Literacy Practice and Research

Volume 47 | Number 4 Article 2

2022

Predictive Values of Factors Affecting Reading Comprehension Assessment

Taylor M. Grantham
Valdosta State University, taylormstruble@gmail.com

Matthew Carter
Valdosta State University, mdcarter@valdosta.edu

Crystal C. Randolph
Valdosta State University, crandolph@valdosta.edu

Gina M. Doepker
The University of Texas at Tyler, gdoepker@uttyler.edu

Predictive Values of Factors Affecting Reading Comprehension Assessment

Cover Page Footnote

Taylor Grantham is a graduate of Valdosta State University in Valdosta, Georgia and a speech-language pathologist in Sweetwater, Texas, United States; email: taylormstruble@gmail.com Matthew D. Carter is an associate professor at Valdosta State University, Valdosta, Georgia, United States; email: mdcarter@valdosta.edu Crystal Randolph is an associate professor at Valdosta State University, Valdosta, Georgia, United States; email: crandolph@valdosta.edu Gina Doepker is an associate professor at The University of Texas at Tyler, Tyler, Texas, United States; email: gdoepker@uttyler.edu

Predictive Values of Factors Affecting Reading Comprehension Assessment

Taylor Grantham is a graduate of Valdosta State University, Valdosta, Georgia, United States and a speech-language pathologist in Sweetwater, Texas, United States; email: taylormstruble@gmail.com

Matthew D. Carter is an associate professor at Valdosta State University, Valdosta, Georgia, United States; email: mdcarter@valdosta.edu

Crystal Randolph is an associate professor at Valdosta State University, Valdosta, Georgia, United States; email: crandolph@valdosta.edu

Gina Doepker is an associate professor at The University of Texas at Tyler, Tyler, Texas, United States; email: gdoepker@uttyler.edu

Children across the United States continue to exhibit difficulty with reading comprehension skills. As reported by Spencer and Wagner (2018), approximately one-third of fourth and eighth grade students achieve reading comprehension scores at a proficient level. Furthermore, more than 50% of children classified as late-emerging poor readers have reading comprehension difficulties (Catts *et al*., 2012). Reading comprehension difficulties extend beyond the academic arena and may affect activities of daily living such as interpreting manuals, contracts, applications, and other documents needed for an occupation and home-based operations. Therefore, continually advancing the knowledge-base related to reading comprehension assessment and intervention is necessary. Selecting reliable assessment tools and interpreting their results in their proper context has major implications for the differential diagnosis and treatment of reading impairments that may result in comprehension difficulties.

Reading Comprehension

Reading comprehension has been defined as "the process of simultaneously constructing and extracting meaning through interaction and engagement with print" (Research and Development Reading Study Group, 2002, p. 11). It is an additive process, which relies on the successful integration of both textual and extra-textual information (e.g., background knowledge and experience) within the bottleneck of a limited capacity working memory (Britton & Graesser, 1996). Reading comprehension is a multifaceted, dynamic, and interactive process that involves characteristics of the reader and the text and that is situated within a larger sociocultural context that interacts with the aforementioned characteristics. The underlying processes of reading comprehension create a challenging feat to those who create, publish, and interpret reading comprehension assessments.

Simple View of Reading

Word reading, fluency, and reading mode (i.e., silent, and oral reading) are a few of the underlying processes that affect reading comprehension. The bottom-up simple view of reading proposed by Gough and Tunmer (1986) categorizes reading into two main components: decoding and linguistic comprehension. Although this view of reading is referred to as the "simple" model, the involvement of multiple subprocesses (e.g., alphabetic knowledge, cognition, and level of language ability) complicates the cognitive construct known as comprehension (Kamhi, 2009). As a result, comprehension is rendered exceedingly difficult to assess. Although the relationship between decoding and comprehension is evident, multiple studies have indicated that this relationship varies as a function of reader proficiency and reader age (Cain, Oakhill, & Bryant, 2004). Cain *et al*. found that the relationship between word reading ability and comprehension was more evident in children ranging from 8 to 9 years old than children ranging from 10 to 11 years old indicating the latter relied less heavily on word accuracy to understand the text.

Reading Fluency

Reading fluency can be viewed as reading accurately, at an appropriate rate and with the appropriate prosody or expression. Fluent reading is mainly measured via reading accuracy and reading rate. As reading skills advance, it is expected that reading rate will increase as expression becomes smoother and more effortless. Reading that is at an appropriate speed, accurate, and with proper expression (fluent) is thought to facilitate comprehension. However, fluent reading does not ensure comprehension, but instead, theoretically allows cognitive resources to be devoted toward the act of comprehending the material that is being read (LaBerge & Samuels, 1974).

Each aspect of reading fluency influences textual comprehension (Jenkins et al., 2003). For instance, words read incorrectly can lead to a setback in the interpretation of the word and comprehension of the overall text (Hudson, Lane, & Pullen, 2005). Furthermore, slow, inaccurate reading is suggestive of dysfluent reading (Breznitz, 2006) and is often indicative of problems in reading comprehension (Başaran, 2013) resulting in constraints in the mental capacity needed to comprehend text (Hudson *et al*., 2005). Once more, the multifaceted nature of a reading comprehension construct (i.e., reading fluency) increases the complexity of measuring and interpreting assessment scores. Findings from a previous study suggested that textual reading fluency predicts reading comprehension ability more so than single-word reading fluency and context accuracy although there was much overlap (Jenkins *et al*., 2003).

Reading Modes

The comprehension of what is read can be achieved through oral and silent reading which adds yet another layer to the dynamics regarding assessment of the underlying comprehension processes. Oral and silent reading each entails multiple underlying processes that work in tandem to allow for meaning extraction. Oral reading relies on two senses and requires the child to see and hear the word whereas silent reading relies solely on seeing the word. As a result, oral reading is thought to be superior to silent and listening comprehension (Elgart, 1978) although it is also suggested that silent reading fluency and silent reading comprehension progresses from oral reading fluency and comprehension (Kim Wagner, & Foster, 2011). Oral reading comprehension has been demonstrated to be highly correlated with word decoding abilities. As children transition toward reading to learn rather than learning to read, they begin to increase silent reading. Silent reading comprehension becomes the preferred method for skilled readers because it usually is a faster method of reading in comparison to oral reading (Kim *et al*., 2011). Furthermore, since oral reading comprehension requires significantly more mental capacity than silent reading comprehension (Vorstius, Radach, & Lonigan, 2015), the load-bearing variables of reading comprehension are likely to vary between the aforementioned reading modes.

Reading Comprehension Assessments

Multiple assessments purport to measure reading comprehension. However, the various underlying constructs of assessments that measure this skill may result in different comprehension scores and interpretations by the examiner. A multitude of factors has been proposed, as discussed above, to affect the results that are obtained on reading comprehension tests (Kamhi, 2012). To this end, there is no universally accepted form of reading comprehension assessment. Generally, assessment tools require the individual to read a text and then answer subsequent questions regarding that text. Tools vary considerably in terms of how much text is to be read (sentences to longer passages), whether texts are read aloud or silently, whether factual recall or inferential analysis is emphasized, and which question type (e.g., multiple choice, cloze, true/false, open-ended) is presented. These varied practices further implicate the need for predicting the variance accounted for by the underlying processes of reading comprehension.

The current study extends and updates findings from Cutting and Scarborough's (2006) research which revealed that underlying reading processes (i.e., decoding, linguistic comprehension, and reading speed) significantly and variably contributed to the prediction of reading comprehension. However, many of the measurements that Cutting and Scarborough used have been demonstrated to be severely flawed in their designs and thus, several of the assessments that were utilized have been drastically revised and updated to address these concerns. The investigators for this study utilized current versions of the assessments to examine the predictive relationship of single-word reading accuracy, single-word reading fluency, and textual reading fluency in oral and silent reading comprehension. By examining these relationships, two goals will be achieved 1) researchers and practitioners will have updated information about the predictive values and relationships of underlying reading processes in relationship to reading comprehension and 2) practitioners will be able to acquire discernment related to the various underlying processes that contribute to reading comprehension.

Theoretical Perspective

Given the complexity and dynamic nature of the underlying processes of reading comprehension, assessing this skill presents with several challenges. First, reading assessments that measure narrowly-defined constructs only account for a solitary aspect of reading comprehension. Second, accounting for the multiple underlying constructs in absence of explaining the predictive value of each could further complicate interpretation of assessment scores. Lastly, varied elicitation styles of stimuli could further muddle the predictive relationship between underlying constructs and reading comprehension; thus, intensifying the confusion in interpreting assessment scores and definitively identifying the impact of the underlying construct being measured. Despite the above concerns and the multi-factorial nature of comprehension, test developers and publishers continue to publish assessment tools that are designed to assess “comprehension” as if it were a singular process.

Assessing the Underlying Constructs of Reading Comprehension

Single Word Reading

Dual-route theories of single word reading propose that word reading can be accomplished by accessing the word’s phonological representation (phonological decoding) or its visual representation (sight-word reading) (Besner & Smith, 1992; Pritchard, Coltheart, Marinus, & Castles, 2018). Tests of single-word reading commonly differentiate between phonological decoding and sight word reading. Subsequently, two distinct tasks have been developed to measure single word reading (Carter, Walker, & O’Brien, 2015). In order to assess decoding ability, nonsense words are commonly used as stimuli since these words are not likely to exist in the reader’s sight-word vocabulary. This task is commonly referred to as word attack. Its counterpart is word identification, which is designed to assess sight word reading abilities. Word identification uses real words as stimuli with a large proportion of those words being considered non-phonetic or exception words such as *yacht*. The logic behind this method is that readers should not be able to rely solely upon phonetic rules to accurately decode non-phonetic words.

Reading Fluency

Reading fluency is frequently assessed in two manners: with single word lists or with connected text. Single-word reading fluency tests require the reader to read lists of words quickly and accurately for a certain amount of time. On the other hand, textual fluency assessment consists of obtaining measures of reading rate and accuracy while reading connected text. Textual reading fluency tasks lend themselves more naturally toward measuring the reader’s prosody although the extra-linguistic context of the text can potentially affect fluency rates. In contrast, single-word reading fluency tasks provide more focal information regarding word-reading abilities since context is not present in a list and therefore, meaning cannot be utilized to scaffold decoding (Torgesen, Wagner, & Rashotte, 2012).

Issues Affecting the Assessment of Reading Comprehension

The underlying constructs measured by reading assessments account for differing predictive values, which makes it difficult to discern which reading skill is accounted for (Nation & Snowling, 1997). Therefore, it is critical to understand the nature of a specific reading comprehension test in order to ascertain what specific reading skills are actually being emphasized (Keenan & Betjemann, 2006). Research does not support the notion of comprehension as a singular construct. Instead, it has been found that slight manipulations in the way comprehension is assessed can drastically affect comprehension assessment results. Assessments utilizing expository texts have been demonstrated to be more reliant upon prior knowledge whereas narrative texts have been shown to potentially reduce this impact (Wolfe & Woodwyk, 2010). If the text is within view when questions are answered, memory demands are greatly reduced (Johnston, 1984). Cloze format questions are highly associated with word recognition skills and multiple-choice formats can be more susceptible to passage dependency effects (Fletcher, 2006; Keenan, Betjemann, & Olson, 2008). Longer text lengths increase the demands of the working memory system (Andreassen & Bråten, 2010) whereas shorter texts are more reliant upon decoding ability (Carter, Walker, O'Brien, & Hough, 2017). This signifies that there are other factors that both hinder and facilitate reading comprehension. Each assessment tool places different weights on these various skills to approximate their view of the construct of comprehension. As such, there is great variation in how comprehension is measured and the results that are obtained from these measures.

Cutting and Scarborough's Original Study

Cutting and Scarborough (2006) investigated the relationships among several assessment tools relative to reading comprehension scores. Their study investigated the contribution of decoding, linguistic comprehension, phonological awareness, vocabulary, reading speed, verbal memory, and reasoning skills to reading comprehension as measured by different instruments. The authors analyzed the data from a sample of 97 children (32 females). The grade level ranged from Grades 1.5 to 10.8, and the age range was from 7.0 years to 15.9 years. Data consisted of the results from various oral language, written language, and cognitive assessments. A summary of these assessment tools is presented in Table I.

Cutting and Scarborough (2006) conducted a series of regression and correlation analyses and found that the comprehension measures obtained on the *G-M* correlated highly with those from the *WIAT*, but the same could not be said for the *GORT-3*. In addition, phonological decoding/word recognition skills contributed to the prediction of reading comprehension regardless of which comprehension measure was entered in the regression models, although the relative contributions varied significantly (Cutting & Scarborough, 2006). However, it should be noted that the measure of word reading abilities that the authors used was a combination of two different measures, the Basic Reading subtest of the *WIAT* and the Word Attack subtest from the *WJPB-R.* Although Cutting and Scarborough utilized what is believed to be a measure of phonological decoding (Word Attack from *WIAT*), the content validity of the sight word reading measure they utilized is uncertain. The Basic Reading subtest of the *WIAT* requires the individual to "identify sounds to decode words" (Cohen, 1993). This description does not

Table I
Summary of the relevant assessment tools used in Cutting and Scarborough, 2006.

Assessment Tool	Area(s) Assessed	Assessment Method Summary	Comparable Measure in Current Study
Gates-MacGinitie Reading Test – Revised	Silent reading comprehension	Expository and narrative texts are read silently and multiple-choice questions answered while the text is still in view.	Gray Silent Reading Tests uses multiple choice questions while the text is still in view.
Wechsler Individual Achievement Test	Single-word reading Silent Reading	Single words are read aloud and scored for accuracy Expository and narrative passages are read silently and two open-ended questions are answered while the text is in view.	Test of Word Reading Efficiency – 2nd Ed. used to obtain single word reading fluency scores relative to sight-word and phonological decoding processes. This allows for individual scores as well as a normed composite score to be combined from a singular assessment tool using common normative data.

Gray Oral Reading Tests – 3rd Ed.	Oral reading comprehension Oral reading rate	Expository and narrative texts are read aloud while being timed and multiple-choice questions are answered after the text is removed from view.	Gray Oral Reading Tests – 5th Ed. has addressed major issues regarding reliability that were revealed in the 3rd and 4th editions.
Woodcock Johnson Psychoeducational Battery – Revised	Phonological decoding	Nonsense words are read aloud and scored for accuracy.	In addition to previously mentioned Test of Word Reading Efficiency – 2nd Ed., the Word Identification and Word Attack subtests of the Woodcock Reading Mastery Tests – 3rd Ed. was administered to provide an untimed measure of single word reading abilities.

necessarily entail the common format associated with word identification subtests which could affect content validity. In addition, there is cause for concern when forming a composite score from two unrelated tools that were normed on completely different samples. In summary, the results from Cutting and Scarborough (2006) indicate that each measure of reading comprehension assesses the construct of comprehension in drastically different manners. The authors concluded that different tests provide discrepant information regarding the component skills that could be targets for remediation.

Current Study Rationale/Purpose

Cutting and Scarborough's results provide appropriate applications to current practice, but there are limits upon how generalizable their results remain. Perhaps the most concerning limitation is the datedness of the test protocols that were utilized. The *GORT-3* has since undergone two additional revisions. At the time of Cutting and Scarborough's (2006) research, the norms which the *GORT-3* relied upon were approximately 24 years old (Wiederholt & Bryant, 2012). The *WIAT* has also undergone two additional revisions with the third edition being published in 2009. The authors were using the most recent 4th edition of the *GM* although normative updates have been provided for that version. Finally, multiple tests bearing the name Woodcock-Johnson have been produced and updated since the 1989 publishing of the *WJPB-R*.

In addition to the issues associated with the datedness of the measures used by Cutting and Scarborough, the format of each measure has drastically changed since 2006. This greatly reduces the readers' ability to apply the study's specific findings to the current versions of the measures. For example, the current *GORT-V* no longer applies basal or ceiling rules based upon comprehension performance. In the 3rd edition, the *GORT* used procedures to identify basals and ceilings for both fluency (accuracy and rate) and comprehension. The current edition obtains measures of comprehension but the test procedures are not directed by examinee's comprehension performance. In addition, a lack of passage dependency has been reported on the texts used by the various editions of the *GORT* (Keenan & Betjemann, 2006). Although the *GORT-V* has maintained the use of the questionable passages, it has changed the format and wording of the questions in order to reduce the likelihood that the questions can be answered without reading the passage. The *GORT-3* utilized multiple choice questions whereas open ended questions are currently utilized. In total, the *GORT-V* is a rather different test than was the *GORT-3*. Questions remain as to whether the current versions of these tests are reliant upon the same subprocesses as were the previous versions.

Many assessment tools claim to measure the same generic concept known as comprehension but significant dissimilarities exist between the results that are obtained. Much of this depends on the format each assessment uses and the subprocesses that are subsequently emphasized. This renders confusion in interpreting the results in order to develop treatment plans. The examiner must understand what principle components each comprehension test is measuring in order to provide appropriate, focused evidence-based interventions. The purpose of this study was to investigate the predictive value of single-word reading accuracy, single-word reading fluency, and textual reading fluency to oral and silent reading comprehension.

The current study seeks to investigate the following experimental question: What is the predictive value of single-word reading accuracy, single-word reading fluency, textual reading fluency in relationship to oral and silent reading comprehension. It is hypothesized that single-word reading accuracy would have the highest predictive value for oral reading comprehension. This prediction is based on the notion that young readers depend on lower-level, established skills, such as decoding, to develop higher cognitive skills, such as reading comprehension. In regard to silent reading comprehension, it is hypothesized that textual reading fluency will most be associated with comprehension measures. Reading fluency is a hallmark feature of appropriate reading development and the ability to read silently is a skill that develops with age as well. Therefore, it is hypothesized that those who are more developed in regard to fluency will be more prepared for silent reading tasks.

Methodology

Participants

This study including child assent and parental consent protocols was approved by the BLINDED University Institutional Review Board prior to recruitment of participants. The sample included 39 participants (29 males, 10 females) (17 Caucasian, 22 African-American) ranging from first through fifth grade. Ages ranged from 7.0 years to 12.58 years (mean age= 8.98 years). All participants were native English speakers.

Procedures

Following each evaluation, each child received a monetary compensation of $15 and each parent received a detailed reading evaluation report. Each participant’s hearing was screened according to ASHA standards (ASHA, 1997). Each child’s vision was informally screened utilizing the Eye Chart Pro iPad app from Dok LLC. All participants were required to pass both screenings. Participant responses were collected in real-time on the corresponding record forms for each test. The participants completed all four tests in one sitting with breaks provided as necessary.

Measures

Each participant completed four measures of reading abilities. Administration order of the tests was counterbalanced. To obtain a measure of single word reading fluency, the Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subtests of the *Test of Word Reading Efficiency-Second Edition (TOWRE-2)* (Torgesen et al., 2012) was administered. The composite Total Word Reading Efficiency Index was also calculated from these two subtests to provide an overall measure of single word reading. To obtain an untimed measure of single-word reading accuracy, the Word Identification (WI) and Word Attack (WA) subtests of the *Woodcock Reading Mastery Tests-Third Edition (WRMT-III)* (Woodcock, 2011) were administered. The composite Basic Reading score was also obtained from these measures. To obtain textual reading measures, the Rate, Accuracy, Fluency composite, and Comprehension scores were calculated from the *Gray Oral Reading Tests-Fifth Edition (GORT-V)* (Wiederholt & Bryant, 2012). In addition, the composite Oral Reading Index was calculated from the results of the fluency and comprehension scores. Finally, the Silent Reading Quotient (SRQ) from the *Gray Silent Reading Tests (GSRT)* (Wiederholt & Blalock, 2000) was obtained. Further information regarding each assessment tool is presented in Table II.

Statistical Analysis

Once the standardized scores from all four tests were recorded, the distributions of the scores on each test were analyzed for skewness and outliers. Initially, the scores from the different subtests were subjected to Pearson product-moment correlations analyses. In addition, the prediction of oral reading comprehension (*GORT-V* Comprehension) was investigated utilizing a series of step-wise multiple regression analyses. In the first analysis, a step-wise linear regression model was utilized on the *GORT-V* Comprehension scaled scores with all available scores serving as potential predictors (*GSRT* Silent Reading Quotient, *GORT-V* Rate, Accuracy, Fluency, *TOWRE-2* Sight-Word Efficiency and Phonemic Decoding, and *WRMT-III* Word Identification and Word Attack). In the subsequent analysis, all scores that were obtained on the *GORT-V* were excluded from the step-wise regression analysis which was repeated. To further investigate the relationship between oral reading abilities and the various subskills, the composite scores were entered into a hierarchical regression analysis. Composite scores included sight word reading fluency from the *TOWRE-2*, sight word reading accuracy from the *WRMT-III*, silent reading comprehension from the *GSRT*, and textual reading fluency from the *GORT-V*. The same series of analyses was followed when investigating the Silent Reading Quotient that was obtained from the *GSRT*.

Table II
Assessment tools used in the current study.

Assessment Tool	Subtest/Score	Summary
Test of Word Reading Efficiency-Second Edition (TOWRE-2)	• Sight Word Efficiency (SWE)* • Phonological Decoding Efficiency (PDE)* • Total Word Reading Efficiency Index (TWRE)*	Participant is allowed 45 seconds to accurately read as many words aloud as possible. Real words are used for SWE. Nonsense words are used for PDE. TWRE is a composite score.
Woodcock Reading Mastery Tests-Third Edition (WRMT-III)	• Word Attack (WA)* • Word Identification (WA)* • Basic Skills Quotient*	Participant is required to accurately read aloud real words (WA) and nonsense words (WI). Basic Skills is a composite score.
Gray Silent Reading Tests (GSRT)*	• Silent Reading Quotient*	Participant reads passages of increasing complexity silently and answers five multiple choice comprehension questions after reading with the text in view.
Gray Oral Reading Tests-Fifth Edition (GORT-V)	• Rate^{+} • Accuracy^{+} • Fluency^{+} • Comprehension^{+} • Oral Reading Index (ORI)*	Participant reads aloud passages of increasing complexity aloud and answers five open ended comprehension questions after reading without the text in view. Rate and accuracy scores combine to derive fluency score. ORI is a composite score.

Note. * indicates a standard score with a mean of 100 and a standard deviation of 15
+ indicates a scaled score with a mean of 10 and a standard deviation of 3

Results

No violations in terms of skewness or outliers were noted. In addition, no missing data was present. Table III presents a summary of the descriptive statistics from the scores that were obtained from each of the tests.

Correlations

Results of the correlation values are presented in Table IV. The oral comprehension measures that were obtained on the *GORT-V* were significantly correlated with all scores that were obtained on all other assessment tools. The highest correlation values were the values that were also obtained on the *GORT-V*. In order, they were the Oral Reading Index, Fluency, Rate, and then accuracy scores. The correlation values between these scores and comprehension ranged from .66 to .95 indicating high degrees of correlation. When comparing non-*GORT-V* scores, the Word Identification subtest from the *WRMT-III* yielded a strong relationship (.60) to the *GORT-V* comprehension scores.

Regarding silent reading comprehension as measured by the *GSRT* Silent Reading Quotient, significant correlations were found at the alpha = .05 level for Phonemic Decoding, TWRE Index, Word Identification, *WRMT* Basic Skills, and all the scores that were obtained from the *GORT-V*. The range of significant correlation values between silent reading comprehension and the various other tools was .33 to .41, indicating moderate degrees of correlation.

Oral Reading Comprehension Regression Results

When regressing *GORT-V* comprehension scores including all *GORT* subtests as possible predictors along with all other individual subtest scores, results indicated that *GORT-V* Rate accounted for the most variance, $F(1, 37) = 51.364$, $p < .001$, $R^2 = .762$, $R^2_{Adjusted} = .570$. The step-wise procedures entered no other variables.

When all *GORT-V* related variables were removed from the analysis as potential predictors of oral reading comprehension, the *WRMT-III* Word Identification subtest was found to be the leading predictor of oral reading comprehension as measured by the *GORT-V*, $F(1, 37) = 21.225$, $p < .001$, $R^2 = .604$, $R^2_{Adjusted} = .347$. The step-wise procedure entered no other variables into the equation.

Results of the hierarchical regressions between the composite scores and the comprehension scores are reported in Table 5. Analysis revealed *GORT-V* Fluency was the leading predictor of oral reading comprehension as measured by the *GORT-V*, $F(1, 37) = 49.15$, $p < .001$, $R^2 = .571$, $R^2_{Adjusted} = .559$. Adding additional variables into the model did not account for additional variance.

Table III

Descriptive Statistics: Mean Standard and Scaled Scores.

Subtest	Mean Score (SD)	Minimum	Maximum
TOWRESWE+	99.41 (12.90)	70	123
TOWREPD+	93.18 (14.31)	62	121
TOWREOVERALL+	96.15 (13.38)	64	119
WRMTID+	103.97 (15.36)	70	132
WRMTATTACK+	99.44 (13.43)	72	124
WRMTBASICSKILLS+	101.90 (14.72)	72	126
GSRTSRQ+	93.79 (13.34)	71	123
GORTRATE-	9.31 (2.17)	5	13
GORTACC-	9.41 (2.53)	4	14
GORTFLU-	9.18 (2.23)	5	14
GORTCOMP-	8.56 (2.71)	2	14
GORTOVERALL+	93.67 (12.20)	73	118

Note. + denotes standard score, average = 100, standard deviation = 15
- denotes scaled score, average = 10, standard deviation = 2

Table IV
Correlations among Predictors.

	1	2	3	4	5	6	7	8	9	10	11
TOWRE SWE	-										
TOWRE PD	.75**	-									
TOWRE Index	.93**	.94**	-								
WRMT ID	.67**	.79**	.78**	-							
WRMT WAT	.63**	.86**	.81**	.82**	-						
WRMT BASIC	.68**	.86**	.83**	.96**	.95**	-					
GSRT SRQ	.26	.37*	.33*	.38*	.30	.35*	-				
GORT RATE	.79**	.70**	.79**	.78**	.69**	.77**	.37*	-			
GORT ACC	.61**	.74**	.72**	.71**	.71**	.75**	.39*	.70**	-		
GORT FLU	.74**	.80**	.82**	.78**	.76**	.81**	.40*	.88**	.95**	-	
GORT COMP	.51**	.49**	.53**	.60**	.43**	.55**	.39*	.76**	.66**	.76**	-
GORT ORI	.65**	.67*	.71**	.73**	.62**	.71**	.41**	.87**	.84**	.92**	.95**

Note. N = 39. ** indicates correlation is significant at the $p < .01$ level, 2-tailed. * indicates correlation is significant at the $p < .05$ level, 2-tailed.

Silent Reading Comprehension Regression Results

Results of the initial step-wise regression in which all subtest scores were entered into the step-wise analysis revealed that *GORT-V* Fluency was the strongest predictor of silent reading abilities, $F\ (1, 37) = 6.964$, $p = .012$, $R^2 = .158$, $R^2_{Adjusted} = .136$. No other variables were entered into the step-wise procedure.

Results of the hierarchical regressions between the composite scores and the comprehension scores are also reported in Table V. As previously mentioned, *GORT-V* Fluency, which is a composite score complied from the Rate and Accuracy scores, was the strongest predictor of silent reading comprehension as measured by the *GSRT*.

Discussion

The current research sought to investigate the predictive values contributed to oral and silent reading comprehension by word reading and fluency. Additionally, the researchers investigated the relationship between the subprocesses of reading and oral and silent reading comprehension.

Oral Reading Comprehension

The researchers hypothesized that single-word reading accuracy would have the highest predictive value for oral reading comprehension. Although the majority of the variance observed in the *GORT-V* comprehension values can be explained from other *GORT-V* measures, it was still of interest to investigate the relationships between the *GORT-V* comprehension and non-*GORT-V* scores. Without the *GORT-V* measurements, single-word reading accuracy as measured by the *WRMT-III* was shown to be highly predictive of oral reading comprehension as well. It was somewhat surprising that an untimed single-word reading test would be more highly associated with this comprehension measure than any of the timed single-word reading measures obtained from the *TOWRE-2*. Previous studies have demonstrated that oral reading places a high demand on single-word decoding skills (Carter *et al*., 2017; Oakhill *et al*., 2003). Oakhill *et al.* found that word reading accuracy is a better indicator of reading comprehension abilities in young children ages 8 to 9 years. The average age of the children in the current study was 8.92 years of age. It is possible that if there is discernible shift in cognitive requirements from decoding to fluency, that it had not yet become robust in this sample of children.

When examining the relationships between the variables, it was not surprising that the Rate subtest of the *GORT-V* revealed to have the strongest associations with the comprehension scores that were obtained from the *GORT-V*. Within any standardized test, it is expected that the subtests correlate highly with one another in measuring the skill the test is intended to measure. However, the fact

Table V

Hierarchical Regression Models of Composite and Comprehension Scores.

	Variable Entered	*R*	*R2*	*R2+*	*β*	*p*
***GORT-V* Comprehension**						
Model 1						
	GORT-Fluency	.755	.571	.559	.755	.000
	TWRE Index	.772	.596	.573	-.281	.000
	WRMT Basic Skills	.773	.597	.563	-.346	.000
Model 2						
	TWRE Index	.532	.284	.264	.532	.000
	WRMT Basic Skills	.566	.32	.282	.254	.001
Model 3						
	WRMT Basic Skills	.547	.299	.28	.547	.000
	GORT-Fluency	.763	.582	.559	.904	.000
***GSRT* Silent Comprehension**						
Model 1						
	GORT-Fluency	.398	.158	.136	.398	.012
	TWRE Index	.398	.158	.112	.017	.045
	WRMT Basic Skills	.402	.162	.09	.110	.100
Model 2						
	TWRE Index	.333	.111	.087	.333	.038
	WRMT Basic Skills	.361	.131	.082	.247	.08
Model 3						
	GORT-Fluency	.354	.125	.102	.354	.027
	WRMT Basic Skills	.402	.161	.115	.323	.042

that the assessment tool provides two separate scores would indicate that additional information is being obtained from these two scores. The strong correlation and regression coefficients that were revealed as part of this analysis perhaps would indicate that the information being provided by these two supposedly distinct measures is somewhat redundant. An explanation for this can be found in examining the basal and ceiling procedures incorporated by this assessment. The examinees are required to meet a basal of two consecutive raw scores of 9 or 10 on the Fluency measure, which is a combination of the Rate and Accuracy scores. The ceiling is met when the examinees reach two consecutive scores of a 2 or below on the Fluency measure, regardless of comprehension performance.

The emphasis of the *GORT-V* is therefore not placed on how well the reader performs on the comprehension portion of the test, but rather on textual reading fluency (reading accuracy + reading rate). It is possible that an examinee could be reading at high comprehension levels (answering all comprehension questions correctly) and meet ceiling criteria if that individual is reading slowly and inaccurately. This would drastically underestimate comprehension abilities since the remaining items would all be scored as 0s. In addition, those readers who are reading quickly and/or accurately can be allowed to continue to proceed through the test with low levels of comprehension. If the examinee reads 8 texts and answers an average of 3 questions correct out of the 5 possible (only 60% accuracy) then they would have obtained a raw score of at least 24 depending on the entry point. Compare this to the hypothetical examinee who slowly read 4 passages and answered on average 4 of the 5 questions correctly (80%). This person's raw score would be 8 points lower than would the faster reading individual who is comprehending little of what was read. To put this into perspective, a 9-year-old child with a raw score of 24 would be assigned a scaled score of 8, which typically qualifies as within the normal range, whereas a 9-year-old child who obtains a raw score of 16 would receive a scaled score of 6 which is below the average range. If considering eligibility from these numbers alone, the child who correctly answered 60% of the comprehension questions would not be diagnosed with a comprehension deficit whereas the child who answered 80% of the questions correctly could be diagnosed with a comprehension deficit. This is of great concern and could result in both the under- and over-identification of children with comprehension problems.

At first glance, these results are slightly at odds with Cutting and Scarborough who found only moderate relationships between reading rate and *GORT* results. On the *GORT-3* which was utilized by Cutting and Scarborough, reading speed accounted for only 56% of the variance whereas in the current results, the Rate subtest accounted for 76% of the variance. The most blatant difference between these two editions comes in terms of the basals and ceilings. The *GORT-3* based comprehension testing procedures off comprehension performance. The

GORT-V does not. Although there are other differences between the two editions, the removal of this procedure appears to have drastically altered the test's implicit definition of comprehension. This alteration is so severe, that perhaps it should be considered that this specific assessment tool serves as a more accurate indicator of textual reading fluency, and should be thought of as measuring comprehension indirectly, at best. It is questionable whether the authors of the *GORT-V* have produced an assessment tool which independently assesses multiple reading constructs, or if it simply assesses the fluency by which one reads.

Cutting and Scarborough (2006) emphasize the importance of knowing how tests measure overall reading ability and reading comprehension so that scores can be interpreted and understood appropriately. They state that reading comprehension scores can vary by how this complex skill is measured through the various demands each test places on the subskills of reading comprehension. This is completely apparent when considering the newest version of the *GORT-V*, which places enormous weight upon reading rate when defining "comprehension". The relationship between oral reading rate, oral reading fluency, and reading comprehension performance in young readers is well established (Ashby *et al*., 2013; Kim *et al*., 2011; and Kim, 2015), but these constructs should not be considered as synonymous.

Silent Reading Comprehension

The silent reading comprehension measures revealed less of a pattern than did the oral reading comprehension analyses. Although associations were present, they were far less strong than were the relationships with oral reading comprehension. The strongest relationship was revealed between the Fluency score that was obtained from the *GORT-V* and the *GSRT* SRQ. This finding agrees with the current authors' hypothesis which stated that fluency would be most predictive of silent reading abilities. It is possible that the current results capture a cross-section of time in the development of reading skills in which both reading fluency and the ability to read silently are progressing in a parallel fashion.

From a clinical perspective, the lower levels of predictability for the silent reading quotient indicate that unique information is perhaps being provided by the *GSRT*. As previously stated, it appears that the *GORT-V* Comprehension score is mostly based on other skills. However, this does not seem to be the case with the *GSRT*. That is most likely due to the nature of the test. This test is silent and therefore, possibly less reliant upon decoding skill (Ashby *et al*., 2013; Kim *et al*., 2011; Kim, 2015). In addition, this test allows the reader to refer to the text that is being read. This manipulation could potentially place the emphasis more on strategic analysis that could be mitigated by the cognitive processes of memory and attention (Johnston, 1984). The current findings are similar to those Cutting and Scarborough (2006) reported, when attention was found to add to the predictive model for a similar silent reading comprehension test that also allowed the text to remain in view (*WIAT*). It is possible that the *GSRT* and other similar tests are more reliant upon the appropriate utilization of cognitive processes in addition to the bottom-up processes that oral reading is so reliant upon. Future research should investigate this link.

Limitations

The results of this study are limited to the four assessment tools that were included. Future studies investigating the relationship between the skills measured by these assessment tools in relation to reading comprehension are needed to add to the discussion of the predictive relationships between oral reading comprehension and silent reading comprehension in hopes to provide beneficial information regarding comprehensive reading assessments. This would allow for recommendations during the evaluation process including deciding on which assessment tools to include over others and a clearer focus on treating reading deficits.

Implications

The results indicate that oral reading comprehension abilities as measured by the *GORT-V* are highly related to other skills that are measured on the same test: fluency. The strong correlations that were found between comprehension and fluency, more specifically reading rate is a cause for concern. Although reading comprehension is thought to represent the culmination of many skills, the current results indicate that how the *GORT-V* assesses comprehension is highly predicated upon something else. It is possible that this finding is a result of the *GORT-V*'s utilization of fluency abilities to establish basals and ceilings which dictate test administration procedures. This test is proposed to be capable of providing a reliable measurement of comprehension ability through strengths and weaknesses noted throughout administration of the test. However, it appears from these results that the *GORT-V* is possibly more sensitive to fluency instead. For instance, a hyperlexic child who decodes without comprehension could advance to higher-level stories and achieve an inflated score that inaccurately represents the child's comprehension abilities. On the other hand, a child who has poor word decoding skills but can comprehend well could obtain lower scores that underestimate their comprehension abilities. Similar profiles were observed throughout data collection and future studies could perhaps quantify the likelihood of this occurrence. This finding indicates that potentially the only individuals who are accurately identified according to the *GORT-V* are those whose fluency levels and comprehension levels closely approximate each other. Although fluency and comprehension have been demonstrated to be highly associated, they are not interchangeable and the potential for varying proficiencies between the two skills most certainly exists. Clinicians who consider the *GORT-V* should be aware that the test is, perhaps, best suited to identify only those students who exhibit oral reading fluency deficits as opposed to reading comprehension deficits due to the *GORT-V* essentially conflating the two separable skills into a singular reading construct.

Regarding silent reading comprehension, clinicians should most certainly consider the cognitive capabilities of the individual who is being assessed. Silent periods allow the mind to roam (Carter *et al.,* 2017) and the importance of decoding is lessened in silent reading tasks (Kim *et al*., 2011). Therefore, those who exhibit attention and memory problems are at risk for performing poorly on silent reading tasks that allow the reader to refer to the text while answering questions.

For many years, clinicians have been selecting and administering assessment tools with the word "comprehension" in their titles, and assuming that they were gaining an accurate quantitative depiction of what the reader is experiencing daily during reading activities. As a result, students have been potentially inaccurately identified and more importantly, inaccurately treated. Clinicians must become acquainted with the actual content validity of the tests that they administer in order to provide individualized plans of care in an accurate and efficient manner. The current results indicate that the measures obtained on the *GORT-V* are significantly altered by fluency abilities whereas results on the *GSRT* are impacted by strategic factors.

References

American Speech-Language and Hearing Association. (1997). *Guidelines for audiologic screening: ASHA panel on audiologic assessment*. Retrieved from http://www.asha.org/policy/GL1997-00199.htm#sec1.8

Andreassen, R. & Bråten, I. (2009). Examining the prediction of reading comprehension on different multiple-choice tests. *Journal of Research in Reading*, 33, 263-283

Ashby, J., Dix, H., Bontrager, M., Dey, R., & Archer, A. (2013). Phonemic awareness contributes to text reading fluency: Evidence from eye movements. *School Psychology Review, 42*, 157–170.

Başaran, M. (2013). Reading fluency as an indicator or reading comprehension. *Educational Sciences: Theory & Practice, 13*(4), 2287-2290.

Besner, D., & Smith, M.C. (1992). Models of visual word recognition: When obscuring the stimulus yields a clearer view. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3)*,* 468-470.

Breznitz, Z. (2006). *Fluency in Reading: Synchronization of Processes*. New York, New York: Psychology Press.

Britton, B. K., & Graesser, A. C. (1996). *Models of Understanding Text*. New York, NY: Psychology Press.

Cain, K., Oakhill, J, & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42.

Carter, M. D., Walker, M. M., & O'Brien, K. (2015). The effects of rate on single-word reading assessment. *American Journal of Speech-Language Pathology, 24*, 13-23.

Carter, M. D., Walker, M. M., O'Brien, K., & Hough, M. S. (2017). The effects of text length on reading abilities in accelerated reading tasks. *Speech, Language and Hearing*, *22*(2), 111-121, DOI: 10.1080/2050571X.2017.1415086

Catts, H. W., Compton, D., Tomblin, J. B., & Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, *104*(1), 166-181.

Cohen, L. G. (1993). Test review: Wechsler Individual Achievement Test. *Diagnostique, 18*, 255-268.

Cutting, L. E. & Scarborough, H. S. (2006). 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.

Elgart, D. B. (1978). Oral reading, silent reading, and listening comprehension: A comparative study. *Journal of Reading Behavior*, *10*(2), 203-207.

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

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

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment: What, why, and how? *The Reading Teacher, 58*(8), 702-714.

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

Johnson, P. H. (1984). Assessment in reading. In P. D. Pearson (Ed.). *Handbook of reading research* (pp. 147-182). Mahwah, NJ: Lawrence Erlbaum Associates.

Kamhi, A. G. (2009). The case for the narrow view of reading. *Language, Speech, and Hearing Services in Schools, 40*, 174-177.

Kamhi, A. G. (2012). Perspectives on assessing and improving reading comprehension. In A. G. Kamhi & H. W. Catts (Eds.), *Language and Reading Disabilities* (3rd ed., pp. 146-159). Boston, MA: Pearson.

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

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*, 281–300.

Kim, Y. S. (2015). Developmental, component-based model of reading fluency: An investigation of predictors of word-reading fluency, text-reading fluency, and reading comprehension. *Reading Research Quarterly, 50*(4), 459-481.

Kim, Y. S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading, 15*(4), 338-362.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory automatic information processing in reading. *Cognitive Psychology, 6*(2), 293-323.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L. G. (2000). *Gates–MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside.

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.

Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*(4), 443-468.

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science, 42*, 722-770.

RAND Reading Study Group (RRSG). (2002). *Reading for understanding: Toward a research and development program in reading comprehension.* Santa Monica, CA. Retrieved September 2015, from http://www.rand.org/content/dam/rand/pubs/monograph_reports/MR1465/MR1465.ch2.pdf

Spencer, M., & Wagner, R. K. (2018). The comprehension problems of children with poor reading comprehension despite adequate decoding: A meta-analysis. *Review of Educational Research*, 88(3), 366-400.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2012). *Test of word reading efficiency-Second edition.* Austin, TX: Pro-Ed.

Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, *22*(3-4), 458-485.

Wechsler, D. L. (1992). *Wechsler individual achievement test*. San Antonio, TX: Psychological Corporation.

Wiederholt, J. L., & Blalock, G. (2000). *Gray silent reading tests.* Austin, TX: Pro-Ed.

Wiederholt, L., & Bryant, B. (1992). *Gray oral reading tests-Third edition.* Austin, TX: Pro-Ed.

Wiederholt, L., & Bryant, B. (2012). *Gray oral reading tests-Fifth edition.* Austin, TX: Pro-Ed.

Wolfe, M. B. W., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology, 80*, 341–362.

Woodcock, R. W. (2011). *Woodcock reading mastery tests-Third edition.* Bloomington, MN: Pearson.

Woodcock, R.W., & Johnson, M. B. (1989). *Woodcock–Johnson psychoeducational battery-Revised.* Circle Pines, MN: American Guidance Service.