Research Article

Hammill Institute on Disabilities

Journal of Learning Disabilities
2021, Vol. 54(3) 155–169

DOI: 10.1177/0022219420932139
journaloflearningdisabilities.sagepub.com
SAGE

The Differential Importance of Component Skills on Reading Comprehension Test Performance Among Struggling Adolescent Readers

Nathan H. Clemens, PhD[1], Yu-Yu Hsiao, PhD[2], Kejin Lee, PhD[1], Amanda Martinez-Lincoln, PhD[1], Clinton Moore, MA[1], Jessica Toste, PhD[1], and Leslie Simmons, MEd[3]

Abstract

Reading comprehension tests vary in format and characteristics, which may influence the extent to which component skills are involved in test performance. With students in Grades 6 to 8 with reading difficulties, dominance analyses examined the differential importance of component reading and language skills (word- and text-reading fluency, vocabulary, listening comprehension, and working memory) on several standardized tests of reading comprehension: *The Gates-MacGinitie Reading Test*, 4th edition (GMRT), *Group Reading Assessment and Diagnostic Evaluation*, *Gray Oral Reading Test*, 5th edition (GORT-5), and the *Test of Silent Reading Efficiency and Comprehension* (TOSREC). Students' word- and text-reading fluency skills were generally the most dominant predictors of performance on most reading comprehension tests, especially those with a time limit (GMRT and TOSREC). Listening comprehension was most important on the GORT-5, a test in which students read passages orally and listen to questions read by an examiner. Working memory was the least important component skill across the reading comprehension tests. Overall, results were consistent with previous work indicating that reading comprehension measures vary with regard to the skills or knowledge sources that are most important for test performance and extend these findings to struggling adolescent readers. Implications for research and practice are discussed.

Keywords

assessment, comprehension, reading, middle school, age

Reading is one of the primary ways in which students in the middle and secondary grades are expected to acquire new knowledge, which makes reading comprehension central to academic success and post-secondary opportunities. Unfortunately, many adolescents experience difficulties in reading comprehension (National Assessment of Educational Progress, 2019). In the middle grades and beyond, increasing text complexity and expectations that students independently learn from print (Swanson et al., 2016) make reading comprehension a primary area of interest when assessing adolescents' reading skills.

Several standardized, commercially available measures of reading comprehension exist; most of which involve answering questions about a text or supplying a missing word. Although many tests make the measurement of reading comprehension appear simple, conclusions regarding the true status of students' reading comprehension skills are complicated by the complex nature of the construct and the role of underlying component skills that can differentially influence students' test scores. With a group of struggling readers in Grades 6 to 8, this study investigated the differential influence of several reading and language skills on several tests of reading comprehension.

The Challenge of Measuring Reading Comprehension

Several factors make measuring reading comprehension difficult. First, any attempt to measure reading comprehension is

[1]The University of Texas at Austin, USA
[2]The University of New Mexico, Albuquerque, USA
[3]Texas A&M University, College Station, USA

Corresponding Author:
Nathan H. Clemens, PhD, Department of Special Education, University of Texas at Austin, SZB 408M, 1 University Station, D5300, Austin, TX 78712, USA.
Email: nathan.clemens@austin.utexas.edu

challenging because a reader's accurate mental representation of text is not something that can be directly observed as it occurs. Consequently, measurement relies on indirect assessment of students' understanding after reading (Pearson & Hamm, 2005), which requires an examiner to make assumptions about the quality and adequacy of reading comprehension that occurred (Fletcher, 2006).

The second aspect that makes measuring reading comprehension difficult is its complexity—it is the product of multiple student-, task-, and situation-level factors. Important student-level skill and knowledge variables include word recognition (i.e., decoding), text-reading fluency, vocabulary knowledge, linguistic comprehension and reasoning, background knowledge, inference-making, working memory, and attention (Ahmed et al., 2016; Cain & Oakhill, 2009; Collins et al., 2019; Cromley & Azevedo, 2007; Gough & Tunmer, 1986; Peng et al., 2018; C. Perfetti & Stafura, 2014). These student-level skills interact with text features (e.g., genre, text structure, text complexity) and the demands of the reading situation (e.g., acquire knowledge, being tested, or for pleasure) to influence the extent to which a reader understands and learns from text (Miller et al., 2014; RAND Reading Study Group, 2002).

A third problematic factor in measuring reading comprehension is the significant variability in test characteristics. Reading comprehension tests differ widely in their response formats (e.g., multiple-choice, oral or written open-ended response, sentence verification, cloze, retell); whether students are asked to read orally or silently; the amount of text students are asked to read (e.g., single sentences, multiparagraph passages); group or individual administration format; and whether administration is timed. A result of this variability is that reading comprehension test scores do not correlate with one another as strongly as would be expected of measures that purport to assess the same construct (Cutting & Scarborough, 2006; Francis et al., 2006; Keenan et al., 2008; Keenan & Meenan, 2014).

In addition, the variability in reading comprehension test formats can minimize or magnify the influence of student-level skill and knowledge variables on test performance. For example, students' word-reading skills tend to play a greater role in explaining performance on reading comprehension measures that use single sentences or very short passages (Andreassen & Bråten, 2010; Francis et al., 2006; Keenan et al., 2008) because a single reading error is more likely to disrupt comprehension in contrast to longer texts that provide more context to mitigate the influence of a decoding error (Keenan et al., 2008). Word-reading skills also play a greater role in explaining reading comprehension test performance for younger or weaker readers, on tests that use cloze compared with multiple-choice response formats, when students read silently compared with reading aloud, and when tests consist of narrative texts rather than expository texts (Best et al., 2008; Francis et al., 2006; García & Cain, 2014; Nation & Snowling, 1997; Spear-Swerling, 2004).

Other skill and knowledge variables have been shown to differentially affect reading comprehension test performance. Listening comprehension and linguistic knowledge appear to have a greater influence (relative to other skills, such as word reading) on measures that include longer passages (Nation & Snowling, 1997) or those in which questions are read aloud to students (Keenan et al., 2008). Best and colleagues (2008) observed that general knowledge was more important on tests that used expository texts compared with narrative texts. Reading fluency may play a greater role in performance on tests that are scored based on correct responses within a time limit (Kendeou et al., 2012). Test achievement also appears to vary in terms of the contribution of working memory, with some evidence suggesting a stronger influence of working memory on measures with shorter passages (Keenan & Meenan, 2014) and tests in which the passages are removed from sight while students respond to questions (Andreassen & Bråten, 2010; but see conflicting evidence from Cutting & Scarborough, 2006).

In summary, the complex and covert nature of reading comprehension, combined with the broad variability in how it is measured, poses considerable difficulties for decision-making based on the results of reading comprehension tests. These issues have prompted concerns about reducing reading comprehension to a single test score which may inadequately represent the construct (Fletcher, 2006; Kamhi & Catts, 2017; Wixson, 2017). Reading comprehension measures lack agreement in identifying students with poor reading comprehension (Keenan & Meenan, 2014), and some tests may make struggling readers look like typical achievers (Collins et al., 2018). In short, reading comprehension tests may result in misleading conclusions regarding whether poor test performance is the result of problems with high-order text processing or some other basic reading skill or knowledge variable.

Unanswered Questions

At least three factors warrant additional research. The majority of the participants in previous studies were in middle elementary grades. Some studies included adolescents as part of a range spanning elementary to secondary grades (e.g., Cutting & Scarborough, 2006; Keenan et al., 2008); however, studies have not investigated the differential relations of component skills on test performance specifically among adolescent readers. Adolescents are an important population for reading comprehension research given the common expectations that they independently learn from text to succeed in school.

Second, the majority of studies have included students representing a full range of reading skill levels. These studies

have been important in revealing the relative importance of component skills to reading comprehension test performance across achievement levels, but students with reading difficulties are often the main recipients of reading comprehension tests in evaluation or intervention research contexts. Test results may be used alone or with other tests to determine the need for continued intervention support, evaluate the efficacy of experimental interventions, inform instructional adjustments, or evaluate eligibility for special education. In short, struggling readers are a group in which reading comprehension tests are most often used and for whom interpretation of test results has significant implications. Better understanding what skills are most influential to performance across various tests of reading comprehension is important for informing assessment, intervention, and research with struggling adolescent readers. To date, no study has investigated the differential importance of reading and language component skills across various tests of reading comprehension specifically with a sample of struggling adolescent readers.

Third, studies that have investigated the differential influence of component skills on reading comprehension test performance have primarily relied on correlation and multiple regression analyses. Although these analyses reveal the strength of relations among variables and the proportion of unique variance accounted for by the predictors in a model, they may not indicate the variable(s) that are *most important* in explaining performance on different tests of reading comprehension. An alternative exists that can provide this type of insight.

Dominance Analysis

Dominance analysis (Azen & Budescu, 2003; Budescu, 1993) is an application of multiple regression and is used to determine the relative importance of individual predictors of a dependent variable. Dominance analysis involves testing all possible combinations of predictors in separate regression models (i.e., subset models) to reveal the additional contribution of each predictor relative to other predictors under study. Dominance analysis has been used in several studies to identify the skills or measures that are most important in predicting reading skills on a concurrent or longitudinal basis (e.g., Clemens et al., 2019; Fuchs et al., 2004; Mellard et al., 2012; Schatschneider et al., 2004; Tighe & Schatschneider, 2014). However, dominance analysis has not been used to investigate the extent to which component reading skills and knowledge sources differentially predict performance across different tests of reading comprehension.

Study Purpose

The purpose of this study was to investigate the differential importance of reading and language skills for performance on a set of standardized, commercially available tests of reading comprehension among struggling adolescent readers. We examined the influence of word-reading fluency, oral reading fluency (ORF), vocabulary, listening comprehension, and working memory on tests of reading comprehension that differed in question and response format, passage length, and timing. We focused on struggling readers to inform future work with students who are often the recipients of such assessments and for whom test results have considerable consequences for research and educational decision-making.

Method

Participants

Participants were drawn from an initial sample of 233 students in sixth through eighth grades from two schools (one rural and one suburban) in the southwest United States. Students were participating in a randomized controlled trial (RCT) of a reading comprehension intervention (Fogarty et al., 2017). The present analyses utilized only pretest data (i.e., before intervention began).

From the sample of 233 students participating in the RCT, we identified a subsample of 180 students with below-average and well below average reading comprehension skills for the present analyses. The rationale and procedures for the identification of the low-comprehension subsample are detailed below. The analysis sample was 51.7% female, and 30.6% Black, 28.3% Hispanic, 23.9% White, 2.8% Asian, and 13.9% other/multiple ethnicities. Students who were eligible for special education services (7.8%) or identified as an English learner while receiving reading instruction in English (14.4%) were included.

Reading Comprehension Measures

Gates-MacGinitie Reading Test, 4th edition (GMRT) reading comprehension. The Reading Comprehension subtest from the GMRT (MacGinitie et al., 2002) contains 11 narrative and informational passages ranging from 3 to 15 sentences. Each passage is followed by three to six multiple-choice questions, and the score is based on the number of questions answered correctly in 35 min. Students read and answer the questions silently during a group-administered session. Kuder–Richardson 20 (KR-20) reliability coefficients range from .88 to .94 for Form S in Grades 6 to 8 (Maria & Hughes, 2008).

Group Reading Assessment and Diagnostic Evaluation (GRADE). The GRADE (Williams, 2001) includes two untimed, group-administered subtests of reading comprehension. Given that the subtests are different in format we investigated the contributions of the component skills to each

subtest, as well as on the *Total Comprehension Scale* which is a composite of the two subtests. On the *Sentence Comprehension subtest*, students read 19 sentences, each with a missing word, and select the word that best completes the sentence from five answer choices. Test authors reported coefficient alpha ranging from .83 to .88 for Grades 6 to 8. On the Passage Comprehension subtest, students read six passages of narrative or expository text and answer five multiple-choice questions per passage. Passages ranged from 8 to 30 sentences in length. Coefficient alpha estimates range from .85 to .88 (Williams, 2001). Grade-level versions of Form A were administered.

Gray Oral Reading Test, 5th edition (GORT-5) comprehension. The GORT-5 (Wiederholt & Bryant, 2012) is an individually administered assessment in which students orally read a set of narrative and informational passages (6–12 sentences in length) while the examiner records reading errors and rate. The passage is removed from view after reading and the examiner asks the student a series of open-ended questions. Responses are scored as correct or incorrect based on test guidelines and passages are administered until a fluency ceiling is reached. The Comprehension scaled score was used in our analyses, which ranges from 1 to 20 (50th percentile = 10). Coefficient alpha for the Comprehension ranges from .93 to .95 for students ages 11 to 14 (Wiederholt & Bryant, 2012).

Test of Silent Reading Efficiency and Comprehension (TOSREC). The TOSREC (Wagner et al., 2010) is a sentence verification task. Students are allotted 3 min to read a series of individual sentences and indicate whether each is factually true. Scores consist of the number of correct minus incorrect responses. Alternate-form reliability estimates for Grades 6 to 8 range from .85 to .93 (Wagner et al., 2010). Grade-level versions of Form O were administered to each participant. The TOSREC has been used in studies as an index of reading comprehension (e.g., Ahmed et al., 2016; Lonigan & Burgess, 2017).

Component Constructs and Measures

Sight Word Efficiency (SWE). The SWE subtest of the *Test of Word Reading Efficiency 2nd edition* (TOWRE-2; Torgesen et al., 2012) was used to assess decontextualized word recognition proficiency. SWE scores consist of the number of words read correctly in 45 s from a list that increases in difficulty. Word-list reading is considered a "pure" form of word identification because it does not afford the benefits of contextual understanding, and the combination of accuracy and speed on SWE offers a more complete view of word-reading proficiency than accuracy alone (Torgesen et al.). Test–retest reliability for students' ages 8 to 18 years ranges from .84 to .93. Form A was administered to all students.

ORF. ORF was measured using a passage reading fluency probe from the EasyCBM system (University of Oregon, 2008). Students orally read an assigned passage while the examiner records the number of words read correctly in 1 min. We administered a seventh-grade passage to students in Grades 6 to 8 (see Note 1). EasyCBM passages demonstrate alternate-form reliability coefficients ranging from .75 to .96 with an average of .91.

Vocabulary. *The Vocabulary subtest* from the GRADE is an untimed, group-administered multiple-choice test. Each item includes a short phrase with a target word in bold type, and students select the closest synonym of the target word from five answer choices. Coefficient alpha ranges from .86 to .88 for students in Grades 6 to 8 (Williams, 2001). Students were administered Form A corresponding to their grade-level.

Listening comprehension. The listening comprehension subtest from the GRADE is an untimed, group-administered multiple-choice test. For each item, the examiner reads a 1 to 2 sentence passage aloud, and students choose a picture from a set of four choices that best represents the information in the sentence(s). Coefficient alpha estimates range from .63 to .81 for students in Grades 6 to 8 (Williams, 2001). Students were administered grade-level Form A.

Working memory. Working memory was assessed with the individually administered digits reversed subtest from the Woodcock-Johnson III Normative Update (Woodcock et al., 2006). For each item the examiner reads a set of digits, after which the student is asked to repeat the digits in reverse order. The number of digits in the set increases across items. Median split-half reliability for ages 2 to 19 is .87. Reversed digit-span tasks are one of the most widely used measures in research and practice to evaluate working memory (Gathercole et al., 2004; Kasper et al., 2012).

Procedures

Assessments were conducted within a 3-week period at the start of the school year by research staff experienced in data collection. Staff received 7 hr of training and practice on the specific measures. Initial reliability was established in mock testing sessions, which were repeated until each test administrator demonstrated 100% procedural and scoring fidelity. In-field reliability was established whereby test administrators were required to demonstrate 100% procedural fidelity and 95% inter-scorer agreement before being permitted to administer assessments independently to participants.

Data Analyses

Identification of the low reading comprehension subgroup. Participants were originally recruited for the RCT using school-determined need of supplemental intervention in reading comprehension (a decision schools often based on failing the state reading proficiency exam the previous school year). Despite school designation as low-achieving, the initial sample of 233 in the RCT included students who scored within the average or above-average range on some of the pretest reading comprehension measures. Given our intent to focus the present analyses on struggling readers, it was necessary to limit the sample to those with low reading comprehension. In a previous study with the same sample (Clemens et al., 2017), we used latent class analysis (LCA) to empirically identify students with low reading comprehension. LCA uses multiple measures to identify classes (i.e., subgroups) of individuals that demonstrate similar skill profiles. The advantages of LCA include (a) the ability to use students' performance from multiple measures rather than single measures which may under-represent a complex construct like reading comprehension (Fletcher, 2006); and (b) use of an empirical approach to identify subgroups rather than relying on predefined cut scores.

We utilized the LCA-identified subgroups used in Clemens et al. (2017), but briefly describe the process here. Three reading comprehension measures were used in the LCA (GMRT, GRADE Comprehension, GORT-5 Comprehension). We omitted the TOSREC from the LCA given that researchers have questioned whether it is more a measure of silent reading fluency versus reading comprehension (Denton et al., 2011). However, because the TOSREC has been used as an index of comprehension in some studies (e.g., Ahmed et al., 2016), it was included in our subsequent dominance analyses to investigate the subskills that most strongly influence performance (and thereby inform future studies). The LCA included systematically testing solutions with increasing numbers of classes beginning with a two-class solution. At each step, we followed recommended statistical fit criteria (Asparouhov & Muthén, 2012; Nylund et al., 2007) and guidelines for evaluating whether the extracted classes were interpretable and theoretically plausible (Geiser, 2013). The LCA identified two subgroups with below-average scores on the reading comprehension measures, which represented the 180 students included in the present study.

Treatment of missing data. The mean percentage of cases with missing data on at least one measure was 7.07% (range = 0%–13.8%). All students had scores on at least one comprehension measure and at least one component skill measure. Data were missing due to student absences on testing days that could not be made up. Patterns of missingness, as examined with the MCAR (missing completely at random) test (Little, 1988), indicated data were missing completely at random. To yield more accurate and efficient parameter estimates in the dominance analysis, the missing cases were handled by conducting multiple imputations via chained equations (MICE) in R (van Buuren & Groothuis-Oudshoorn, 2011). Based on the average missingness of 7.07% we conducted the recommended number of imputations (Bodner, 2008; Graham et al., 2007; Royston & White, 2011) and current results were summarized from 20 imputed data sets.

Dominance analyses. We conducted dominance analyses (Budescu, 1993) using the SAS Macro developed by Azen and Budescu (2003) to determine the relative importance of each component skill for students' performance on each reading comprehension test. Dominance analyses involve a series of comparisons of models that use all possible combinations of predictors of a dependent measure. Each combination of predictors is referred to as a subset model. Via pair-wise comparisons of the separate subset models, dominance analysis reveals the unique contribution of each predictor compared with other predictors in the same model. Subsequent ranking of variables based on their importance across all possible models allows for determining whether one predictor "dominates" (i.e., has a larger additional contribution) over other predictors.

The importance of a predictor can be further interpreted by how consistently it dominates another predictor with a set of hierarchical terms (Azen & Budescu, 2003; Budescu & Azen, 2004). "Complete dominance" is the strongest level of dominance and occurs when the unique contribution of a predictor is always greater than another predictor in every possible subset model. "Conditional dominance" occurs when the average unique contribution of one predictor is greater than the unique contribution of another predictor across subset models with the same number of predictors. "General dominance" is the lowest level of dominance and is observed when the average unique contribution of one predictor across all the possible subset models is greater than that of another predictor, even though its unique contribution was not greatest in every subset model. Additional technical details on dominance analysis are provided in Azen and Budescu (2003) and Budescu and Azen (2004).

Results

Descriptive data and intercorrelations are reported in Table 1. Predictor contributions to the subset models across each reading comprehension test are reported in Table 2. Each coefficient in Table 2 can be interpreted as the average proportion of variance (i.e., R^2) accounted for by that specific predictor (see rows) that was unique and independent from

Table 1. Intercorrelations and Descriptive Statistics.

Variable		1	2	3	4	5	6	7	8	9	10	11	*M*	*SD*	Min.	Max.
1	Sight word efficiency	—											88.70	9.70	59	119
2	Oral reading fluency	.79**	—										123.93	30.09	13	212
3	Vocabulary	.16*	.11	—									89.40	9.08	55	111
4	Listening comprehension	.01	.03	.17*	—								10.87	2.67	1	17
5	Working memory	.18*	.09	.08	.11	—							88.47	12.16	26	114
6	GMRT reading comp	.28**	.33**	.22**	.14	.13	—						82.13	8.95	65	104
7	GRADE sentence comp	.17*	.27**	.34**	.18*	.11	.21**	—					8.54	3.17	1	17
8	GRADE passage comp	.05	.13	.13	.11	.13	.23**	.19*	—				13.62	4.30	2	25
9	GRADE comp total	.11	.19*	.29**	.14	.11	.28**	.67**	.82**	—			86.82	7.53	58	104
10	GORT-V reading comp	.14	.10	.21*	.21**	.03	.18**	.09	.002	.12	—		6.38	1.14	3	9
11	TOSREC	.42**	.38**	.28**	.02	.03	.17*	.14	−.04	.16*	.33*	—	86.11	9.44	54	112

Note. GMRT = *Gates-MacGinitie Reading Tests*; GRADE = *Group Reading and Diagnostic Evaluation*; Comp = comprehension; GORT-V = *Gray Oral Reading Test* (5th edition); TOSREC = *Test of Silent Reading Efficiency and Comprehension*. Standard scores reported for Sight Word Efficiency, Working Memory, GMRT, GRADE Total, and TOSREC; scaled score reported for GORT-V (which range from 1 to 20); raw scores reported for all other variables (standard scores were not available).
*$p < .05$. **$p < .01$.

the effect of other predictors in models of each size (see columns). For example, ORF alone accounted for 12.04% of the variance in GMRT, an average of 9.45% variance unique from the other predictors when included with one of the other four subskill variables, an average of 7.34% variance unique from the other predictors when included with every combination of two other predictors, and so on. Overall, as indicated in the last column of Table 2, ORF accounted for an average of 7.69% variance unique from the other predictors across all the subset models. A summary of the dominance analyses is reported in Table 3. In the following sections, we the describe the results for each comprehension measure.

Predicting GMRT Comprehension

As reported in Table 2, ORF made the largest average unique contribution in explaining GMRT Comprehension scores relative to the other predictors (7.69%). ORF demonstrated complete dominance over all other component skills. The second-most dominant predictor of GMRT performance was SWE, with an average of 4.08% explained variance, followed by vocabulary with an average of 3.54%. SWE demonstrated general dominance over vocabulary, listening comprehension, and working memory; while vocabulary demonstrated complete dominance over listening comprehension and working memory. Listening comprehension dominated working memory.

Predicting GRADE Sentence Comprehension

Vocabulary (10.21%) accounted for the largest average unique contribution in explaining GRADE Sentence Comprehension scores relative to other predictors (see Table 2) and demonstrated complete dominance over the other component skills (see Table 3). The second-most dominant predictor of GRADE Sentence Comprehension was ORF (6.07%), which demonstrated complete dominance over SWE, listening comprehension, and working memory. Listening comprehension (2.16%) demonstrated complete dominance over SWE and working memory, and SWE (1.43%) was conditionally dominant over working memory (.71%).

Predicting GRADE Passage Comprehension

ORF (2.06%) accounted for the largest average unique contribution in explaining GRADE Passage Comprehension scores relative to the other predictors (see Table 2). ORF was completely dominant over SWE, listening comprehension, and working memory; and demonstrated general dominance over vocabulary (see Table 3). Vocabulary was the second-most dominant predictor (1.53%) of the component skills, demonstrating complete dominance over SWE and listening comprehension, and general dominance over working memory. Working memory (1.34%) demonstrated complete dominance over SWE and listening comprehension, and listening comprehension (.80%) was generally dominant over SWE.

Predicting GRADE Total Comprehension

As previously noted, GRADE Total Comprehension is a composite score of the Sentence and Passage Comprehension subtests. Vocabulary (7.57%) accounted for the largest average unique contribution in explaining GRADE Total Comprehension scores relative to the other predictors, and was completely dominant over the other component skills.

Table 2. Average Unique Contributions of Each Predictor to Reading Comprehension Test Performance: All Subset Models and Overall Average.

Reading comp measure	Predictor	1 IV	2 IVs	3 IVs	4 IVs	5 IVs	*M*
GMRT reading comp	Oral reading fluency	.1204	.0945	.0734	.0557	.0405	.0769
	Sight word efficiency	.0865	.0592	.0370	.0185	.0027	.0408
	Vocabulary	.0520	.0400	.0323	.0275	.0249	.0354
	Listening comprehension	.0222	.0184	.0160	.0140	.0118	.0165
	Working memory	.0149	.0089	.0058	.0043	.0039	.0076
GRADE sentence comp	Vocabulary	.1233	.1082	.0983	.0919	.0886	.1021
	Oral reading fluency	.0782	.0646	.0566	.0525	.0516	.0607
	Listening comprehension	.0327	.0261	.0210	.0164	.0118	.0216
	Sight word efficiency	.0286	.0166	.0102	.0077	.0082	.0143
	Working memory	.0128	.0077	.0055	.0047	.0048	.0071
GRADE passage comp	Oral reading fluency	.0193	.0185	.0194	.0215	.0245	.0206
	Vocabulary	.0196	.0165	.0146	.0134	.0127	.0153
	Working memory	.0158	.0135	.0125	.0124	.0130	.0134
	Listening comprehension	.0111	.0093	.0079	.0065	.0051	.0080
	Sight word efficiency	.0024	.0026	.0044	.0074	.0114	.0057
GRADE reading comp total	Vocabulary	.0885	.0792	.0732	.0695	.0678	.0757
	Oral reading fluency	.0421	.0361	.0338	.0342	.0368	.0366
	Listening comprehension	.0200	.0158	.0124	.0093	.0063	.0128
	Working memory	.0139	.0099	.0081	.0075	.0078	.0095
	Sight word efficiency	.0111	.0066	.0058	.0077	.0117	.0086
GORT-V reading comp	Listening comprehension	.0431	.0390	.0356	.0327	.0299	.0361
	Vocabulary	.0463	.0398	.0347	.0304	.0268	.0356
	Sight word efficiency	.0192	.0146	.0113	.0090	.0073	.0123
	Oral reading fluency	.0096	.0062	.0038	.0021	.0008	.0045
	Working memory	.0025	.0011	.0005	.0003	.0004	.0010
TOSREC	Sight word efficiency	.1708	.1327	.0987	.0680	.0405	.1022
	Oral reading fluency	.1332	.0968	.0644	.0352	.0092	.0678
	Vocabulary	.0798	.0672	.0588	.0538	.0520	.0623
	Listening comprehension	.0019	.0017	.0019	.0019	.0018	.0019
	Working memory	.0021	.0011	.0011	.0015	.0019	.0015

Note. For each reading comprehension criterion measure, predictor variables are rank ordered according to their average unique contribution to predicting performance, from strongest to weakest, of performance on each respective reading comprehension test. Comp = comprehension; GMRT = *Gates-MacGinitie Reading Tests*; GRADE = *Group Reading and Diagnostic Evaluation*; GORT-V = *Gray Oral Reading Test* (5th edition); TOSREC = *Test of Silent Reading Efficiency and Comprehension*; IV = independent variable.

The second-most dominant predictor was ORF (3.66%), which was completely dominant over SWE, listening comprehension, and working memory. Listening comprehension (1.28%) was generally dominant over SWE and working memory, and working memory (.95%) was generally dominant over SWE.

Predicting GORT-5 Comprehension

Listening comprehension (3.61%) made the largest average unique contribution in explaining GORT-5 Comprehension scores. This predictor demonstrated complete dominance over ORF, SWE, and working memory, and general dominance over vocabulary. Vocabulary (3.56%) followed closely in terms of its relative importance and demonstrated complete dominance over ORF, SWE, and working memory. SWE (1.23%) was completely dominant over ORF and working memory, and ORF (.45%) was conditionally dominant over working memory.

Predicting TOSREC

SWE (10.22%) made the largest average unique contribution to TOSREC performance. SWE was completely

Table 3. Dominance Analysis Results for Each Subskill Predicting Performance on Each Reading Comprehension Measure.

Reading comp measure	Predictor	Complete dominance over	Conditional dominance over	General dominance over
GMRT reading comp	ORF	SWE, Vocab, Listen, WM	—	—
	SWE	—	—	Vocab, Listen, WM
	Vocab	Listen, WM	—	—
	Listen	WM	—	—
	WM	—	—	—
GRADE sentence comp	Vocab	ORF, SWE, Listen, WM	—	—
	ORF	SWE, Listen, WM	—	—
	Listen	SWE, WM	—	—
	SWE	—	WM	—
	WM	—	—	—
GRADE passage comp	ORF	SWE, Listen, WM	—	Vocab
	Vocab	SWE, Listen	—	WM
	WM	SWE, Listen	—	—
	Listen	—	—	SWE
	SWE	—	—	—
GRADE comp total	Vocab	ORF, SWE, Listen, WM	—	—
	ORF	SWE, Listen, WM	—	—
	Listen	—	—	SWE, WM
	WM	—	—	SWE
	SWE	—	—	—
GORT-V comprehension	Listen	ORF, SWE, WM	—	Vocab
	Vocab	ORF, SWE, WM	—	—
	SWE	ORF, WM	—	—
	ORF	—	WM	—
	WM	—	—	—
TOSREC	SWE	ORF, Listen, WM	—	Vocab
	ORF	Listen, WM	—	Vocab
	Vocab	Listen, WM	—	—
	Listen	—	—	WM
	WM	—	—	—

Note. For each reading comprehension measure, predictors are rank ordered from strongest to weakest in terms of their dominance over the other predictors. Comp = Comprehension; GMRT = *Gates-MacGinitie Reading Tests*; GRADE = *Group Reading and Diagnostic Evaluation*; GORT-V = *Gray Oral Reading Test* (5th edition); TOSREC = *Test of Silent Reading Efficiency and Comprehension*; ORF = oral reading fluency; SWE = sight word efficiency; Listen = listening comprehension; WM = working memory; Vocab = vocabulary.

dominant over ORF, listening comprehension, and working memory, and generally dominant over vocabulary. ORF and vocabulary were similarly important; ORF (6.78%) was completely dominant over listening comprehension and working memory, and generally dominant over vocabulary; whereas vocabulary (6.23%) demonstrated complete dominance over listening comprehension and working memory. Listening comprehension (.19%) was generally dominant over working memory.

Discussion

The skills that influence students' performance on standardized, commercially available tests of reading comprehension vary based on the format and characteristics of each test. The purpose of the present study was to further investigate and expand the literature on this issue, specifically with struggling readers in Grades 6 to 8. We used dominance analyses to determine the relative importance of component skills on tests of reading comprehension that varied in their reading demands, response formats, and other characteristics.

Dominant Predictors of Reading Comprehension Scores

Results revealed that component skills differentially contributed to test performance across the six reading comprehension measures, and that specific test characteristics

may dictate which component skill is most influential. In short, ORF was the most dominant predictor for reading comprehension tests that involved reading passages of text and answering multiple-choice questions (GMRT and GRADE Passage Comprehension). Listening comprehension was the most dominant predictor of GORT-5 Comprehension, which was the only test in which students read passages aloud and answered open-ended questions that were presented orally by the examiner. Vocabulary was the most dominant predictor of performance on GRADE Sentence Comprehension, a sentence-level cloze task in which students selected the missing word from a set of multiple-choice options. Fluency reading individual words (SWE) played its largest roles in predicting performance on tests with a time limit (GMRT and TOSREC). Working memory was least important relative to the other subskills. Following, we discuss the results in depth.

ORF tended to be the most dominant predictor overall. It was the most dominant predictor of GMRT and GRADE Passage Comprehension subtests, and was second-most dominant in predicting performance on the GRADE Sentence Comprehension subtest, GRADE Total Comprehension (as a function of its role in predicting performance on the two subtests that make up the composite), and the TOSREC. That ORF was the most dominant overall predictor should not come as a surprise given evidence of the robust relations of text-reading fluency to reading comprehension (Jenkins et al., 2003a, 2003b; Kim et al., 2012; Reschly et al., 2009; Shinn et al., 1992). Reading comprehension is impaired if words are not read accurately, and word-level automaticity allows finite cognitive resources to be devoted to higher-order reasoning and not expended by costly decoding processes (Klauda & Guthrie, 2008; C. A. Perfetti, 1985). Furthermore, this relationship is reciprocal, because text comprehension also influences reading fluency (Eason et al., 2013; Jenkins et al., 2003a, 2003b). In short, the importance of ORF to performance on the reading comprehension tests was likely due to the ways in which reading fluency is both indicative of, and influenced by, overall reading proficiency. Nevertheless, it is important that researchers, program developers, and clinicians consider this finding, especially when they might assume that students' performance on reading comprehension tests to be driven primarily by linguistic reasoning, inference-making, and other higher-order text processing skills. Our results diverge somewhat from those of Cutting and Scarborough (2006), who included a sample of students representing a range of reading skill levels in Grades 1 through 10. Although reading speed accounted for unique additional variance in the prediction of GMRT scores over decoding and oral language skills, Cutting and Scarborough found that oral language skills accounted for more overall unique variance than word-level reading skills, including reading speed. As we will discuss later, basic skills in word- and text-level reading remain a primary and defining feature of adolescents' reading difficulties, and it is possible that the broad age range and skill levels included in the Cutting and Scarborough study were a reason for the somewhat different outcomes we observed on the importance of ORF for GMRT performance.

ORF was not the most dominant predictor of *all* reading comprehension measures in this study. On the GORT-5, for example, listening comprehension was most important, followed by vocabulary. There are at least two possible reasons for this finding, and both relate to the specific characteristics of the GORT-5 Comprehension scale. First, it is important to reiterate that after students read the passages aloud, the passage is removed and students answer open-ended questions posed orally by the examiner. Unlike all of the other reading comprehension tests included in this study, students do not read the questions. Oral responding to verbal questions removes the challenges of having to read the questions and answer choices, thus it would make sense that linguistic comprehension skills, such as oral comprehension and vocabulary knowledge, would be critical to performance on the GORT-5. In a multiple regression analysis, Keenan et al. (2008) observed that decoding skills were less involved in performance on the previous 4th edition of the GORT-5 Comprehension scale than listening comprehension, a pattern that increased with older readers. The authors suggested the finding was due to the tendency for students to answer questions correctly on the 4th edition without having read the passages, as previously observed by Keenan and Betjeman (2006). Our study used a more recent version of the GORT (5th edition), in which items were revised by the publishers in efforts to make them passage-dependent (Wiederholt & Bryant, 2012). However, students' word- and text-reading skills remained less important than their listening comprehension, as in Keenan et al. (2008). If the test developers have indeed removed the possibility that students can answer questions correctly without having read the passages, it may mean that the question/answer format that does not involve reading may make language comprehension processes a primary driver of GORT-5 Comprehension scores.

A second possibility for the lower relative importance of word- and text-reading skills on the GORT-5 Comprehension scale is based on results of the meta-analysis by García and Cain (2014), who observed that word-reading skills were less predictive of reading comprehension on tests in which students read the comprehension passages aloud as opposed to silently. They suggested that reading comprehension tests that students read aloud often involve ceiling rules based on accuracy and/or fluency in reading the passages. In other words, like on the GORT-5, testing is stopped when a student's rate of correct words per minute falls below a predetermined criterion; as such, comprehension questions are limited to the passages in which text-reading was adequate.

This inherently limits the extent to which word- and text-reading skills would explain individual differences in reading comprehension. Regardless of the reason, the results underscore the tendency for reading comprehension test features to maximize or minimize the influence of component skills, like text-reading fluency, on test outcomes.

Vocabulary was the most dominant predictor of performance on the GRADE Sentence Comprehension subtest. Examining the format of this subtest offers clues to this finding. The subtest involves reading a sentence and selecting from a set of four one-word answer choices that best completes each sentence. Because students are only provided with a single sentence, they cannot rely on additional context clues that in a larger passage might aid in their overall comprehension, and they must be familiar with the meanings of the individual words in the answer choices. Therefore, scoring well is highly dependent on semantic understanding of a given set of words, which makes vocabulary knowledge the prime component.

Fluency with word recognition out of context, as measured by SWE in this study, played its largest role relative to the other subskills on measures that involved a time limit. Notably, SWE was the most dominant predictor of performance on the TOSREC. Considered a hybrid measure of reading fluency and comprehension (Cirino et al., 2013), we included it to inform subsequent work given the fact that it has been used in some studies as a measure of reading comprehension (e.g., Ahmed et al., 2016; Lonigan & Burgess, 2017; Petscher et al., 2017). In our sample of struggling adolescent readers, TOSREC performance was influenced the most by fluency reading words in list form (SWE) and in connected text (ORF). The brief 3-min time limit on the TOSREC prioritizes rapid word identification skills, therefore it makes sense that a measure like SWE in which performance is driven entirely by efficient word recognition would be the most influential of scores on the TOSREC. In addition, the TOSREC requires efficient sentence-level comprehension, in which word-reading accuracy and efficiency are critical because additional context is not available for supporting comprehension as it is in longer passages.

Another notable finding was that SWE was a more important predictor (relative to listening comprehension, vocabulary, and working memory) on the GMRT, another measure with a time limit. As observed by Kendeou et al. (2012), time-limited tests increase the importance of text-reading rate. Text-reading rate is primarily made possible by word-level automaticity, especially for struggling readers; Jenkins et al. (2003b) found that among students with reading difficulties, poor fluency in connected text was strongly linked to their difficulties with automaticity in decontextualized word reading (i.e., fluency reading words in list form). Although untimed measures are good indices of reading accuracy, automaticity in word reading captures an additional dimension that is important for better ascertaining word-reading proficiency (Protopapas et al., 2018). Difficulties with word-reading accuracy and rate are highly common among adolescents with reading difficulties (Brasseur-Hock et al., 2011; Cirino et al., 2013), an area of difficulty that is often overlooked by educators who perceive adolescents' reading difficulties as being driven primarily by insufficient skills in higher-order comprehension skills (Biancarosa & Snow, 2004; Meltzer et al., 2002). The present results build on previous studies that investigated the relative influence of word-reading skills on reading comprehension test performance and either used untimed measures of word reading, or included timed word reading as part of a composite variable (Best et al., 2008; Keenan et al., 2008; Nation & Snowling, 1997). We believe that the inclusion of a speeded measure of word reading is important for understanding the extent to which inefficient word recognition skills contribute to reading comprehension difficulties, which is particularly important for understanding the needs of adolescent struggling readers. It is also important to recognize that difficulties with word-level reading skills are probably magnified on comprehension tests with a time limit, such as the GMRT and the TOSREC, which is key for accurately interpreting a student's low scores on a test of reading comprehension.

Working memory was the least dominant predictor of performance on the reading comprehension measures. Involving simultaneous storage and manipulation of information, working memory is critical to many complex tasks, and evidence indicates that working memory is associated with reading comprehension (Peng et al., 2018). However, in the current study, the relations among working memory and the measures of reading comprehension were negligible. It is possible that the digit-span task, although popular as a working memory task, was insufficient for evaluating working memory, a possibility that will be discussed further below (Conway et al., 2005; Wells et al., 2018). It is also possible that working memory may have played a smaller role in test performance relative to the other measures because working memory is already involved in tests of oral reading, vocabulary knowledge, or listening comprehension. For instance, the GRADE listening comprehension subtest requires the student to listen to a 1 to 2 sentence passage and select a picture from four choices that best represents what they heard moments ago. Language comprehension itself involves working memory (Baddeley, 2003), and this subtest probably further implicates working memory by requiring the student to maintain information from the passage while processing the answer choices. Multiple-choice tests of vocabulary likely also involve working memory in similar ways. Therefore, it is unlikely that an isolated and discrete working memory task like digits-reversed will contribute uniquely to reading comprehension

test performance over other predictors that also involve working memory.

Contributions and Implications

Results of this study contribute to the research base in several ways. First, we included reading comprehension tests not previously used in similar studies (e.g., TOSREC and GRADE), as well as component skills that had been less-commonly examined in terms of their influence on reading comprehension test performance (e.g., oral reading fluency, working memory). Second, this study focused exclusively on struggling adolescent readers, which is a highly relevant population for these types of analyses given the extent to which tests may be used to inform high-stakes educational decisions or evaluations of intervention efficacy in research contexts. Thus, the interpretation of reading comprehension test results has arguably greater implications for students with reading difficulties compared with average- or above-average readers. The weak to modest correlations we observed among reading comprehension measures, which is consistent with previous studies with samples of broader range of learners (e.g., Cutting & Scarborough, 2006; Francis et al., 2006; Keenan et al., 2008, 2014), lends further support to the notion that tests of reading comprehension measure different skills—and we extend this evidence to low-achieving adolescents. Third, our use of dominance analysis is unique from previous studies, which have relied primarily on multiple regression analyses. Although dominance analysis is an extension of multiple regression, it allows for more definitive conclusions regarding the relative importance of individual predictors. Fourth, our results converge with previous studies of struggling adolescent readers (e.g., Brasseur-Hock et al., 2011; Cirino et al., 2013; Clemens et al., 2017) and underscore the possibility that adolescents' difficulties with basic reading and language skills may be primary factors driving adolescents' low reading comprehension, as opposed to higher-order text comprehension skills or lack of effective reading comprehension strategies, which middle-school educators may assume to be their primary sources of difficulty.

The findings also have several practical implications for educators, school leaders, researchers, and students with learning difficulties. Most notably, based on the characteristics of reading comprehension tests (e.g., passage length, item type, timing), the skills primarily implicated in test performance may be quite different from those an examiner expected. For researchers conducting intervention studies, who traditionally have not interpreted their results in terms of measurement-related issues, it is important to consider that the dependent measures they selected to evaluate the effects of a "reading comprehension" intervention may be influenced by a much different set of skills (e.g., word- or text-reading efficiency, vocabulary knowledge) than what was targeted in their intervention (e.g., main-idea generation). Clearly, different reading comprehension tests measure different things—and clinicians and researchers must be aware of these differences.

Second, the findings provide additional evidence that oral reading fluency is a robust index of overall reading proficiency across grade levels (Reschly et al., 2009). In the present study, ORF was the most dominant or second-most dominant predictor of comprehension test performance on all measures except for the GORT-5 (the one measure in which responding to questions did not involve reading). Our results converge with observations that difficulties reading words and text with efficiency are one of the most commonly observed skill deficits among struggling adolescent readers (Brasseur-Hock et al., 2011; Cirino et al., 2013; Clemens et al., 2017). The findings indicate that educators and clinicians should continue to consider oral reading fluency as an important part of assessment with struggling adolescent readers. In practice, this may mean including an assessment of reading fluency when evaluating students' reading skills, and when a student has a history of difficulties with word- or text-reading, perhaps periodically monitoring oral reading fluency. Inclusion of oral reading fluency in these contexts can indicate when arduous, error-prone reading may be at the core of a student's difficulties in understanding and learning from text. It should be remembered, however, that text-reading fluency is influenced by comprehension, because readers often slow down when encountering an unknown vocabulary term or when a portion of text does not make sense. Therefore, oral reading fluency assessment should be viewed as a source of information to consider alongside measures of vocabulary knowledge, background knowledge, and text comprehension to provide a complete picture of students' reading skills. In terms of instructional implications for these findings, the present results do not necessarily argue for reducing attention to reading comprehension, but rather emphasize the importance of continuing to target skills that improve students' ability to read text with accuracy and efficiency, which are critical for supporting text comprehension and important for students' academic success in subsequent grades.

Third, the results underscore the importance of carefully considering how reading comprehension test results are interpreted for adolescents with reading difficulties. Specifically, it should be considered that low scores on a test of reading comprehension do not necessarily mean a student has a problem with higher-order text processing. Once again, we point out the tendency for adolescents with reading difficulties to experience low word- and text-reading skills, inadequate vocabulary knowledge, or difficulties in both areas (Brasseur-Hock et al., 2011; Cirino et al., 2013; Clemens et al., 2017; Hock et al., 2009). Efficiency in connecting printed words to meanings with

little conscious effort is essential for allowing higher-order language comprehension processes to take place (C. Perfetti & Stafura, 2014; C. A. Perfetti, 1985). Nevertheless, as observed in our study, the relative importance of these skills varied across the tests of reading comprehension. A key implication is that for educators and researchers who work with struggling readers in the middle grades and beyond, results of reading comprehension tests should be interpreted with the knowledge that difficulties with basic reading and language skills are common and may be the primary reasons for why reading comprehension skills are weak. Instruction in or expectation of higher-order text processing would be of little use if those basic skills are inadequate.

Fourth, the timed nature of some tests should be considered. Consistent with Kendeou et al. (2012) and as evidenced by our results, tests with time limits may magnify the influence of reading fluency. Situations in which low reading fluency is suspected to be a primary factor in a students' reading comprehension test performance might call for "testing the limits," which in educational or psychological testing refers to allowing an examinee to continue responding past a standardized time limit or ceiling rule. When testing the limits, a student's responses up to the limit are recorded to allow for determining standard scores and percentiles. Observation of the student's responses beyond the time limit or ceiling can inform clinical impressions regarding the student's test performance and sources of difficulty. Testing the limits with a timed test of reading comprehension may allow the examiner to ascertain the extent to which low reading fluency may explain low test scores.

Concerns regarding the variability of skills measured by tests of reading comprehension may be reduced by administering multiple measures that vary in format and characteristics, which allows users to examine convergence in scores across tests and reduces the possibility of under-representing the construct (Fletcher, 2006). We acknowledge that the administration of multiple measures of reading comprehension is not always feasible, or even necessary in some situations. The use of multiple measures might be more important in situations involving high-stakes decisions, such as evaluation for special education eligibility or determination of educational services and placement. In addition, in research studies in which reading comprehension is an important independent or dependent variable, multiple measures reduce mono-method or mono-operation bias, and affords the use of latent variables (Fletcher, 2006). On the contrary, in situations involving lower-stakes decisions such as routine formative assessment to inform ongoing adjustments to instruction, a single measure of reading comprehension administered periodically may be sufficient, provided the user is confident the test measures skills of primary interest. A key point for educators, clinicians, and researchers to recognize is that single measures of reading comprehension provide a narrow view of the construct, and test performance may be driven primarily by skills not initially perceived as important on measures marketed as tests of "reading comprehension."

Finally, the accumulated body of evidence that reading comprehension tests measure different skills has implications for test developers. We believe it is important for test developers to examine the relative importance of component reading and language skills for performance on their tests. These analyses should include a range of achievement levels with disaggregation of results for students with low reading achievement. Information like this can improve test development and can allow users to make more informed decisions on test selection.

Limitations

Several limitations of this study should be acknowledged. First, although we included four different commercially available tests of reading comprehension, our list was not exhaustive of the different ways in which reading comprehension can be assessed. For example, retell is common form of assessment (Reed & Vaughn, 2012), and we also did not include the Passage Comprehension subtest from the Woodcock-Johnson family of assessments (an open-ended cloze technique within short passages), which is one of the most frequently used measures in experimental and correlational studies of reading comprehension. Future dominance analyses might investigate what reading subskills are most dominant in predicting performance on other test formats among struggling adolescent readers.

There were limitations to our battery of component skills. The GRADE Vocabulary subtest is partially influenced by word reading because students must read each word in the multiple-choice format as they take the test independently. We did not include a measure of inference-making, a key aspect in reading comprehension (Barth et al., 2015). We also did not include measures of sustained attention or self-regulation; the duration of the reading comprehension tests in our battery ranged from 3 to 45 min and evidence indicates that longer tests may magnify the role of students' attention to task (Collins et al., 2019; DiCerbo et al., 2004; Vavassoeur, 2016). We used a digits-reversed working memory task, which is common in research and practice (Kasper et al., 2012), however, evidence suggests that digit-span tasks may measure short-term memory more than working memory because they do not introduce new stimuli within the primary storage set such as in reading span tasks (Conway et al., 2005; Wells et al., 2018). Nevertheless, our findings are consistent with those of Tighe and Schatschneider (2014), who observed

that working memory measured by reading and listening span tasks was the least predictive of reading comprehension across grade levels compared with other reading, language, and reasoning skills.

Conclusion

Results of this study stress the need for caution when interpreting tests of reading comprehension for struggling adolescent readers. Our results converge with previous findings on the differential influence of component skills and knowledge variables for performance on commercially available tests of reading comprehension. Our study extends this work to a sample of struggling adolescent readers, a population for whom results on tests of reading comprehension can have significant educational and research implications. In short, educators, clinicians, and researchers should be aware that reading comprehension test performance may be driven more by passage length, how questions are asked, student response formats, or whether the measure is timed, rather than comprehension of the text itself. These issues are relevant for evaluating individual student achievement, system-wide outcomes, and the effects of interventions designed to improve reading comprehension. Struggling readers are of primary interest in all three situations.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Preparation of this article was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305F100013 to the University of Texas at Austin as part of the Reading for Understanding Research Initiative. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

ORCID iDs

Nathan H. Clemens https://orcid.org/0000-0002-8361-1303

Amanda Martinez-Lincoln https://orcid.org/0000-0002-6271-0353

Jessica Toste https://orcid.org/0000-0002-6327-0054

Note

1. A common passage was administered as part of the intervention trial to more clearly ascertain intervention effects. The Grade 7 passage had a Lexile score (an index of text difficulty and complexity) of 960, which fell within the Lexile range observed on the EasyCBM passages across Grades 6 to 8 (600–1180).

References

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., & Kulesz, P. (2016). Validation of the direct and inferential mediation (DIME) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology*, *44*, 68–82.

Andreassen, R., & Bråten, I. (2010). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading*, *33*(3), 263–283.

Asparouhov, T., & Muthén, B. (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes. *Mplus Web Notes*, *14*, 22–32.

Azen, R., & Budescu, D. V. (2003). The dominance analysis approach for comparing predictors in multiple regression. *Psychological Methods*, *8*, 129–148.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Barth, A. E., Barnes, M., Francis, D., Vaughn, S., & York, M. (2015). Inferential processing among adequate and struggling adolescent comprehenders and relations to reading comprehension. *Reading and Writing*, *28*(5), 587–609.

Best, R. M., Floyd, R. G., & Mcnamara, D. S. (2008). Differential competencies contributing to children's comprehension of narrative and expository texts. *Reading Psychology*, *29*(2), 137–164.

Biancarosa, G., & Snow, C. E. (2004). *Reading next: A vision for action and research in middle and high school literacy: A report from Carnegie Corporation of New York*. Alliance for Excellent Education.

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling: A Multidisciplinary Journal*, *15*, 651–675.

Brasseur-Hock, I. F., Hock, M. F., Kieffer, M. J., Biancarosa, G., & Deshler, D. D. (2011). Adolescent struggling readers in urban schools: Results of a latent class analysis. *Learning and Individual Differences*, *21*(4), 438–452.

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, *114*, 542–551.

Budescu, D. V., & Azen, R. (2004). Beyond global measures of relative importance: Some insights from dominance analysis. *Organizational Research Methods*, *7*(3), 341–350.

Cain, K., & Oakhill, J. (2009). Reading comprehension development from 8 to 14 years. The contribution of component skills and processes. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 143–175). Guilford Press.

Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing*, *26*, 1059–1086.

Clemens, N. H., Hsiao, Y. Y., Simmons, L. E., Kwok, O. M., Greene, E. A., Soohoo, M. M., . . . Otaiba, S. A. (2019). Predictive validity of kindergarten progress monitoring measures across the school year: Application of dominance analysis. *Assessment for Effective Intervention*, *44*(4), 241–255.

Clemens, N. H., Simmons, D., Simmons, L. E., Wang, H., & Kwok, O. M. (2017). The prevalence of reading fluency and vocabulary difficulties among adolescents struggling with reading comprehension. *Journal of Psychoeducational Assessment*, *35*(8), 785–798.

Collins, A. A., Compton, D. L., Lindström, E. R., & Gilbert, J. K. (2019). Performance variations across reading comprehension assessments: Examining the unique contributions of text, activity, and reader. *Reading and Writing*, *33*, 1–30.

Collins, A. A., Lindström, E. R., & Compton, D. L. (2018). Comparing students with and without reading difficulties on reading comprehension assessments: A meta-analysis. *Journal of Learning Disabilities*, *51*, 108–123.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769–786.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, *99*(2), 311–325.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*(3), 277–299.

Denton, C. A., Barth, A. E., Fletcher, J. M., Wexler, J., Vaughn, S., Cirino, P. T., . . . Francis, D. J. (2011). The relations among oral and silent reading fluency and comprehension in middle school: Implications for identification and instruction of students with reading difficulties. *Scientific Studies of Reading*, *15*, 109–135.

DiCerbo, K. E., Oliver, J., Albers, C., & Blanchard, J. (2004). Effects of reducing attentional demands on performance of reading comprehension tests by third graders. *Perceptual and Motor Skills*, *98*(2), 561–574.

Eason, S. H., Sabatini, J., Goldberg, L., Bruce, K., & Cutting, L. E. (2013). Examining the relationship between word reading efficiency and oral reading rate in predicting comprehension among different types of readers. *Scientific Studies of Reading*, *17*(3), 199–223.

Fletcher, J. M. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, *10*(3), 323–330.

Fogarty, M., Clemens, N. H., Simmons, D., Simmons, L., Anderson, L., Oslund, E., Davis, J., & Smith, A. (2017). The impact of a technology-mediated intervention on adolescents' reading comprehension. *Journal of Research in Educational Effectiveness*, *10*, 326–353.

Francis, D. J., Snow, C. E., August, D., Carlson, C. D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension. *Scientific Studies of Reading*, *10*(3), 301–322.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children*, *71*(1), 7–21.

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, *84*(1), 74–111.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177–190.

Geiser, C. (2013). *Data analysis with Mplus*. Gilford Press.

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

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*, 206–213.

Hock, M. F., Brasseur, I. F., Deshler, D. D., Catts, H. W., Marquis, J. G., Mark, C. A., & Stribling, J. W. (2009). What is the reading component skill profile of adolescent struggling readers in urban schools? *Learning Disability Quarterly*, *32*(1), 21–38.

Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003a). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*(4), 719–729.

Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003b). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice*, *18*(4), 237–245.

Kamhi, A. G., & Catts, H. W. (2017). Epilogue: Reading comprehension is not a single ability—Implications for assessment and instruction. *Language, Speech, and Hearing Services in Schools*, *48*(2), 104–107.

Kasper, L. J., Alderson, R. M., & Hudec, K. L. (2012). Moderators of working memory deficits in children with attention-deficit/hyperactivity disorder (ADHD): A meta-analytic review. *Clinical Psychology Review*, *32*, 605–617.

Keenan, J. M., & Betjemann, R. S. (2006). Comprehending the gray oral reading test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies of Reading*, *10*(4), 363–380.

Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, *12*(3), 281–300.

Keenan, J. M., & Meenan, C. E. (2014). Test differences in diagnosing reading comprehension deficits. *Journal of Learning Disabilities*, *47*(2), 125–135.

Kendeou, P., Papadopoulos, T. C., & Spanoudis, G. (2012). Processing demands of reading comprehension tests in young readers. *Learning and Instruction*, *22*(5), 354–367.

Kim, Y. S., Wagner, R. K., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from Grade 1 to Grade 2. *Journal of Experimental Child Psychology*, *113*, 93–111.

Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology*, *100*, 310–321.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198–1202.

Lonigan, C. J., & Burgess, S. R. (2017). Dimensionality of reading skills with elementary-school-age children. *Scientific Studies of Reading*, *21*, 239–253.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2002). *Gates-MacGinitie Reading Tests–4th Edition: Technical manual, forms S & T*. Riverside.

Maria, K., & Hughes, K. E. (2008). *Gates-MacGinitie reading tests: Technical report supplement*. The Riverside Publishing Company.

Mellard, D. F., Anthony, J. L., & Woods, K. L. (2012). Understanding oral reading fluency among adults with low literacy: Dominance analysis of contributing component skills. *Reading and Writing*, *25*(6), 1345–1364.

Meltzer, J., Cook Smith, N., & Clark, H. (2002). *Adolescent literacy resources: Linking research and practice*. LAB, Brown University.

Miller, A. C., Davis, N., Gilbert, J. K., Cho, S.-J., Toste, J. R., Street, J., & Cutting, L. E. (2014). Novel approaches to examine passage, student, and question effects on reading comprehension. *Learning Disabilities Research & Practice*, *29*(1), 25–35.

Nation, K., & Snowling, M. (1997). Assessing reading difficulties: The validity and utility of current measures of reading skill. *British Journal of Educational Psychology*, *67*(3), 359–370.

National Assessment of Educational Progress. (2019). *The nation's report card: 2019 Reading national achievement level results*. https://nces.ed.gov/nationsreportcard/

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*, 535–569.

Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices—Past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 31–88). Routledge.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., . . . Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, *144*(1), 48–62.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*(1), 22–37.

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

Petscher, Y., Al Otaiba, S., Wanzek, J., Rivas, B., & Jones, F. (2017). The relation between global and specific mindset with reading outcomes for elementary school students. *Scientific Studies of Reading*, *21*, 376–391.

Protopapas, A., Katopodi, K., Altani, A., & Georgiou, G. K. (2018). Word reading fluency as a serial naming task. *Scientific Studies of Reading*, *22*(3), 248–263.

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

Reed, D. K., & Vaughn, S. (2012). Retell as an indicator of reading comprehension. *Scientific Studies of Reading*, *16*(3), 187–217.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology*, *47*, 427–469.

Royston, P., & White, I. R. (2011). Multiple imputation by chained equations (MICE): Implementation in Stata. *Journal of Statistical Software*, *45*(4), 1–20.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology*, *96*(2), 265–280.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. I. L. (1992). Oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review*, *21*(3), 459–479.

Spear-Swerling, L. (2004). Fourth graders performance on a state-mandated assessment involving two different measures of reading comprehension. *Reading Psychology*, *25*(2), 121–148.

Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., & Hairrell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly*, *32*(3), 199–222.

Tighe, E., & Schatschneider, C. (2014). A dominance analysis approach to determining predictor importance in third, seventh, and tenth grade reading comprehension skills. *Reading and Writing*, *27*, 101–127.

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency—Second edition (TOWRE-2)*. Pearson Clinical Assessment.

University of Oregon. (2008). *Passage reading fluency* (Assessment measures).

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67.

Vavassoeur, L. C. (2016). *Predictive Power of the Test of Everyday Attention for Children (TEA-CH) on Various Methods of Reading Comprehension Assessment among Low-Income Fourth Grade Children of Color* [Unpublished doctoral dissertation]. ProQuest LLC, Howard University.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2010). *TOSREC: Test of sentence reading efficiency and comprehension*. Pro-Ed.

Wells, E. L., Kofler, M. J., Soto, E. F., Schaefer, H. S., & Sarver, D. E. (2018). Assessing working memory in children with ADHD: Minor administration and scoring changes may improve digit span backward's construct validity. *Research in Developmental Disabilities*, *72*, 166–178.

Wiederholt, J. W., & Bryant, B. R. (2012). *Gray oral reading tests–5th edition*. Pro-Ed.

Williams, K. T. (2001). *Group reading assessment and diagnostic evaluation*. American Guidance Service.

Wixson, K. K. (2017). An interactive view of reading comprehension: Implications for assessment. *Language, Speech, and Hearing Services in Schools*, *48*(2), 77–83.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2006). *Woodcock-Johnson® III NU tests of achievement*. Riverside.