View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

NIH Public Access

Author Manuscript

Sci Stud Read. Author manuscript; available in PMC 2014 May 01.

Published in final edited form as:

Sci Stud Read. 2013 May 1; 17(3): 199–223. doi:10.1080/10888438.2011.652722.

Examining the Relationship Between Word Reading Efficiency and Oral Reading Rate in Predicting Comprehension Among Different Types of Readers

Sarah H. Eason,
University of Maryland

John Sabatini,
Educational Testing Service, Princeton, New Jersey

Lindsay Goldberg,
University of North Carolina at Chapel Hill

Kelly Bruce, and
Educational Testing Service, Princeton, New Jersey

Laurie E. Cutting
Kennedy Krieger Institute, Baltimore, Maryland

Abstract

To further explore contextual reading rate, an important aspect of reading fluency, we examined the relationship between word reading efficiency (WRE) and contextual oral reading rate (ORR), the degree to which they overlap across different comprehension measures, whether oral language (semantics and syntax) predicts ORR beyond contributions of word-level skills, and whether the WRE–ORR relationship varies based on different reader profiles. Assessing reading and language of average readers, poor decoders, and poor comprehenders, ages 10 to 14, ORR was the strongest predictor of comprehension across various formats; WRE contributed no unique variance after taking ORR into account. Findings indicated that semantics, not syntax, contributed to ORR. Poor comprehenders performed below average on measures of ORR, despite average WRE, expanding previous findings suggesting specific weaknesses in ORR for this group. Together, findings suggest that ORR draws upon skills beyond those captured by WRE and suggests a role for oral language (semantics) in ORR.

Many different components contribute to a child's ability to comprehend written text. It is well documented that word recognition and oral language are significant predictors of general reading comprehension (Hoover & Gough, 1990; Storch & Whitehurst, 2002). In addition, reading fluency has been suggested as another factor underlying comprehension (Adlof, Catts, & Little, 2006). However, there has been considerably less focus on fluency than on word recognition, and debates exist concerning how best to define and measure reading fluency.

For some, fluency is essentially defined as automatized word recognition (LaBerge & Samuels, 1974), or fast and accurate word reading (e.g., Torgesen, Wagner, & Rashotte, 1999). High correlations between sight word efficiency and continuous text rate—measured

Copyright © 2012 Society for the Scientific Study of Reading

Correspondence should be sent to Laurie E. Cutting, Vanderbilt University, 230 Appleton Place, Nashville, TN 37203. laurie.cutting@vanderbilt.edu.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

as words correct per minute when reading continuous text passages—suggest to some that automatic word recognition is the primary element of a reading fluency construct and that the two are essentially measures of the same construct (e.g., Barth, Catts, & Anthony, 2009). This viewpoint stems from theories related to automaticity and verbal efficiency theory (LaBerge & Samuels, 1974; Perfetti, 1985, 2003; Perfetti & Hogaboam, 1975). Reading words automatically (i.e., accurately, rapidly, and with relatively little expenditure of attentional resources or conscious awareness) frees up cognitive resources for comprehension processes. If a reader does not have automatized word recognition, then a "bottleneck" is created and comprehension suffers. Thus, efficient word reading facilitates comprehension in that it is less taxing of shared resources between the two sets of cognitive processes (Perfetti, 1985, 2003). Therefore, a reader with poor word recognition skills would be expected to struggle with reading words automatically (both in and out of context) and, as a consequence, would have poor comprehension.

Other views of reading fluency suggests that although word reading efficiency (WRE) is a component of reading fluency, it does not encompass all elements of fluency (Daane, Campbell, Grigg, Goodman, & Oranje, 2005; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Klauda & Guthrie, 2008; Samuels, 2006). For example, Wolf and Katzir-Cohen (2001) suggested a model of fluency comprising an extensive list of processes: "Lower level attention and visual perception, orthographic (letter-pattern) representation and identification, auditory perception, phonological representation and phoneme awareness, short-term and long-term memory, lexical access and retrieval, semantic knowledge, and connected-text knowledge and comprehension" (p. 220). In addition, they proposed that aspects of executive function are necessary to coordinate the various processes and that deficits in any one or more of these processes may lead to dysfluency. Others have emphasized the role and measurement of prosody or expressiveness (Daane et al., 2005; Kuhn, Schwanenflugel, & Meisinger, 2010). This expanded view of fluency suggests that word recognition is a necessary prerequisite, but not sufficient, for fluency.

Clearly, understanding the elements of fluency has significant implications for the development of both assessments and interventions of reading fluency. In this article, we focus on an important but specific issue related to fluency: What (if anything) distinguishes WRE from continuous-text passage oral reading rate (ORR)? For the purposes of this study, WRE was defined as fast, accurate reading of single words presented either in a list or one at a time (with a measure of voice onset time as an index of efficiency). ORR was defined as fast, accurate reading of a story or paragraph aloud.

Although some work has focused on the issue of what may distinguish WRE and ORR, further elucidation is needed. For example, although several investigations have shown that ORR contributes unique variance to reading comprehension beyond WRE (Denton et al., 2011; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003a; Klauda & Guthrie, 2008; Samuels, 2006); others have not reported distinctions between WRE and ORR (Barth et al., 2009). These conflicting findings may have to do with varying measures of reading comprehension and/or the populations studied, which at this time has not been fully explored. Furthermore, studies distinguishing WRE and ORR have hypothesized—but not fully tested—the origins of the distinction between the two and have suggested that ORR, but not WRE, captures elements of oral language (Denton et al., 2011; Jenkins et al., 2003a). Thus, even though some have reported distinctions between ORR and WRE, very little information exists as to what may be associated with these differences. In sum, we may gain further clarification about the relationships between WRE–ORR and, in turn, to reading comprehension by understanding what (if anything) distinguishes WRE from ORR with regard to (a) contributions to different reading comprehension measurements, (b) the influence of oral language, and (c) how the relationship varies for different reading profiles.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

READING COMPREHENSION MEASUREMENT

Previous studies have found that reading comprehension performance can vary by type of comprehension measure used; factors such as text characteristics and text length can influence readers' performance on comprehension tests by differentially drawing upon various cognitive abilities (see Cutting & Scarborough, 2006; Francis, Fletcher, Catts, & Tomblin, 2005; Keenan, Betjemann, & Olson, 2008). Nevertheless, although the issue of reading comprehension formats has been explored with relation to WRE and oral language, whether WRE and ORR contribute similarly to reading comprehension across multiple measures of reading comprehension, especially to those with different formats, has not been examined. To date, investigations showing that ORR contributes unique variance in predicting reading comprehension beyond WRE have been confined to passage-length multiple-choice formats (Denton et al., 2011; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003b). It is plausible that both the specific portions of shared and unique variance as well as the overall amount of variance that WRE and ORR account for may be different depending on the comprehension assessment; finding that ORR contributes additional variance beyond WRE to multiple measures of reading comprehension would reveal how robust ORR's uniqueness from WRE is, and whether this varies depending on measure. Such findings would be valuable for understanding what types of reading comprehension ORR may be capturing (e.g., formats similar to those on state comprehensive tests, or other formats not typically used in high-stakes testing).

ORAL LANGUAGE AND ORR

A small but growing literature of research comparing WRE and ORR's contributions to reading comprehension report that ORR contributes additional variance beyond WRE to comprehension, but not vice versa (Denton et al., 2011; Jenkins et al., 2003b; Klauda & Guthrie, 2008), supporting the possibility that ORR captures an element not in WRE. Each of these investigations has suggested that ORR may capture oral language processes critical for reading comprehension; however, this supposition has not been comprehensively explored with consideration to different elements of oral language. For example, Jenkins et al. (2003b) hypothesized, but did not test, that ORR's additional contribution of variance to reading comprehension reflected oral language, specifically semantics and syntax. Recently, Denton et al. (2011) found that a measure of vocabulary/verbal reasoning was related to ORR; however, more refined and specific measures of oral language, particularly syntax, were not included in this study. In addition, there is some indication that syntax may relate to reading words in context; Cain (2007) found that syntactic awareness predicted contextual word reading, although it is important to note that the reading measure assessed *accuracy* of words in context, not ORR. Thus, it is an open question to what extent semantics and syntax may contribute to ORR; understanding which specific elements of oral language are related to ORR would provide more fine-grained information about the types of cognitive processes it captures.

ORR AND SUBTYPES OF READING DEFICITS

Finally, although reading fluency has been explored in reader subtypes, no investigation as of yet has directly compared WRE and ORR across various well-defined reader subtypes. In particular, the inclusion of impaired readers—distinguishing between those with documented word-level deficits and those with specific reading comprehension impairment (i.e., no word-level deficits)—may further clarify the differences between WRE and ORR. Those with word-level deficits would be presumed to have poor WRE and, as a consequence, poor ORR; however, poor comprehenders, by definition, do not have difficulty

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

with WRE. Therefore, whether or not they perform poorly on ORR may help shed light on ORR processes.

It is important to note that although some studies have investigated whether the relationship between word recognition in and out of context differs as a function of child's reading ability, these findings do not yet provide a comprehensive understanding of how the relationship may vary across different types of readers (e.g., Jenkins et al., 2003a; Nation & Snowling, 1998a). For example, Jenkins et al. (2003a) examined readers with varying profiles of reading deficits (RD) in a study comparing skilled and impaired readers, utilizing both WRE and ORR. Students with RD, as identified by performing at or below the 25th percentile on a standardized reading measure, consistently performed more poorly than skilled readers when reading both lists of words and passages. No students demonstrated a specific impairment in ORR. However, it is important to note that word recognition/decoding was not used as a grouping variable in this study; therefore, it is probable that many of those children with RD had word reading deficits (as the findings showed). Therefore, it is not surprising that their RD group showed *both* WRE and ORR deficits, as WRE is a necessary component of ORR. It is possible, of course, that their RD group consisted of some with word reading deficits and some without, but this was not examined in the study. This is an important point to consider, as children who struggle with comprehension due to word difficulties have shown to have different profiles than children who are poor comprehenders despite average word-level abilities (specifically poor comprehenders; Cain & Oakhill, 2006; Cutting, Materek, Cole, Levine, & Mahone, 2009; Nation & Snowling, 1998a, 1998b, 2000). In particular, poor comprehenders show weaknesses in oral language; therefore, it is possible that if ORR captures oral language processes, the RCD subtype would show impairments in ORR but not WRE.

Indeed, findings from studies that have used specific, well-defined grouping criteria for RD subtypes have found that children with RCD do not show the same level of contextual facilitation as would be expected given their word-level ability (Cutting et al., 2009; Nation & Snowling, 1998a). For example, Nation and Snowling (1998a) examined contextual facilitation for specific reading profiles and defined RCD as poor performance on both listening and reading comprehension measures, but with decoding skills on par with average readers. After listening to a sentence and then reading the final word in print, they found that poor comprehenders' reading was not facilitated by context to the same extent as that of average readers. A more recent study by Cutting et al. (2009) compared children with specific reading comprehension deficits (i.e., those without decoding deficits), children with general reading deficits (both decoding and comprehension), and average readers by investigating their oral language, print, and executive function skills. Children with specific reading comprehension deficits read at the same level as average readers on measures of WRE but demonstrated a relative weakness on a measure of ORR.

Although these studies suggest that poor comprehenders have difficulty with ORR, questions still remain regarding ORR for these readers. For example, although the Nation and Snowling (1998a) study revealed that contextual facilitation at the sentence level is not equal for all readers, the design did not include a measure of ORR. Providing the context (a sentence stem) orally allowed readers to avoid obstacles of decoding and comprehending the entire sentence before reaching the final key word. Therefore, there are limits to what conclusions can be made regarding how poor comprehenders navigate a passage as a whole, and understanding any potential differences between WRE and ORR in poor comprehenders cannot be concluded from the Nation and Snowling (1998a) study. In Cutting et al.'s 2009 study, the measures of WRE and ORR were from separate standardized tests, and the differences in poor comprehenders' performance could be attributed to test characteristics other than whether words were presented in isolation or context (e.g., word regularity). In

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

addition, the measures of reading comprehension were limited, and examination of the relationship of oral language to ORR was not addressed. Thus, exploration of whether children with poor comprehenders, in contrast to those with deficits in WRE, exhibit specific difficulties with ORR, and whether oral language abilities may account for these difficulties, is needed.

THE PRESENT STUDY

The current study aimed to further clarify the relationship between WRE and ORR by examining the overarching question of what (if anything) distinguishes WRE from ORR. To investigate this question, we focused on three specific subquestions, each of which contributed to addressing our overarching question:

1. Is the relationship between WRE and ORR similar across reading comprehension measures that vary in format? Showing that ORR contributes additional variance beyond that of WRE across different types of reading comprehension measures would reveal the robustness and consistency of the relationship between the two. Nevertheless, although we expected based on previous findings that ORR would contribute additional variance beyond WRE for all reading comprehension measures, given the differences among reading comprehension formats, we hypothesized that the contributions would vary depending on measure used. In particular, we hypothesized that ORR may predict more unique variance for measures that consisted of longer passages versus those that were shorter and have been shown to rely heavily on WRE (e.g., cloze format).

2. If, as expected, ORR contributes additional variance to reading comprehension measures beyond WRE, do oral language skills explain the differences between the two? We hypothesized that both semantics and syntax would explain unique variance in ORR, given that efficiency of reading text orally should draw upon both semantic and syntactic skills.

3. Does the relationship between WRE and ORR vary depending on reader characteristics? Finding that the relationship between WRE and ORR varies based on reader characteristics would provide another form of evidence for the distinctness between the two. Clearly those with WRE deficits would be expected to show poor ORR largely as a result of poor WRE. However, understanding ORR performance in poor comprehenders (who show adequate WRE) would provide further clarification as to the relationship between WRE and ORR. If poor comprehenders showed poor ORR, such findings would suggest that ORR captures cognitive processes beyond those captured by WRE. Furthermore, if findings showing that poor comprehenders' weaknesses in oral language account for their poor ORR, this would provide further suggestion that ORR captures elements of oral language not captured by WRE. Alternatively, children with specific comprehension deficits may not show weaknesses in ORR; if such a result were revealed, along with finding that ORR did not contribute unique variance beyond WRE across various reading comprehension measures (Question 1), such findings would suggest redundancy between WRE and ORR. In sum, findings of varying WRE–ORR patterns between different reader profiles (i.e., poor decoders vs. poor comprehenders) would provide further clarification as to the relationship between the two, and in particular may provide insights as to why some studies have shown redundancy between WRE and ORR (Barth et al., 2009), whereas others have not (Denton et al., 2011; Jenkins et al., 2003b).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

METHOD

Participants

Children ages 10 to 14 years were recruited through flyers distributed in the community to participate in a 2-day study of reading comprehension. Potential participants were screened prior to participation and were excluded if they met any of the following criteria: (a) first language other than English; (b) previous diagnosis of mental retardation or pervasive developmental disorder; (c) known, uncorrected visual impairment; (d) documented hearing loss of 25 decibels or more in either ear; (e) history of known neurological disorder (e.g., epilepsy, cerebral palsy); (f) treatment with psychotropic medications for any psychiatric disorder other than attention deficit hyperactivity disorder (ADHD); and (g) both Verbal Comprehension Index (VCI) and Perceptual Reasoning Index (PRI) scores below 80 from the Wechsler Intelligence Scale for Children–Fourth Edition (WISC–IV; Wechsler, 2003). Given the high comorbidity between reading disabilities and ADHD, children with a previous diagnosis of ADHD were included in the study; similarly, children with a diagnosis of oppositional defiant disorder and/or adjustment disorder were not excluded from participation. Children taking stimulants for the treatment of ADHD were asked to stop taking them the day before and the days of testing. Children with ADHD who were treated with medications other than stimulants were excluded.

The sample for this study included 88 children with a mean age of 11.78 (SD = 1.31). Fifty-two (59%) participants were male; 59 (67%) were Caucasian, 21 (24%) were African American, 2 (2%) were Asian, and 3 (3%) were identified by their parents as more than one race; 50 (57%) described themselves as non-Hispanic/Latino, 3 (3%) were Hispanic/Latino. There were significantly fewer African American children in the group of average readers than in either of the RD groups ($\chi^2 = 8.36$, $p = .015$); all other ethnicities were represented equally among the three reading groups.

Grouping Criteria

To address the first question, participants were assigned into three groups based on word reading and comprehension ability. The criterion for word reading was based on the Basic Skills Cluster from the Woodcock Reading Mastery Test–Revised/Normative Update (WRMT–R/NU), a composite of the Word Identification and Word Attack subtests. Comprehension was assessed by five reading comprehension measures: Reading Comprehension from the Stanford Diagnostic Reading Test, Fourth Edition (SDRT–4), Comprehension from the Gates-MacGinitie Reading Tests–Fourth Edition (GMRT–4), Reading Comprehension from the Diagnostic Achievement Battery (DAB), the Comprehension score from the Gray Oral Reading Test–Fourth Edition (GORT–4), and Passage Comprehension from the WRMT–R/NU. *Poor decoders* were defined as children with a WRMT–R/NU Basic Skills Cluster Score at or below the 25th percentile. Children who scored at or above the 37th percentile on the Basic Skills Cluster but scored at or below the 25th percentile on two or more of the reading comprehension measures were classified as *poor comprehenders*. The *average reader* group consisted of children who scored at or above the 37th percentile on the Basic Skills Cluster, as well as on at least four of the comprehension tests, with the fifth comprehension score being no lower than the 26th percentile.

In addition, participants were categorized based on ADHD diagnosis. To meet the criteria for ADHD for the study, children must have a T score of 65 or greater on the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) Inattentive or Hyperactive/Impulsive scale of the Conners' Parent Rating Scale (Conners, 1997) *and* have scores of 2 or 3 on at least six of the nine items for either the Inattentive or

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Hyperactive/Impulsive scales on the ADHD Rating Scale Home Version (DuPaul, Power, Anastopoulos, & Reid, 1998).

Screening Measures

At the start of participation, children were screened for eligibility to ensure that they met criteria for one of the reading groups, as well as ensuring that their general intellectual functioning was in the appropriate range. Children were required to have a Full-Scale IQ (FSIQ) of 80 or above. Children with an FSIQ below 80 were accepted on a case-by-case basis if their VCI or PRI was above 80.[1] Although children with an FSIQ above 120 were included in the larger study, they were excluded from the analyses discussed in this article so that the control group was representative of a more average population; similarly, children with a WRMT–R/NU Word Identification Standard Score above 120 were not included in analyses.

WISC–IV (Wechsler, 2003)—FSIQ along with VCI and PRI were used to assess overall intellectual ability. According to technical reports for the WISC–IV internal consistency reliability ranges from .92 for the PRI to .97 for the VCI, and test–retest reliability ranges from .89 to .93. Due to the obvious linguistic nature of reading, when analyses required covarying for IQ, a non-verbal subtest of the WISC–IV, Matrix Reasoning was used rather than FSIQ.[2] Matrix Reasoning consists of a series of increasingly complex matrices with a part missing. Children are required to select the missing part of the matrix from five options. Internal consistency for Matrix Reasoning is reported to be .89.

WRMT–R/NU (Woodcock, 1998)—To screen for word reading ability, children were assessed by two subtests: Word Identification, a measure of word recognition in which children are given a list of words of increasing difficulty, and Word Attack, a measure of decoding in which children read a series of pseudowords to read aloud. Both subtests are untimed, and therefore solely assess accuracy (not efficiency). Age-based standard scores from Word Identification and Word Attack were combined to form the Basic Skills Cluster. The Basic Skills Cluster standard score and percentile were used to group participants based on word reading ability.

The Passage Comprehension subtest requires participants to silently read a sentence or short paragraph and supply the missing word. Standard scores and percentiles based on age for Passage Comprehension were used to group participants based on comprehension ability. Internal consistency reliability for the WRMT–R/NU tests has been reported to range from .68 to .98; a split-half reliability range from .87 to .98 has been observed for the Cluster scores.

SDRT–4 (Karlsen & Gardner, 1995)—The Reading Comprehension subtest is timed and requires participants to read both narrative and expository texts prior to answering multiple-choice questions about the passages. Percentiles, based on grade-level norms, were used to group participants based on comprehension ability. Reliability estimates for the SDRT–4 ranged from .91 to .93.

[1]There were six participants included despite having a FSIQ below 80. When there are great discrepancies (i.e., more than 20 points) between index scores on the WISC–IV, it is recommended that the FSIQ not be used because it does not accurately reflect a child's level of cognitive functioning; instead, the individual index scores should be examined. This was the case for five of the participants. An additional child did not actually show a 20-point discrepancy but had index scores that were all close to 80 (two indexes were above 80, and two were just below).

[2]Both scaled scores and raw scores for WISC–IV Matrix Reasoning were used for analyses, depending on whether the particular analysis included only standardized measures or both standardized and experimental measures.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

GMRT–4 (MacGinitie, MacGinitie, Maria, & Dreyer, 2002)—The Comprehension sub-test is a timed assessment of reading comprehension, during which participants read passages silently and respond to multiple-choice questions. Scores are based on grade-level norms; participants were grouped based on percentiles. For all levels, the manual reports internal consistency reliability as ranging from .91 to .93 and alternate form reliability as ranging from .80 to .87.

DAB (Newcomer, 2001)—The Reading Comprehension subtest involves silent reading of passages, followed by open-ended questions that are administered orally by the examiner. Age-based scaled scores were used for grouping participants. Reliability estimates for the DAB were reported as ranging from .89 to .93.

GORT–4 (Wiederholt & Bryant, 2000)—Participants read a passage out loud and then respond to orally administered multiple-choice questions. Reading rate, accuracy, and fluency, as well as comprehension, are assessed. Age-based scaled scores from the Comprehension subscore were used to group participants. The internal consistency of the GORT–4 Comprehension subtest has been reported to be .93.

Fluency Measures

Both standardized and experimental measures were used to assess WRE and ORR. The standardized measures have the benefit of having normative distributions of scores based on large samples; however, as mentioned earlier, because they are not designed in tangent, they do not necessarily present text in a similar fashion or share a substantial number of the same words. Thus, the experimental measures, which address these issues, were also included.

GORT–4 (Wiederholt & Bryant, 2000)—The Fluency score, a composite of both accuracy and rate, was used as a measure of ORR. When children read the GORT stories out loud, both errors and time taken to complete the story are recorded. Each story has an individual rubric ranging from 0 to 5 for rate (based on number of seconds taken to read the story) and accuracy (based on number of errors); the rate and accuracy scores are combined to calculate the raw fluency score. Stories increase in difficulty; because the starting point is determined by a child's age and testing is discontinued when a child reaches a certain rate of error, which stories are presented and how many stories presented varies from child to child. Raw scores are then converted to age-based scaled scores. Internal consistency for the Fluency subtest of the GORT–4 is reported as .93.

Test of Word Reading Efficiency (TOWRE; Torgesen et al., 1999)—To measure WRE, the Sight Word Efficiency (SWE) subtest was administered; during this task the participant reads a list of words of increasing difficulty. Word efficiency rate is measured by the number of words read correctly in 45 s. The raw scores from the SWE subtests are converted to age-based standard scores. The TOWRE manual reports reliability ranges of .90 to .99.

SARA Battery (Sabatini, 1997, 1998, 2000; Sabatini, Venezky, Kharik, & Jain, 2000)—Two subtests from the SARA battery were used to assess reading rate/efficiency. The SARA battery is a computer-administered experimental task. For all subtests, children sit in front of a computer and listen to prerecorded instructions for each task. The Word Recognition subtest measures rate and accuracy for WRE by presenting 76 words one at a time on the computer monitor; 35 of the 76 words are also used from one of the two passages used for the Passage Reading task; this is the passage used for analyses in the study. Children read the words aloud into a microphone; both reaction time (voice onset time) and accuracy are recorded and the next word appears automatically. Passage Reading

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assesses rate and accuracy of contextual reading by having participants read several short stories out loud. The stories are presented one paragraph at a time, and participants press the space bar to move forward to the next paragraph. The tester records words that were skipped or read incorrectly. For each subtest, fluency was analyzed as words correct per minute (WCPM). Analyses of the reliability for the SARA battery have yielded alpha coefficients ranging from .86 to .95.

Rapid Naming Measure

CTOPP Rapid Letter Naming (Wagner, Torgesen, & Rashotte, 1999)—The rapid letter naming subtest of the CTOPP requires children to name out loud four rows of letters as quickly as possible; there are two trials. The sum of the time taken on both trials is the raw score, which is then converted to age-based scaled scores. This was included as a measure of naming speed. The internal consistency coefficient alpha is reported as .82.

Language Measures

In addition to the measures used for screening and assessing fluency, other tests were used to assess language skills.

Peabody Picture Vocabulary Test–Third Edition (PPVT-III; Dunn & Dunn, 1997)—The raw score and age-based standard scores from the PPVT–III were used as a measure of vocabulary. Participants select one of four pictures that corresponds with a word read by the examiner. Alpha coefficients for the PPVT–III range from .92 to .98.

WISC–IV Vocabulary subtest—The WISC–IV Vocabulary test was also used as a measure of word knowledge. Participants provide definitions for words, which are presented both in text and orally by the examiner. Raw scores and age-based scaled scores were used in analyses. Internal consistency for the WISC–IV Vocabulary has been reported as .89.

Test of Language Development–Intermediate, Third Edition (TOLD-I:3; Newcomer & Hammill, 1997)—The Grammatic Comprehension subtest of the TOLD requires children to listen to a sentence and decide whether it is grammatically correct or incorrect. The raw score from Grammatic Comprehension was used to assess syntactic ability. Reliability for the Grammatic Comprehension subtest has been reported as .95.

Syntactic comprehension—To further assess syntactic awareness, we administered an experimental measure of syntax in which children listened to sentences with two levels of syntactic complexity (Subject-Subject vs. Subject-Object) and answered true–false questions about each one (based on a task by Booth, MacWhinney, & Harasaki, 2000). Performance was assessed by number of total items correct.

Procedure

Each child's participation took place over the course of 2 days; in most instances, the 2 days of testing were no more than a week apart. Behavioral testing lasted for approximately 4 to 5 hr each day. A psychology associate with a master's degree administered all standardized tests. The experimental measures were administered by a psychology associate or a research assistant with a bachelor's degree.

Data Analyses

To address the first subquestion, hierarchical regressions were used to examine the unique contributions of WRE versus ORR to reading comprehension. To address the second subquestion, whether language skills might account for differences between participants'

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

WRE and ORR, hierarchical regressions were run to examine the contributions of vocabulary and syntax to ORR, and then to WRE. Finally, to address the last subquestion, whether the relationship between WRE and ORR varies depending on reader profiles, we ran two repeated measures analyses of variance (RM-ANOVAs) to examine within- and between-group differences on WRE and ORR. Note that for analyses in which only standardized measures were included, standard scores were used; for analyses including at least one experimental measure, age was controlled for and raw scores were used for all measures.

RESULTS

Sample Characteristics

Overall means and standard deviations of all measures for the whole group, which were used to answer the first two subquestions, can be found in Table 1. To answer the third subquestion, we grouped participants based on reading ability. After grouping participants based on reading ability, 36 (41%) children were categorized as average readers, 34 (39%) were categorized as poor decoders, and 18 (21%) were categorized as poor comprehenders. The groups did not differ on age, $F(2, 85) = 0.87$, $p = .42$; ratio of male to female participants ($\chi^2 = 0.78$, $p = .68$); or percentage of participants meeting criteria for ADHD ($\chi^2 = 0.99$, $p = .610$). Mean age, ratio of male to female participants, and frequency of ADHD for each group are reported in Table 1. There were significant differences in general cognitive ability, as measured by WISC–IV Matrix Reasoning, $F(2, 84) = 18.10$, $p <.001$; the average readers had higher scores than both poor reading groups ($p < .001$). There were also significant differences between the groups on the Basic Skills Cluster from the WMRT–R/NU, $F(2, 85) = 103.95$, $p < .001$, $\eta^2_p = .710$, with both average readers ($p < .001$) and poor comprehenders ($p < .001$) having higher scores than the poor decoders and average readers having higher scores than the poor comprehenders ($p = .008$). When examining the two subtests that compose the Basic Skills Cluster, it was found that scores from both the Word Identification, $F(2, 85) = 113.73$, $p < .001$, $\eta^2_p = .728$, and Word Attack, $F(2, 85) = 80.84$, $p < .001$, $\eta^2_p = .655$, subtests were significantly different between groups. Specifically, the poor decoders had lower scores than the average readers and poor comprehenders on both subtests (all $p < .001$). The poor comprehenders had lower scores than the average readers on Word Identification ($p = .001$); however, the average readers and poor comprehenders performed similarly on Word Attack ($p = .284$). It is not surprising that this pattern of findings would exist, as poor comprehenders should be unimpaired in their phonological decoding skills but would likely exhibit some weaknesses in real word identification, due to their poor semantic skills. Mean scores and group differences for all screening measures are listed in Table 1.

Contributions of WRE and ORR to Reading Comprehension

A series of hierarchical regressions was conducted to examine the contributions of WRE and ORR to performance on each of the five comprehension assessments (GMRT–4, SDRT–4, WRMT Passage Comprehension, DAB, GORT–4 Comprehension). Mean scores and correlations for the comprehension and fluency measures are summarized in Table 2; Table 3 provides a summary of the regression analyses. WRE (TOWRE SWE) and ORR (GORT Fluency) were entered in separate steps and then reversed in order to determine the amount of unique variance for which each variable accounted. The overall model was statistically significant for all five comprehension measures, accounting for 24% (on the DAB) to 60% (on the WRMT Passage Comprehension) of the variance. ORR added 19% (on the DAB) to 25% (on the GORT–4 Comprehension) unique variance when entered in the second step (all $p < .001$). The unique variance of WRE ranged from 0.1% (on the SDRT–4) to 3% (on the GORT–4 Comprehension) and was nonsignificant.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

When using the experimental SARA measures to predict reading comprehension, a similar pattern was observed. After accounting for age, SARA Word Recognition and SARA Passage Reading together accounted for 16% (on the DAB) to 61% (on the WRMT Passage Comprehension) of the variance in comprehension scores. SARA Passage Reading added 8% (on the DAB) to 28% (on the GMRT–4) unique variance when entered in the last step (all $p < .01$). The unique variance of SARA Word Recognition ranged from 0% to 1.8% and was nonsignificant.

Predictors of ORR

Hierarchical regressions were run to examine the contributions of language skills to ORR, measured by GORT–4 Fluency, after taking word-level and rapid naming abilities into account. Age, Letter-Word Identification, and Word Attack from the WRMT–R/NU, and Rapid Letter Naming from the CTOPP were entered in the first step. The second step added vocabulary, as measured by the WISC Vocabulary subtest and the PPVT. The third step added syntax, measured by TOLD Grammatic Comprehension and the Sentence Comprehension task. The overall model was significant ($R^2 = .766$), $F(8, 74) = 30.29$, $p < .001$, and accounted for 77% of the variance for ORR. After taking word-level and rapid naming skills into account, language contributed an additional 6% of the variance; however, only vocabulary contributed a significant amount of unique variance (4%, $p < .01$).

Similar regressions were run to examine predictors of Passage Reading from the SARA Battery. The overall model was significant ($R^2 = .772$), $F(8, 74) = 31.32$, $p < .001$, and predicted 77% of variance on SARA Passage Reading. After taking word-level and rapid naming skills into account, language contributed an additional 3% of the variance; however, only vocabulary contributed a significant amount of unique variance (2%, $p < .05$). Results are summarized in Table 4.

To confirm that the influence of language is unique to reading in context, we ran the same regression models to predict WRE. When WRE was measured by the TOWRE, the model was statistically significant ($R^2 = .728$), $F(8, 74) = 24.81$, $p < .001$, and accounted for 73% of the variance in TOWRE scores. After taking word-level and rapid naming skills into account, neither vocabulary nor syntax contributed a significant amount of unique variance. Regressing the SARA measure of single word fluency on the same model yielded similar results. For SARA Word Recognition, the overall model was statistically significant ($R^2 = .409$), $F(8, 74) = 6.39$, $p < .001$, and accounted for 41% of the variance in WCPM in isolation. Neither vocabulary nor syntax contributed a significant amount of unique variance after accounting for word-level and rapid naming ability.

To verify that the variance accounted for by language was not reflective of the influence of overall cognitive ability, the regressions predicting the ORR and WRE measures were rerun, adding Matrix Reasoning as a measure of general cognitive ability. There were no changes in the contributions of vocabulary to ORR and WRE, even after taking general cognitive ability into account.

Between-Group and Within-Group Differences on Measures of Fluency

Two RM-ANOVAs were run to explore the within-subject factor of fluency type (WRE vs. ORR) and the between-subject factor of reading group (average readers vs. poor decoders vs. poor comprehenders; see Tables 5 and 6). The first RM-ANOVA used the two standardized fluency measures—the TOWRE SWE (WRE) and the GORT Fluency (ORR) scores and, because of the clear cognitive ability differences between groups, covaried for Matrix Reasoning. There was a main effect of group, $F(2, 83) = 69.14$, $p < .001$, $\eta^2_p = .625$, but no main effect of fluency type, $F(1, 83) = 2.73$, $p = .102$, $\eta^2_p = .032$. Of greater interest,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

however, was a significant Fluency Type × Group interaction, $F(2, 83) = 21.36$, $p < .001$, $\eta^2_p = .340$. Although the average readers performed similarly on the TOWRE and the GORT, the poor decoders and poor comprehenders both performed more poorly on the GORT ($p < .001$); see Figure 1. When comparing the groups on each fluency type, the poor comprehenders' TOWRE scores were not significantly lower than those of the average readers; however, they did perform more poorly than the average readers on the GORT ($p < .001$). The poor decoders consistently performed at a lower level than both the average readers and poor comprehenders (all $p < .001$).

The second RM-ANOVA compared the two experimental fluency measures—the SARA Word Recognition (WRE) and SARA Passage Reading (ORR)—and covaried for age and Matrix Reasoning. There were significant main effects between groups, $F(2, 82) = 44.36$, $p < .001$, $\eta^2_p = .520$, as well as within groups, $F(1, 82) = 4.68$, $p < .05$, $\eta^2_p = .054$. The within-group main effect is not surprising; as performance was measured in raw scores (WCPM), it would be expected that all groups would show contextual facilitation to some extent. As with the standardized measures, a Fluency Type × Group interaction was observed, $F(2, 82) = 13.65$, $p < .001$, $\eta^2_p = .250$. All of the groups had higher WCPM on SARA Passages, although the difference between WRE and ORR was larger for average readers than it was for the poor comprehenders and poor decoders. It is important to note that, as seen with the TOWRE and GORT, the average readers and poor comprehenders performed similarly on the WRE task, but the poor comprehenders' rate for ORR was significantly lower than that of the average readers ($p < .05$); see Figure 1. Not surprisingly, the poor decoders' mean rates were lower than the average readers and poor comprehenders for both fluency measures (all $p < .001$).

Additional Analyses

Based on the observed influence of language skills—specifically semantics—on ORR, a series of follow-up multivariate analyses of covariance were conducted to confirm if vocabulary accounted for the differences between the average readers and poor comprehenders on measures of ORR. First, the two groups' vocabulary scores (WISC–IV Vocabulary and PPVT) were compared. After covarying for general cognitive ability (measured by WISC–IV Matrix Reasoning), there was a significant multivariate main effect for group, $F(2, 50) = 5.93$, $p < .01$, $\eta^2_p = .192$. Significant univariate effects for group were observed for each language measure such that average readers performed better on both the WISC–IV Vocabulary and PPVT (both $p < .01$).

Next, we ran a multivariate analysis of covariance comparing the average readers' and poor comprehenders ORR scores on the GORT and the SARA passages while taking into account the difference in vocabulary for the two groups. After covarying for age, general cognitive ability (WISC–IV Matrix Reasoning) and vocabulary (WISC–IV Vocabulary and PPVT), there was no significant main effect of group, $F(2, 47) = 1.78$, $p = .180$, $\eta^2_p = .070$. Taken together, these analyses suggest that poor comprehenders' slower ORR can be accounted for by, in part, their oral language abilities, specifically semantics.

DISCUSSION

This study sought to investigate one central but important issue related to fluency: What (if anything) distinguishes reading words quickly in isolation versus in context? Although results of some studies have suggested that ORR does not contribute additional variance to reading comprehension above and beyond WRE, others have found that ORR captures additional variance beyond WRE, raising questions about the WRE–ORR relationship (e.g., Barth et al., 2009; Jenkins et al., 2003b). In addition, little is known about the oral language processes that may be related to ORR versus WRE. The current study focused on clarifying

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

the WRE–ORR relationship by posing three subquestions, each of which addressed a different facet of our overarching question: (a) the issue of formats for reading comprehension measures, (b) the role of oral language in ORR, and (c) reader profiles.

Is the relationship between WRE and ORR similar across reading comprehension measures that vary in format?—We sought to replicate and extend previous findings showing that ORR contributes additional variance to reading comprehension, above and beyond WRE across various measures, as it is possible that some conflicting findings to date have to do with the use of different reading comprehension measures, which are known to vary in cognitive demands (Cutting & Scarborough, 2006; Francis et al., 2005; Keenan et al., 2008). To this end, findings suggested that ORR *is* strongly related to, but not synonymous with, WRE. ORR accounted for a substantial amount of variance above and beyond the contribution of WRE, ranging from 7.9% to 28.2% of the variance across five different measures of reading comprehension. In contrast, after accounting for ORR, WRE accounted for a nonsignificant amount of unique variance (0 to 1.9% of the variance). The common variance in comprehension predicted by the two fluency types ranged from 2.5% to 35.3% of the variance; that there was overlapping variance between these two measures is not surprising, as clearly both types of fluency, as well as comprehension, share the requirement of at least a basic level of word recognition skill.

Overall, our findings that the relationship between WRE, ORR, and reading comprehension holds even when varying the format of a comprehension measure indicate that others' findings (e.g., Denton et al., 2011; Jenkins et al., 2003b) are not merely an effect of how reading comprehension was measured. At the same time, it is important to note that there was a substantial range in the amount of variance that the overall model (WRE and ORR) accounted for, as well as in the amount of unique variance that each measure of reading efficiency contributed. The fact that ORR accounted for as little as 7.9% of unique variance for the DAB but more than 28% of unique variance for the GMRT–4 Passage Comprehension indicates that, as suggested by the findings of us as well as others (e.g., Cutting & Scarborough, 2006; Francis et al., 2005; Keenan et al., 2008), different measures of comprehension place varying levels of demands of cognitive skills. One possible explanation is that the underlying skills that distinguish ORR from WRE (such as oral language) may not be as essential when comprehension is assessed in particular formats. This suggests that poor ORR may selectively impede children's comprehension as measured by particular tasks; particularly notable was the fact that ORR consistently contributed substantially for timed tests with longer passages and multiple choice questions (similar to the formats of many high-stakes tests). This suggests that children with poor ORR will encounter difficulties with reading comprehension tasks that are most similar in format to the types of text they encounter in school.

If, as expected, ORR contributes additional variance to reading comprehension measures beyond WRE, do oral language skills explain the differences between the two?—Our findings also supported and extended previous suggestions that oral language is related to ORR. After accounting for rapid naming and basic word-level skills, oral language was found to be a predictor of ORR, whether ORR was measured by the GORT or SARA Passage Reading. Furthermore, oral language did *not* contribute to children's WRE. These results are consistent with previous results indicating a greater link between ORR and oral language than for WRE and oral language (e.g., Barth et al., 2009). However, although our findings support previous studies, it is important to note that the current study examined oral language at a more fine-grained level, including measures of *both* vocabulary and syntax. Although we had expected that both semantics and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

syntax would contribute to ORR, only measures of vocabulary made significant unique contributions. The finding that vocabulary, but not syntax, contributed to ORR highlights the importance of semantics in reading comprehension. Such findings are in line with the lexical quality hypothesis (Perfetti, 2007), which states that the quality of semantic representations (as well as phonological and orthographic representations) in word recognition is a critical component of reading comprehension, and in essence is an important link between word-level and comprehension abilities. The current findings suggest that ORR may be reflective of this role of semantics in reading comprehension.

It is interesting to note that our findings that vocabulary, but not syntax, contributed to reading in context could be viewed as somewhat reversed from findings with contextual word reading *accuracy* (not rate). For example, Ricketts, Nation, and Bishop (2007) found that oral vocabulary contributed to comprehension but not text reading accuracy, whereas Cain (2007) found that a measure of syntactic awareness predicted text reading accuracy but not comprehension. The differences between these previous findings and ours suggest that accurate text reading does not equate to ORR and, furthermore, that ORR may be more closely aligned with comprehension than reading accuracy, even when accuracy is being assessed in context.

Does the relationship between WRE and ORR vary depending on reader characteristics?—Overall, our findings suggested distinct patterns for different reader profiles. For average readers, ORR was faster than WRE, consistent with the well-known effect (especially for typical readers) of contextual facilitation. Nevertheless, for those with word-level deficits, poor performance for both WRE and ORR were shown across all measures, consistent with the supposition that poor WRE invariably leads to poor ORR. Of most interest, however, was understanding the WRE–ORR relationship in poor comprehenders, as prior studies have suggested that poor comprehenders do not benefit from contextual facilitation to the degree that their WRE skills would predict. In fact, as seen in Figure 1, the pattern for RCD is clearly one in which these readers show similar WRE to average readers but poorer ORR. Although this pattern is shown in the figure, it is also helpful to consider the patterns within the context of standardized age-normed tasks (TOWRE and GORT; see Table 5): if a child's WRE and ORR were comparable, there should be no significant differences between their standard scores on the two measures. Indeed, this was the case for average readers: Although their ORR was slightly higher than their WRE, the difference was not significant. In contrast, the poor comprehenders, as well as the poor decoders, both performed more poorly on the GORT when compared to their same-age peers than would be expected based on their TOWRE standard scores. This discrepancy between WRE and ORR is particularly interesting for the poor comprehenders, because their WRE was comparable to that of average readers. A comparable pattern on the experimental SARA tasks emerged, with poor comprehenders scoring similarly to average readers on WRE but not ORR. Taken together, these findings are in agreement with previous suggestions that poor comprehenders fail to benefit from contextual cues (Nation & Snowling, 1998a). Our hypothesis that was poor ORR in poor comprehenders may be due to their oral language skills, as studies examining oral language skills in relation to reading comprehension have found that children who are poor comprehenders exhibit deficits in oral language, including vocabulary (Cain & Oakhill, 2006; Cutting et al., 2009; Nation, Clarke, & Snowling, 2002; Oakhill & Yuill, 1996; Ricketts et al., 2007). Therefore, these oral language difficulties may be what prevent poor comprehenders from maintaining an average rate of word reading when words are presented in context. Indeed, results of additional analyses we conducted suggested that oral language does appear to be a contributing factor to ORR performance in this group: Consistent with previous studies, poor comprehenders demonstrated weaknesses in vocabulary as compared to average readers, even when general non-verbal ability was accounted for (Cain & Oakhill, 2006, 2011; Nation et al., 2002;

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Ricketts et al., 2007). Providing additional support to this supposition was that when covarying for vocabulary, the differences between the average readers' and poor comprehenders' ORR scores were no longer significant. These findings suggest that poor comprehenders' lower reading rates of continuous texts, as well as their comprehension difficulties, may be—in part—accounted for by their oral language skills. In summary, findings from the present study provide converging evidence that ORR contributes unique variance to reading comprehension beyond WRE, and that this unique variance, in part, can be explained by oral language abilities, specifically vocabulary/semantics.

Future Directions and Limitations

A strength of this study was the inclusion of multiple measurements for both single word and contextual fluency. The consistent pattern of our results across both experimental and standardized assessments suggests that the findings were not due to the design of a particular measure of WRE, ORR, or reading comprehension. Specifically, the SARA Word Recognition and SARA Passage reading share similar format and keywords. In contrast, the TOWRE and the GORT aren't designed to be overlapping in this manner. However, similar results were found between SARA Word Recognition-Passage Comprehension versus TOWRE–GORT analyses, suggesting that the patterns found with regard to the WRE–ORR relationship cannot be attributed to overlap/nonoverlap words. At the same time, it is important to consider differences among the tests. For example, the SARA Passages requires all children to read the same two passages, regardless of grade level or age. In contrast, the GORT, designed to adapt to ability level, starts out at a given story based on grade level, and then progresses to easier or more difficult passages based on performance.

Another difference between these two tests of ORR is what counts as an error. The SARA Passages only count misread or skipped words as errors, whereas the GORT errors include insertions and repetitions in addition to misread words and omissions. Different error types may be indicative of distinct deficits, whereas misreading a word may be connected to poor decoding, errors such as inserting unrelated words could be related to issues such as poor semantic abilities. Thus, although patterns were similar across all of our findings, some of the variability in results between the standardized and experimental measures of ORR may be related to differences in the test formats. Related to this, it is also interesting to consider Walczyk et al.'s (2007) suggestion that some errors, as well as corresponding compensatory devices, may be more disruptive in comprehension than others. Future studies incorporating qualitative analyses of errors, particularly for poor comprehenders, may shed light on the sources of different types of errors, as well as whether different errors have the same impact on comprehension of a passage.

It is important to mention that the current study relied solely on oral reading response measures of efficiency/fluency and included no measures of silent reading rate/efficiency (e.g., Woodcock–Johnson Reading Fluency subtest), nor of prosody or expressiveness. Research investigating the relations between oral and silent reading measures and how task/construct differences are related to comprehension have yielded mixed results; some studies suggest that oral reading may be a stronger predictor of comprehension than silent reading (Berninger et al., 2010; Fuchs, Fuchs, Eaton, & Hamlett, as cited in Fuchs et al., 2001), whereas others found that after controlling for reading ability, comprehension did not vary between oral and silent reading conditions (McCallum, Sharp, Bell, & George, 2004). It is possible that the relationship between comprehension and oral versus silent reading may vary depending upon reading level (Berninger et al., 2010; Miller & Smith, 1990), but a task/response effect cannot be ruled out. Although silent reading fluency may be more of a challenge to assess, most reading in the "real world" is done silently, and understanding the relation between silent reading fluency and comprehension may have important applicable implications, particularly for older readers with comprehension deficits. Also, in future

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

studies, prosody, or reading with appropriate prosody and expression (e.g., pauses, pitch, phrasing) needs to be considered; adding prosody as a variable has shown to contribute additional variance to reading comprehension in some studies (Klauda & Guthrie, 2008), but the robustness of this effect needs further study.

Finally, there are likely several possibilities regarding the nature of the directional relationship between ORR, language, and reading comprehension. Jenkins et al. (2003b) found that the relationship between ORR and comprehension was reciprocal; comprehension was a unique predictor of ORR, just as ORR was a predictor of comprehension. It could be that readers who easily comprehend a text (potentially driven by language skills) and are therefore able to read quickly. Alternatively, language may facilitate efficient word recognition, which in turn frees up resources for deeper comprehension. In either case, it is difficult to determine causality or directionality in the absence of longitudinal or experimental studies. Regressing comprehension on WRE, ORR, and language for each group of readers may clarify these relationships to some extent; however, our sample size in the present study prohibited such analyses. To this end, although our study has many strengths, it is important to note the limitations, one of which certainly was sample size, particularly for analyses that categorized readers into groups. Furthermore, in future studies it also may be helpful to include expanded measures of fluency (i.e., silent and oral, as well as prosody measures) as well as more fine-tuned measures of oral language, which may further clarify the relationship between WRE and ORR.

Conclusion

The results of this study provide further evidence that WRE reading and ORR are related, but not synonymous, and, more important, contribute greater clarity into the nature of the WRE–ORR relationship. Although there is no question that WRE is a strong predictor of ORR, ORR's contribution of unique variance above and beyond WRE to reading comprehension was robust and was upheld across five reading comprehension measures with varying formats. Such findings indicate that previous conflicting findings with regard to the WRE–ORR relationship are likely not due to format of reading comprehension test. Instead, we showed that oral language skills, particularly vocabulary/semantics, contributed ORR performance; these findings were further supported by results that revealed that poor comprehenders' weaknesses in ORR could be accounted for by their semantic knowledge. Our findings also provide some clarity as to why there may be conflicting findings in the literature with regard to WRE and ORR: Although most children with poor word recognition will do poorly on *both* WRE and ORR, not all children demonstrating poor ORR exhibit decoding or WRE deficits. Instead, some children struggle with ORR, as well as reading comprehension, despite having sufficient WRE. From a practical standpoint, this suggests that using measures of ORR taps into constructs that capture both word recognition and higher level comprehension skills (Fuchs et al., 2001). Therefore, a child who performs poorly on an ORR test may or may not need word-level intervention; only by administering a test of WRE would the focused of the needed intervention be clarified, that is, a child with adequate WRE but poor ORR may need an intervention focused on bolstering semantic representations, versus a pure decoding intervention per se. Further refinement of WRE and ORR measures may be valuable and allow for better identification and treatment of children who struggle with reading comprehension.

References

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

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: Author; 1994.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

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

Berninger VW, Abbott RD, Trivedi P, Olson E, Gould L, Hiramatsu S, Westhaggen SY. Applying the multiple dimensions of reading fluency to assessment and instruction. Journal of Psychoeducational Assessment. 2010; 28:3–18.

Booth JR, MacWhinney B, Harasaki Y. Developmental differences in visual and auditory processing of complex sentences. Child Development. 2000; 71:979–1001.

Cain K. Syntactic awareness and reading ability: Is there evidence for a special relationship? Applied Psycholinguistics. 2007; 28:679–694.

Cain K, Oakhill J. Profiles of children with specific reading difficulties. British Journal of Educational Psychology. 2006; 76:683–696. [PubMed: 17094880]

Cain K, Oakhill J. Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. Journal of Learning Disabilities. 2011; 44:431–443. [PubMed: 21772058]

Conners, CK. Conners' Parent Rating Scale–Revised (L). North Tonawanda, NY: Multi-Health Systems; 1997.

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

Cutting LE, Scarborough HS. Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. Scientific Studies of Reading. 2006; 10:277–299.

Daane, MC.; Campbell, JR.; Grigg, WS.; Goodman, MJ.; Oranje, A. Fourth-grade students reading aloud: NAEP 2002 special study of oral reading (No NCES 2006–469). Washington, DC: U. S. Department of Education. Institution of Education Sciences, National Center for Educational Statistics; 2005.

Denton CA, Barth AE, Fletcher JM, Wexler J, Vaughn S, Cirino PT, Francis DJ. 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. 2011; 15:109–135. [PubMed: 21637727]

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test–Third edition. Circle Pines, MN: American Guidance Service; 1997.

DuPaul, GJ.; Power, TJ.; Anastopoulos, AD.; Reid, R. ADHD Rating Scale–IV. New York, NY: Guilford; 1998.

Francis, DJ.; Fletcher, JM.; Catts, HW.; Tomblin, JB. Dimensions affecting the assessment of reading comprehension. In: Paris, SG.; Stahl, SA., editors. Children's reading comprehension and assessment. Mahwah, NJ: Erlbaum; 2005. p. 369-394.

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

Hoover WA, Gough PB. The simple view of reading. Reading and Writing. 1990; 2:127–160.

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

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

Karlsen, B.; Gardner, EF. Stanford Diagnostic Reading Test–Fourth edition. San Antonio, TX: Psychological Corporation; 1995.

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

Klauda SL, Guthrie JT. Relationships of three components of reading fluency to reading comprehension. Journal of Educational Psychology. 2008; 100:310–321.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Kuhn MR, Schwanenflugel PJ, Meisinger EB. Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. Reading Research Quarterly. 2010; 45:230–251.

LaBerge D, Samuels SJ. Toward a theory of automatic information processing in reading. Cognitive Psychology. 1974; 62:293–323.

MacGinitie, W.; MacGinitie, R.; Maria, K.; Dreyer, L. Gates–MacGinitie Reading Tests–Fourth edition. Itasca, IL: Riverside; 2002.

McCallum RS, Sharp S, Bell SM, George T. Silent versus oral reading comprehension and efficiency. Psychology in the Schools. 2004; 41:241–246.

Miller SD, Smith DE. Relations among oral reading, silent reading and listening comprehension of students at differing competency levels. Reading Research and Instruction. 1990; 29:73–84.

Nation K, Clarke P, Snowling MJ. General cognitive ability in children with reading comprehension difficulties. British Journal of Educational Psychology. 2002; 72:549–560. [PubMed: 12495566]

Nation K, Snowling MJ. Individual differences in contextual facilitation: Evidence from dyslexia and poor reading comprehension. Child Development. 1998a; 69:996–1011. [PubMed: 9768483]

Nation K, Snowling MJ. Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. Journal of Memory and Language. 1998b; 39:85–101.

Nation K, Snowling MJ. Factors influencing syntactic awareness skills in normal readers and poor comprehenders. Applied Psycholinguistics. 2000; 21:229–241.

Newcomer, PL. Diagnostic Achievement Battery–Third edition. Columbia, MO: Hawthorne Educational Services; 2001.

Newcomer, PL.; Hammill, DD. Test of Language Development–Intermediate: Third edition. San Antonio, TX: Pro-Ed; 1997.

Oakhill, J.; Yuill, NM. Higher order factors in comprehension disability: Processes and remediation. In: Cornoldi, C.; Oakhill, J., editors. Reading comprehension difficulties: Processes and intervention. Mahwah, NJ: Erlbaum; 1996. p. 137-166.

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

Perfetti CA. The universal grammar of reading. Scientific Studies of Reading. 2003; 7:3–24.

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

Perfetti CA, Hogaboam T. Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology. 1975; 67:461–469.

Ricketts J, Nation K, Bishop DVM. Vocabulary is important for some, but not all reading skills. Scientific Studies of Reading. 2007; 11:235–257.

Sabatini, JP. Unpublished doctoral dissertation. University of Delaware; Newark: 1997. Is accuracy enough? The cognitive implications of speed of response in adult reading disability.

Sabatini, JP. Changing the culture of testing with technology: Computerized assessments of basic reading skills in adults and more. In: Hopey, CH., editor. Adult learning and basic technology skills. Columbus, OH: ERIC Clearinghouse on Adult, Career, and Vocational Education; 1998. p. 71-79.

Sabatini JP. Literacylink: Helping adults build their personal success. Technologia. 2000; 2(5):19–22.

Sabatini, JP.; Venezky, RL.; Kharik, P.; Jain, R. Cognitive assessments for low literate adults: An analytic review and new framework (Tech Rep TR00–01). Philadelphia, PA: University of Pennsylvania, National Center on Adult Literacy; 2000.

Samuels, SJ. Toward a model of reading fluency. In: Samuels, SJ.; Farstrup, AE., editors. What research has to say about fluency instruction. Newark, DE: International Reading Association; 2006. p. 24-46.

Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. Developmental Psychology. 2002; 38:934–947. [PubMed: 12428705]

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

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Wagner, RK.; Torgesen, JK.; Rashotte, C. Comprehensive Test of Phonological Processing. Austin, TX: Pro-Ed; 1999.

Walczyk JJ, Wei M, Griffith-Ross DA, Goubert SE, Cooper AL, Zha P. Development of the interplay between automatic processes and cognitive resources in reading. Journal of Educational Psychology. 2007; 99:867–887.

Wechsler, DL. Wechsler Intelligence Scale for Children–Fourth edition. San Antonio, TX: Psychological Corporation; 2003.

Wiederholt, JL.; Bryant, BR. Examiner's manual: Gray Oral Reading Test–Fourth edition. Austin, TX: Pro-Ed; 2000.

Wolf M, Katzir-Cohen T. Reading fluency and its intervention. Scientific Studies of Reading. 2001; 5:211–239.

Woodcock, RW. Woodcock Reading Mastery Test–Revised/Normative Update. Circle Pines, MN: American Guidance Service; 1998.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

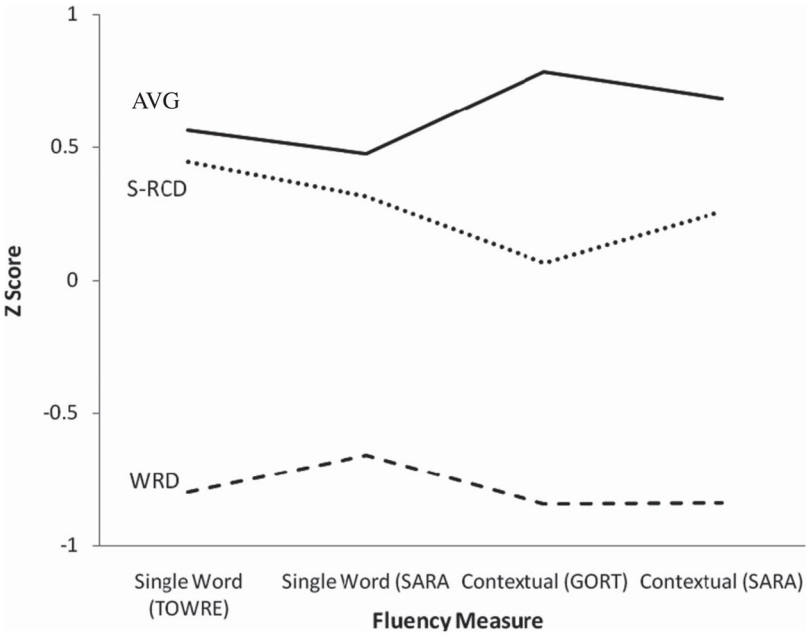

FIGURE 1.

Group comparisons on measures of single word and contextual fluency. *Note.* Within-sample *z* scores demonstrate the different patterns between isolated and contextual fluency measures for different reader profiles. (For the TOWRE and GORT analysis, standardized scores, not within-sample *z* scores were used; however, as expected, results were the same when the analysis used within-sample *z* scores.) TOWRE = Test of Word Reading Efficiency, Standard Scores; GORT = Gray Oral Reading Test Fourth Edition; AVG = average readers; WRD = poor decoders; S-RCD = poor comprehenders.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

TABLE 1

Demographic and Screening Measures by Groups

	Groups															
	Overall (n = 88)			AVG (n = 36)			WRD (n = 34)			S-RCD (n = 18)				Group Contrasts		
Variable	**M**	**SD**	**n (%)**	**M**	**SD**	**n (%)**	**M**	**SD**	**n (%)**	**M**	**SD**	**n (%)**	χ^2	**AVG v. WRD d**	**AVG v. S-RCD d**	**WRD v. S-RCD d**
Male			52 (59)			22 (61)			21 (62)			9 (50)	0.78			
ADHD			21 (24)			7 (19)			10 (29)			4 (22)	0.99			
Age	11.78	1.31		11.59	1.26		12.00	1.45		11.74	1.14			−0.30	−0.12	0.20
WISC–IV MR	10.25	2.85		12.08	2.09		9.18	2.49		8.56	2.83			1.26***	1.14***	0.23
WRMT BS	97.30	12.23		107.03	8.33		84.62	5.11		101.78	5.36			3.24***	0.75**	−3.28***
WRMT WI	96.66	11.52		106.25	6.67		84.68	5.79		100.11	5.29			3.45***	1.02**	−2.78***
WRMT WA	97.90	11.57		105.97	8.81		86.21	4.97		103.83	5.31			2.76***	0.29	−3.43***
WRMT PC	98.02	12.42		108.72	7.96		87.44	9.10		96.61	5.45			2.49***	1.78***	−1.22***
SDRT–4 PR	40.78	30.69		68.14	22.72		20.76	18.26		23.89	20.53			2.30***	2.04***	−0.16
GMRT–4 PR	42.93	36.09		77.22	19.51		16.33	25.14		23.11	18.36			2.71***	2.86***	−0.31
DAB ScS	8.17	3.03		10.19	2.67		6.94	2.35		6.44	2.60			1.29***	1.42***	0.20
GORT–4 ScS	10.30	2.81		12.31	2.23		9.03	1.85		8.67	2.97			1.60***	1.39***	0.15

Note. AVG = Average Readers; WRD = Poor Word Readers; S-RCD = Poor Comprehenders; WISC–IV MR = Wechsler Intelligence Scale for Children, Fourth Edition, Matrix Reasoning, Scaled Scores; WRMT = Woodcock Reading Mastery Test, Standard Scores; BS = Basic Skills Composite; WI = Word Identification; WA = Word Attack; PC = Passage Comprehension; SDRT–4 PR = Stanford Diagnostic Reading Test, Fourth Edition, Percentile; GMRT–4 PR = Gates MacGinitie Reading Tests, Percentile; DAB ScS = Diagnostic Achievement Battery, Reading Comprehension Scaled Score; GORT–4 ScS = Gray Oral Reading Test, Fourth Edition, Comprehension Scaled Score.

** $p < .01$.

*** $p < .001$.

Sci Stud Read. Author manuscript; available in PMC 2014 May 01.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

TABLE 2

Summary of Correlations, Means, and Standard Deviations for Scores on Reading Comprehension, Word Reading Efficiency, and Oral Reading Rate Measures

Variable	1	2	3	4	5	6	7	8	9
1. GMRT–4	—								
2. SDRT–4	.79**	—							
3. WRMT PC	.73**	.77**	—						
4. DAB	.57**	.49**	.46**	—					
5. GORT–4 Comp.	.50**	.52**	.53**	.58**	—				
6. TOWRE SWE	.56**	.49**	.61**	.22*	.25*	—			
7. GORT–4 Fluency	.74**	.67**	.77**	.47**	.53**	.71**	—		
8. SARA Words	.46**	.46**	.58**	.27*	.28**	.56**	.57**	—	
9. SARA Passages	.71**	.66**	.74**	.38**	.43*	.79**	.82**	.63**	—

Note. GMRT–4 = Gates MacGinitie Reading Test; SDRT–4 = Stanford Diagnostic Reading Test, Fourth Edition; WRMT PC = Woodcock Reading Mastery Test, Passage Comprehension; DAB = Diagnostic Achievement Battery; GORT–4 = Gray Oral Reading Test, Fourth Edition; TOWRE SWE = Test of Word Reading Efficiency, Standard Scores; SARA Words = SARA Word Recognition, Words Correct per Minute (partialing out age); SARA Passages = SARA Passage Reading, Bed Story Words Correct per Minute (partialing out age).

* $p < .05$.

** $p < .01$.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

TABLE 3

Summary of Hierarchical Regression Analyses Predicting Reading Comprehension From Word Reading Efficiency and Oral Reading Rate

	Measure of Reading Comprehension									
	GMRT–4		SDRT–4		WRMT PC		DAB		GORT–4	
Regression and Steps	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**
Std. measures										
1. TOWRE SWE	.311***	.558***	.243***	.493***	.372***	.610***	.049*	.221*	.062*	.248*
2. GORT Fluency	.237***	.688***	.212***	.650***	.231***	.678***	.190***	.614***	.246***	.699***
1. GORT Fluency	.545***	.739***	.454***	.674***	.595***	.771***	.216***	.465***	.277***	.527***
2. TOWRE SWE	.003	.071	.001	.035	.009	.132	.023	−.212	.030[a]	−.244
Total R^2	.548***		.455***		.603***		.238***		.307***	
Overall F	50.92		35.49		64.68		13.31		18.86	
Shared variance	.308		.242		.363		.025		.031	
Exp. measures										
1. Age	.017	.129	.004	−.064	.115**	.339**	.007	.085	.028	.168
2. SARA Words	.209***	.462***	.204***	.456***	.290***	.544***	.071*	.269*	.077**	.280**
3. SARA Passages	.282***	.731***	.219***	.644***	.201***	.616***	.079**	.386**	.104**	.444**
2. SARA Passages	.491***	.748***	.420***	.690***	.472***	.732***	.149***	.411***	.181***	.453***
3. SARA Words	.000	.025	.003	.070	.018	.175	.001	.038	.000	.014
Total R^2	.508***		.427***		.605***		.157**		.209***	
Overall F	28.54		20.86		42.90		5.21		7.40	
Shared variance	.209		.201		.271		.070		.077	

Note. GMRT–4 = Gates MacGinitie Reading Test; SDRT–4 = Stanford Diagnostic Reading Test, Fourth Edition; WRMT PC = Woodcock Reading Mastery Test, Passage Comprehension; DAB = Diagnostic Achievement Battery; GORT–4 = Gray Oral Reading Test, Fourth Edition; Std. = standardized; TOWRE SWE = Test of Word Reading Efficiency, Standard Scores; Exp. = experimental; SARA Words = SARA Word Recognition, Words Correct per Minute; SARA Passages = SARA Passage Reading, Bed Story Words Correct per Minute.

*$p < .05$.

**$p < .01$.

***$p < .001$.

[a]$p = .058$.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

TABLE 4

Summary of Hierarchical Regression Analyses for Predicting Oral Reading Rate

	GORT Fluency					SARA Passages				
Regression and Steps	R^2	ΔR^2	F	ΔF	β	R^2	ΔR^2	F	ΔF	β
1. Age	.709	.709	47.62	47.62	.027	.745	.745	56.88	56.88	.051
WRMT LW ID					.790***					.788***
WRMT WA					−.072					−.086
CTOPP RLN					−.177*					−.214**
2. + Semantics	.765	.056	41.30	9.03		.771	.026	42.60	4.33	
WISC Vocab					.285**					.281**
PPVT					.040					−.116
3. +Syntax	.766	.001	30.29	0.13		.772	.001	31.32	0.19	
TOLD GC					.037					.047
Sent Comp					−.018					−.003
2. +Syntax	.728	.018	33.85	2.55		.751	.006	38.22	0.97	
TOLD GC					.144					.090
Sent Comp					.032					.009
3. + Semantics	.766	.038	30.29	6.07		.772	.021	31.32	3.40	
WISC Vocab					.277**					.267*
PPVT					.034					−.129

Note. GORT = Gray Oral Reading Test Fourth Edition, Fluency Scaled Score; WRMT = Woodcock Reading Mastery Test, Standard Scores; LW ID = Letter–Word Identification; WA = Word Attack; CTOPP RLN = Comprehensive Test of Phonological Processing, Rapid Letter Naming Scaled Score; WISC Vocab = Wechsler Intelligence Scale for Children, Vocabulary Scaled Score; PPVT = Peabody Picture Vocabulary Test, Third Edition, Standard Score; TOLD GC = Test of Language Development, Grammatic Comprehension, Scaled Score; Sent Comp = Sentence Comprehension Task, Items Correct.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 5

Mean Standard Scores for Measures of Word Reading Efficiency and Oral Reading Rate as a Function of Reading Group

Group	Word Reading Efficiency (TOWRE)	Oral Reading Rate (GORT)	Within-Group Contrasts[a] (Cohen's d)
AVG[b]	103.31 (10.34)	105.97 (13.14)	−0.22
WRD[c]	86.12 (6.92)	67.39 (10.87)	1.70***
S-RCD[d]	100.00 (8.82)	88.56 (11.83)	1.02***
Between-group contrasts[a] (Cohen's *d*)			
AVG vs. WRD	1.73***	2.91***	
AVG vs. S-RCD	0.26	1.27***	
WRD vs. S-RCD	−1.51***	−1.68***	

Note. TOWRE = Test of Word Reading Efficiency, Standard Scores; GORT = Gray Oral Reading Test Fourth Edition, Scaled Score converted to Standard Score; AVG = average readers; WRD = poor decoders; S-RCD = poor comprehenders.

[a]Contrasts covarying for Matrix Reasoning, scaled score.

[b]$n = 36$.

[c]$n = 33$.

[d]$n = 18$.

***$p < .001$.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

TABLE 6

Mean Words Correct per Minute for Measures of Word Reading Efficiency and Oral Reading Rate as a Function of Reading Group

Group	Word Reading Efficiency (SARA Word Recognition)	Oral Reading Rate (SARA Passage Reading)	Within-Group Contrasts[a] (Cohen's d)
AVG[b]	75.94 (14.75)	154.60 (23.46)	−3.64***
WRD[c]	52.39 (19.59)	100.41 (32.70)	−2.33***
S-RCD[d]	70.02 (14.66)	132.83 (24.12)	−3.22***
Between-group contrasts[a] (Cohen's *d*)			
AVG vs. WRD	1.28***	2.28***	
AVG vs. S-RCD	0.18	0.71*	
WRD vs. S-RCD	−1.13***	−1.63***	

Note. AVG = average readers; WRD = poor decoders; S-RCD = poor comprehenders.

[a]Contrasts covarying for Age and Matrix Reasoning, raw score.

[b]$n = 36$.

[c]$n = 33$.

[d]$n = 18$.

*$p < .05$.

***$p < .001$.