9	D		500	N	0.5	81	50	47	36			214
10	E		510	N	0.5	81	50	43	33			207
11	A	571–680	650	N	1.5	88	53	43	39	81	52	356
12	B		600	E	1.5	78	53	48	36	71	48	334
13	C		660	E	1.0	90	51	46	36	75	56	354
14	D		590	N	1.5	81	50	47	36	72	58	344
15	E		640	N	1.5	81	50	43	33	84	53	344
16	A	681–790	700	E	0.5	88	53	43	39	81	52	356
17	B		740	E	1.5	78	53	48	36	71	48	334
18	C		690	E	1.5	90	51	46	36	75	56	354
19	D		710	N	1.5	81	50	47	36	72	58	344
20	E		780	E	1.0	81	50	43	33	84	53	344
21	A	791–900	840	N	1.0	88	53	43	39	81	52	356
22	B		820	N	1.0	78	53	48	36	71	48	334
23	C		800	E	1.5	90	51	46	36	75	56	354
24	D		800	E	1.5	81	50	47	36	72	58	344
25	E		830	E	1.0	81	50	43	33	84	53	344
26	A	901–1,010	910	E	1.5			43	39	81	52	215

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Passage number	Passage set	Lexile band	Lexile	Text type	Page length	*N* Grade 6 S	Grade 6 T	Grade 7 S	Grade 7 T	Grade 8 S	Grade 8 T	Total
27	B		950	E	1.5			48	36	71	48	203
28	C		920	N	1.5			46	36	75	56	213
29	D		970	N	1.5			47	36	72	58	213
30	E		950	E	1.5			43	33	84	53	213
31	A	1,011–1,120	1020	E	1.0					81	52	133
32	B		1050	E	1.5					71	48	119
33	C		1110	E	1.5					75	56	131
34	D		1030	E	1.0					72	58	130
35	E		1050	E	1.5					84	53	137

Note. N = Narrative; E = Expository. Reader groups: S = Struggling; T = Typical. Students within grade and reader group were randomly assigned to passage sets. ORF-PF = oral reading fluency-passage fluency.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Demographics and Achievement by Passage Set

Demographic/achievement	Passage set A	B	C	D	E
	Grade 6				
N	141	131	141	131	131
M (*SD*)					
Age–Years	11.53 (0.72)	11.35 (0.59)	11.53 (0.71)	11.51 (0.64)	11.46 (0.60)
KBIT Verbal	89.43 (13.27)	91.53 (15.27)	88.03 (15.30)	88.14 (14.10)	89.89 (14.27)
TOWRE	95.43 (14.91)	97.99 (15.10)	96.37 (15.45)	97.49 (15.92)	97.22 (17.36)
LWID	95.74 (13.54)	99.82 (13.74)	95.95 (14.75)	97.39 (13.62)	97.09 (15.53)
Percent					
Female	51	52	47	47	53
Struggling	62	60	64	62	62
Reduced/free lunch	78	82	84	82	88
African American	46	40	49	47	41
Hispanic	34	38	34	32	36
Caucasian	18	19	15	18	20
Asian	1	3	2	2	
	Grade 7				
N	82	84	82	83	76
M (*SD*)					
Age–Years	12.37 (0.72)	12.34 (0.54)	12.39 (0.70)	12.41 (0.62)	12.30 (0.56)
KBIT Verbal	90.91 (14.28)	91.61 (13.70)	91.18 (14.60)	93.04 (17.07)	89.09 (11.99)
TOWRE	97.90 (16.75)	97.55 (16.78)	96.72 (14.76)	99.61 (17.68)	97.39 (13.91)
LWID	97.22 (13.77)	97.67 (13.70)	97.98 (14.85)	99.75 (16.29)	98.17 (13.68)
Percent					
Female	41	51	52	55	49
Struggling	52	57	56	57	57
Reduced/free lunch	79	80	76	77	74
African American	39	37	39	31	43

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Demographic/achievement	Passage set A	B	C	D	E
Hispanic	44	37	44	37	36
Caucasian	13	21	12	25	18
Asian	4	5	2	6	
			Grade 8		
N	133	119	131	130	137
M (*SD*)					
Age–Years[a]	13.40 (0.61)	13.20 (0.48)	13.33 (0.54)	13.24 (0.48)	13.27 (0.57)
KBIT Verbal	89.65 (13.67)	91.72 (13.95)	92.84 (14.98)	93.88 (15.66)	93.26 (15.29)
TOWRE	94.38 (15.51)	97.82 (14.50)	96.48 (16.65)	99.25 (15.23)	95.81 (15.78)
LWID[b]	94.81 (14.52)	96.64 (13.11)	96.78 (12.62)	99.92 (13.51)	97.26 (12.94)
Percent					
Female	51	43	51	52	52
Struggling	61	60	57	55	61
Reduced/free lunch	76	82	82	82	71
African American	41	40	37	42	29
Hispanic	29	40	37	35	43
Caucasian	25	16	21	18	26
Asian	5	3	6	5	

Note. Analysis of variance (ANOVA) and chi-square analyses were conducted for the demographic and achievement measures. For significant ANOVA's, Tukey's post hoc analyses were conducted to determine which pairs of groups differed. Within grade, passage set groups were not significantly different ($p > .05$), except as noted. KBIT Verbal = Verbal Knowledge subtest of the Kaufman Brief Intelligence Test-2 (Kaufman & Kaufman, 2004); TOWRE = Test of Word Reading Efficiency (Torgesen et al., 1999); LWID = Woodcock-Johnson III Letter-Word Identification (Woodcock et al., 2001).

[a] Passage Set Group A > Passage Set Group B ($p < .05$).

[b] Passage Set Group D > Passage Set Group A ($p < .05$).

Table 3

Mean WCPM by Grade, Passage Set, and Passage

Passage set and number	Grade 6				Grade 7				Grade 8			
	M	***SD***	**Skew**	**Kurtosis**	***M***	***SD***	**Skew**	**Kurtosis**	***M***	***SD***	**Skew**	**Kurtosis**
A												
#	*N* = 141				*N* = 82				*N* = 133			
1	128.67	36.72	0.66	0.34								
6	100.11	36.96	0.41	0.10	120.63	39.90	−0.32	−0.28				
11	106.86	38.80	0.66	0.46	115.83	42.65	0.25	0.10	134.40	43.34	−0.34	−0.01
16	108.61	35.07	0.68	0.27	116.77	37.96	−0.15	0.72	128.32	38.47	−0.39	0.06
21	116.20	40.24	1.08	1.07	124.05	43.55	0.65	0.45	138.02	44.93	−0.01	0.10
26					112.66	33.56	0.22	0.62	123.14	33.40	−0.33	0.45
31									124.71	34.13	−0.46	0.75
B												
#	*N* = 131				*N* = 84				*N* = 119			
2	112.38	33.90	0.19	0.01								
7	128.20	32.42	−0.01	0.17	138.13	36.54	−0.24	0.62				
12	106.63	34.15	−0.11	−0.35	114.65	37.63	−0.40	−0.21	131.12	36.72	−0.29	−0.42
17	108.32	35.98	0.03	0.00	114.32	37.87	−0.15	−0.59	127.78	36.13	0.04	−0.54
22	103.82	33.44	0.56	0.24	115.38	38.53	0.12	0.24	125.84	37.62	0.30	−0.41
27					113.21	36.65	−0.30	0.33	121.50	33.43	0.27	0.04
32									126.00	30.39	−0.05	0.52
C												
#	*N* = 141				*N* = 82				*N*= 131			
3	126.14	36.33	−0.20	0.01								
8	111.60	35.12	−0.19	−0.45	128.22	38.55	−0.47	0.16				
13	119.23	31.80	−0.17	0.23	130.60	32.20	−0.05	0.16	145.66	38.80	−0.30	0.04
18	108.80	34.16	−0.14	−0.04	119.77	36.32	−0.02	−0.06	131.67	38.79	−0.27	0.03
23	114.16	32.28	−0.31	−0.31	122.11	32.39	−0.13	−0.19	135.37	36.80	−0.35	0.03
28					121.41	36.00	0.29	−0.22	132.89	40.11	−0.03	−0.08
33									125.18	33.79	−0.73	0.36

Passage set and number	Grade 6				Grade 7				Grade 8			
	M	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis
D												
#	*N* = 131				*N* = 83				*N*= 130			
4	124.24	33.70	−0.34	0.71								
9	125.15	36.52	−0.31	0.84	142.88	46.74	0.29	0.57				
14	132.43	33.24	−0.52	1.83	143.93	40.14	−0.42	−0.04	163.16	35.84	−0.12	0.40
19	123.83	32.22	−0.38	1.14	137.07	41.78	−0.20	−0.61	153.55	38.57	−0.01	0.07
24	123.79	36.17	−0.36	0.56	135.42	43.93	−0.53	−0.38	155.25	36.75	−0.33	−0.28
29					128.82	44.87	−0.23	−0.71	144.74	40.63	−0.14	−0.78
34									139.33	37.45	0.12	0.02
E												
#	*N* = 131				*N*=76				*N*= 137			
5	126.72	43.03	0.18	0.75								
10	145.11	46.22	−0.19	0.26	156.91	38.91	−0.20	0.74				
15	117.98	41.12	0.02	0.11	131.30	38.17	0.17	1.11	152.05	38.85	0.01	−0.21
20	115.93	37.49	−0.45	0.43	128.28	31.38	−0.45	0.56	139.72	34.45	−0.21	0.00
25	101.69	35.31	−0.46	0.16	114.37	26.15	−0.52	0.21	125.05	34.13	−0.03	0.62
30					105.12	31.78	0.01	0.17	114.91	36.98	−0.10	−0.34
35									98.31	29.25	0.08	0.14

Note. Within grade and passage set, the same students read all five passages, and the passages are ordered from least to highest Lexile score. Passages listed in corresponding orders across passage sets within grade fall within the same Lexile band (e.g., Passages 1–5 were the first passages read in Sets A–D, respectively, within Grade 6, and all fall within Lexile Band 350–460; Passages 6–10 were the second passages read in Sets A–D, respectively, within Grade 6, and all fall within Lexile Band 461–570, etc.). WCPM = words correct per minute.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Student-Level Effects on Reading Fluency (WCPM)

Parameter	Fixed effect	*t*	*p*	Variance	*z*	*p*
		Individually				
Intercept (Grade 8)	133.64	95.13	<.01	1130.19	25.19	<.01
Grade 6	−16.14	−8.20	<.01			
Grade 7	−8.01	−3.54	<.01			
Intercept (Male)	120.03	97.94	<.01	1149.90	25.28	<.01
Female	10.88	6.27	<.01			
Intercept (Typical)	145.64	119.08	<.01	900.67	24.17	<.01
Struggling	−33.98	−21.41	<.01			
Intercept (KBIT = 100)	135.33	145.42	<.01	929.79	24.41	<.01
per *SD* (15 SS points)	16.25	20.05	<.01			
Intercept (TOWRE = 100)	130.89	239.90	<.01	348.83	17.95	<.01
per *SD* (15 SS points)	27.49	53.83	<.01			
Intercept (LWID = 100)	130.16	199.71	<.01	560.22	21.49	<.01
per *SD* (15 SS points)	26.69	38.91	<.01			
		Combined: Main effects only				
Intercept	142.21	129.05	<.01	234.11	14.77	<.01
Grade 6	−15.66	−14.52	<.01			
Grade 7	−10.09	−8.16	<.01			
Female	5.51	5.77	<.01			
Struggling	−6.71	−5.82	<.01			
KBIT (per *SD*)	2.89	4.72	<.01			
TOWRE (per *SD*)	21.17	33.33	<.01			
LWID (per *SD*)	6.32	7.80	<.01			
ρ				0.45	22.19	<.01
Residual				369.57	28.74	<.01
		Combined: Main and interaction effects				
Intercept	141.86	75.73	<.01	234.21	14.73	<.01
Grade 6	−16.73	−6.49	<.01			

Parameter	Fixed effect	*t*	*p*	Variance	*z*	*p*
Grade 7	−9.51	−3.27	<.01			
Female	6.20	2.55	.01			
Struggling	−6.54	−2.74	<.01			
KBIT	2.82	4.60	<.01			
TOWRE	21.15	33.25	<.01			
LWID (per *SD*)	6.39	7.88	<.01			
Grade 6 Female	0.27	0.08	.94			
Grade 7 Female	0.90	0.24	.81			
Grade 6 Struggling	1.62	0.51	.61			
Grade 7 Struggling	0.67	0.18	.85			
Female Struggling	−0.44	−0.14	.89			
Grade 6 Female Struggling	−0.24	−0.05	.96			
Grade 7 Female Struggling	−5.71	−1.13	.26			
ρ				0.45	22.19	<.01
Residual				369.83	28.69	<.01

Note. Except as noted in table, the Intercept represents a Grade 8 Typical Male with average IQ, Decoding, and Word Reading Fluency (standard score [SS] = 100). WCPM = words correct per minute; KBIT = Verbal Knowledge subtest of the Kaufman Brief Intelligence Test-2 (Kaufman & Kaufman, 2004); TOWRE = Test of Word Reading Efficiency (Torgesen et al., 1999); LWID = Woodcock-Johnson III Letter-Word Identification (Woodcock et al., 2001). The *SD* for KBIT, TOWRE, and LWID is 15 SS points.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Text Level Effects on Reading Fluency (WCPM)

Model	Fixed effect	*t*	*p*	Variance	*z*	*p*	Variance explained[a] (%)
			Single predictor model				
Lexile	−8.62	−30.41	<.01	64.70	10.66	<.01	51.9
Syntactic simplicity	1.85	7.48	<.01	0.00			11.8
Narrativity	6.37	31.07	<.01	10.37	3.81	<.01	34.2
Text type (Expository)	−7.55	−21.69	<.01	2.22	0.58	.28	16.1
Word concreteness	−0.21	−0.85	.39	2.32	0.85	.20	1.4
Referential cohesion	1.61	9.70	<.01	0.00			0.7
Deep cohesion	−0.51	−2.94	<.01	0.00			0.2
Page length				3.75	0.72	.24	3.6
0.5 (vs. 1.5)	2.48	4.23	<.01				
1 (vs. 1.5)	0.73	1.78	.0759				
			Lexile and syntactic simplicity model				
Lexile	−12.81	−33.53	<.01	68.42	11.28	<.01	
Syntactic simplicity	−5.07	−16.28	<.01	Fixed			
Residual				171.92	34.94	<.01	54.4
			Narrativity and text type model				
Narrativity	6.06	22.64	<.01	10.44	3.82	<.01	
Text type (Expository)	−0.85	−1.87	.06	Fixed			
Residual				247.22	32.41	<.01	34.4
			Combined model				
Intercept	124.65	136.72	<.01	1355.83	28.22	<.01	
Lexile	−5.27	−17.06	<.01	26.39	4.18	<.01	
Narrativity	5.41	24.11	<.01	4.34	1.88	.03	
Referential cohesion	1.50	9.27	<.01	Fixed			
Deep cohesion	−0.62	−3.71	<.01	Fixed			
Page length effect (.5)[b]	4.38	7.36	<.01	Fixed			

Model	Fixed effect	*t*	*p*	Variance	*z*	*p*	Variance explained[a] (%)
Page length effect (1)	0.03	0.07	.94	Fixed			
ρ				0.10	3.56	.0004	
Residual				186.83	27.74	<.0001	51.5

Note. Each text parameter was first evaluated separately (Single predictor model). Lexile and Syntactic simplicity were evaluated together, and Narrativity and Text type were evaluated together. The final model (Combined model) excluded syntactic simplicity, text type, and word concreteness. WCPM = words correct per minute. Dashes represent xxx.

[a]% explained variance was calculated by calculating the difference in residual variance between the unconditional and conditional models and dividing the result by the residual variance from the unconditional model (376.97).

[b]The reference for the page length effects is 1.5 pages.

Table 6

Final Model of Text- and Student-Level Effects on Reading Fluency (WCPM)

Model	Fixed effect	*t*	*p*	Variance	*z*	*p*
Intercept	139.96	119.71	<.01	313.76	23.37	<.01
Student-level predictor						
Grade 6	−25.73	−23.49	<.01			
Grade 7	−16.52	−13.47	<.01			
Gender (Female)	5.59	5.96	<.01			
Reader Type (Struggling)	−6.50	−5.75	<.01			
KBIT Verbal	2.94	5.02	<.01			
TOWRE	22.24	33.94	<.01			
Letter Word ID	5.88	7.92	<.01			
Text-level predictor						
Lexile (L)	−7.86	−12.89	<.01	11.13	1.8	.04
Narrativity (N)	5.89	12.53	<.01	0.004	0	.50
Referential cohesion	1.83	11.41	<.01	Fixed		
Page Length 0.5	6.31	10.32	<.01	Fixed		
Page Length 1	0.08	0.22	.83			
Student/text interactions:						
L × Grade 6	4.65	7.48	<.01			
L × Grade 7	2.90	3.91	<.01			
L × Gender (Female)	−1.15	−2.22	.03			
L × Reader Type (Struggling)	1.08	2.06	.04			
N × Grade 6	0.92	2.04	.04			
N × Grade 7	−1.03	−1.92	.06			
N × Gender (Female)	1.37	3.56	<.01			
N × Reader Type (Struggling)	−1.91	−4.88	<.01			
Autoregressive				0.16	5.61	<.01
Residual				202.49	26.07	<.01

Note. The Intercept represents a Grade 8 Typical Male with IQ, Decoding, and Word Reading Fluency standard scores = 100, reading a 730-Lexile narrative passage that is 1.5 pages in length. WCPM = words correct per minute; KBIT Verbal = Verbal Knowledge subtest of the Kaufman Brief Intelligence Test-2 (Kaufman & Kaufman, 2004); TOWRE = Test of Word Reading Efficiency (Torgesen et al., 1999); Letter Word ID = Woodcock-Johnson III Letter-Word Identification (Woodcock et al., 2001). KBIT Verbal, TOWRE, and Letter Word ID effects are based on *z*-scores using population-normed values ($M = 100$, $SD = 15$).