School Psychology Review,
2013, Volume 42, No. 3, pp. 243–261

Examining Variance in Reading Comprehension Among Developing Readers: Words in Context (Curriculum-Based Measurement in Reading) Versus Words Out of Context (Word Lists)

Scott P. Ardoin
University of Georgia

Tanya L. Eckert
Syracuse University

Theodore J. Christ and Mary Jane White
University of Minnesota

Laura S. Morena, Stacy-Ann A. January, and Jeffrey F. Hine
University of Georgia

Abstract. Curriculum-based measurement in reading (CBM-R) is a widely used measure for identifying students in need of reading intervention and monitoring their progress. Despite a large base of research supporting the efficacy of CBM-R as a measure of comprehension, critics maintain that CBM-R is little more than a measure of word reading. The current study extends existing research by examining the extent to which CBM-R is a measure of student performance beyond word reading rate and accuracy as well as examining the benefits of administering a word list to early elementary students. Participants included 143 first-grade and 147 second-grade students. All students were administered a norm-referenced measure of reading achievement, two CBM-R passages, and two word lists. Results replicate and extend the literature by highlighting potential differences in the reading behavior of developing and fluent readers and demonstrating the potential need to administer measures of word list reading when conducting screenings in early grades.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A090038 to the University of Minnesota. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Theodore J. Christ received consulting income from Pearson AIMSweb, a company that may commercially benefit from the results of this research. This relationship was reviewed and managed by the University of Minnesota in accordance with its Conflict of Interest policies. At the time of this study, Dr. Christ was also the director of the Formative Assessment System for Teachers (FAST). Dr. Christ previously received consulting income related to commercial products from McGraw Hill, Renaissance Learning, and the University of Oregon.

Correspondence regarding this article should be addressed to Scott P. Ardoin, University of Georgia, Department of Educational Psychology, 323 Aderhold Hall, Athens, GA 30602; e-mail: spardoin@uga.edu

Perfetti's (1985) verbal efficiency theory suggests that the demands placed on memory and attention for reading words are decreased as one becomes more proficient at 12 reading words and that when these demands are decreased, greater attention can be given to the meaning of a passage. This theory is well supported and similarly postulated by others (e.g., LaBerge & Samuels, 1974), with studies indicating that word reading and listening comprehension skills account for practically all of the variance in students' reading achievement (Tilstra, McMaster, van den Broek, Kendeou, & Rapp, 2009). Even when controlling for listening comprehension, word reading skills continue to explain significant variance in students' reading comprehension (Tilstra et al., 2009).

Deno (1985) developed Curriculum-Based Measurement of oral reading (CBM-R) based largely upon Perfetti's (1985) verbal efficiency theory and the desire to develop a simple and quick assessment procedure to monitor student progress. CBM-R is a simple yet well-researched procedure that involves having students read passages for 1 min as an examiner records the number of words read correctly (WRCM). There is now an extensive body of research supporting CBM-R and consequently the position by many that reading fluency frees cognitive resources for comprehension (LaBerge & Samuels, 1974; Perfetti, 1985), with studies repeatedly demonstrating a strong relationship between students' WRCM as measured using CBM-R procedures and various norm-referenced measures of reading achievement (Ardoin et al., 2004; Reschly, Busch, Betts, Deno, & Long, 2009).

Meta-analyses by Reschly et al. (2009) indicated a strong relationship between CBM-R and various components of reading achievement. With the exception of word identification, which was found to have the strongest relationship with CBM-R, the strength of the relationship among CBM-R and other components of reading achievement (i.e., vocabulary, comprehension, decoding) did not differ. The nonsignificant difference in the strength of the relationship between CBM-R and various measures of reading achievement suggests that CBM-R is a general outcome measure of reading. However, critics of CBM-R might disagree that it is a general outcome measure because the stronger correlation between CBM-R and word identification measures suggests that CBM-R is simply a measure of students' word identification speed. Moreover, the relationship between CBM-R and measures of reading achievement tends to decrease: (a) beyond fourth grade, when fewer students use decoding skills to read text (Jenkins & Jewell, 1993; Kranzler, Miller, & Jordan, 1999); and (b) when the passages used to assess student performance are easy and thus do not require students to employ their decoding skills (Spear-Swerling, 2006).

Words in Context: Contribution of CBM-R

In attempting to dispel the critics suggesting CBM-R is simply a measure of decoding, Hamilton and Shinn (2003) evaluated differences in rate and comprehension between teacher-identified word callers and students identified by teachers as nonword callers with similar word reading rates to the word callers. All students' skills were assessed using CBM-R, the Maze procedure (January & Ardoin, 2012), a measure of comprehension requiring students to read a passage and respond to comprehension questions, and the comprehension subtest of an individually administered norm-referenced achievement test. Results indicated that the students identified as having better comprehension skills by the teachers demonstrated greater comprehension skills. However, in contrast to teachers' perceptions of the word callers and nonword callers having equivalent oral reading rates, results indicated that WRCM on the CBM-R measure was significantly lower for word callers than for children identified as having good comprehension skills. Although these results may be limited in that subjective judgments of teachers were used to identify students believed to have only word identification skills and poor comprehension skills, other studies have illustrated the benefits of CBM-R over

assessing students' word reading rate (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003a; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003b). Moreover, Compton and colleagues (Compton et al., 2005; Olinghouse, Lambert, & Compton, 2006) found that monitoring students' progress using CBM-R measures explained variance in student reading beyond that explained by various word list measures.

Outside of the CBM-R literature, there is further evidence suggesting that having students read words within context (i.e., CBM-R) measures more than their word identification skills (Cutting & Scarborough, 2006). For instance, research indicates that context assists students in reading words within passages at a significantly greater rate than when they read the same words outside of a passage (Jenkins et al., 2003a; Nicholson, 1991; Nicholson, Bailey, & McArthur, 1991). Jenkins et al. (2003a) explored these findings within the context of CBM-R. Fourth-grade students were administered the Iowa Test of Basic Skills (ITBS; 2001), a CBM-R passage (words within context), and a word list probe consisting of the same words presented in the CBM-R passage. Results revealed that students' reading of words within context (CBM-R passages) was significantly greater than their reading of words out of context (word list) and that both reading words in and out of context explained significant variance in students' reading comprehension as measured by the ITBS. Further analyses revealed that although there was considerable variance shared between CBM-R and word list reading rates, students' CBM-R performance explained variance above and beyond the reading of word lists. Students' word list performance did not, however, explain variance in comprehension above and beyond student CBM-R performance.

Further evidence that CBM-R is not simply a measure of word reading rate is provided by Jenkins et al. (2003b). As opposed to only conducting the typical analysis examining the degree to which variance in comprehension is explained by students' CBM-R performance, these researchers examined the variance in students' CBM-R performance (context rate) explained by the ITBS comprehension subtest and their performance on a word list. Results revealed that although both ITBS reading comprehension and reading of words out of context (word list) explained significant variances in students' CBM-R performance, comprehension explained more than twice as much unique variance in CBM-R performance. The research described earlier suggested that the reading of words in context stimulates comprehension processes and students' WRCM, when administered CBM-R passages, was reflective of those processes. Thus, although CBM-R is not a direct measure of comprehension, reading with fluency requires prerequisite skills (e.g., phonics, sight vocabulary) needed for the successful comprehension of text (LaBerge & Samuels, 1974; Perfetti, 1985) and successful comprehension aids fluent reading (Jenkins et al., 2003a, 2003b). Given the variables that influence fluent reading, CBM-R is often referred to and employed as a general outcome measure.

Despite the strong body of research suggesting that CBM-R is a measure of more than word reading, it is important to recognize that much of this research was conducted with students in the third and fourth grades (Hamilton & Shinn, 2003; Jenkins et al. 2003a, 2003b). In early elementary school (first and second grade), students are primarily at the stage of learning to read in which they are building decoding and sight vocabulary skills that help them to become fluent readers. By middle to late elementary school, many if not most students already have these skills and, when engaged in reading within the classroom, are reading to learn as opposed to learning to read.

Given differences in the general purpose of reading among early and upper elementary students, there may also be differences in the extent to which their performance on CBM-R passages and word lists explains variance in reading comprehension. For example, students' performance when reading connected text (CBM-R passages) in first grade did not increase the prediction accuracy of students identified with a learning disability in second

grade (Compton et al., 2010). However, Compton and colleagues (2010) used a large battery of measures, including a measure of oral vocabulary, to predict learning disability identification in second grade, oversampled low-performing students, and the data collected to predict second-graders identified with a learning disability were collected in the fall of first grade. There is a strong possibility that many students, especially the lowest performing students, were receiving little instruction in connected text and thus reading connected text with minimal to no fluency. Thus, differences between findings of Compton, Fuchs, Fuchs, and Bryant (2006), Compton et al. (2010), and Jenkins et al. (2003a, 2003b) may be from differences in methodology employed across the two studies and/or differences in the developmental reading level of participants, but do indicate a need for further research with students in first and second grades.

Words Out of Context: Contribution of Word Lists

Despite the aforementioned evidence that performance on CBM-R passages may be a stronger predictor of reading comprehension than performance on word lists, studies within and outside of the CBM-R literature provide strong evidence that a measure of student word list reading accuracy and rate does provide meaningful information. Data derived from students' word list reading performance may result in valid decisions for both identifying young students struggling in developing reading skills as well as monitoring their progress (Compton et al., 2006; Compton et al., 2010). Clemens, Shapiro, and Thoemmes (2011) found that among four early reading measures (word identification fluency, letter naming fluency, phoneme segmentation fluency, and nonsense word fluency) administered in the beginning of first grade, word identification fluency most accurately classified those students with weak reading skills at the end of first grade.

Outside of the CBM-R research, empirical studies examining the Simple View of Reading (i.e., the idea that reading comprehension is largely a product of reading accuracy and listening comprehension; Hoover & Gough, 1990) suggest that assessing student performance on word lists is beneficial and that word identification skills even explain significant variance in comprehension above that of listening comprehension (e.g., Adolf, Catts, & Little, 2006). The extent to which word reading accuracy predicts comprehension does, however, decrease as students progress through school. In an examination of a slightly more complex model of the Simple View of Reading, Tilstra et al. (2009) found that fourth-, seventh-, and ninth-grade students' rate of reading words in context explained significant variance in reading comprehension after accounting for students' decoding skills, verbal proficiency, and listening comprehension skills.

Purpose

A substantial research base establishes that measures of students' oral reading rate, as assessed using CBM-R procedures, provide information beyond that of the accuracy and rate at which students read words (Hamilton & Shinn, 2003; Jenkins et al. 2003a, 2003b). However, the existing research base relied exclusively upon middle to upper elementary grade students (i.e., third and above). For example, participants in both studies conducted by Jenkins et al. (2003a, 2003b) were fourth-grade students whose reading skills and purpose for reading words within (CBM-R passages) and outside of texts (word list) likely differed considerably from first- and second-grade students who are in the early development of reading fluency. A primary purpose of this study was to evaluate whether CBM-R data reflect more than word reading rate and accuracy among early elementary students. We achieved this by (a) examining differences in students' speed of reading words in and out of context and (b) examining variance explained in reading comprehension by students' CBM-R and word list reading rates as well as the degree to which word list reading rate and reading comprehension explained variance in

students' CBM-R performance. Evidence that (a) CBM-R explains variance in reading comprehension after accounting for that explained by word list reading rates, and (b) evidence that reading comprehension explains variance in CBM-R after accounting for that explained by word list reading rates, would provide strong evidence that even among early elementary students CBM-R is measuring more than students' word identification skills.

A second purpose of the current study was to evaluate the utility of administering a word list reading rate measure to students in first and second grade. Currently, when conducting universal screenings, schools generally do not administer word list measures to students after the first semester of first grade (Good, 2004), at which time schools begin to use CBM-R as a means of assessing students' reading. There is, however, research supporting the use of word lists for explaining variance in students' reading comprehension beyond that which might be explained by students' CBM-R performance (Compton et al., 2005; Compton et al., 2010; Tilstra et al., 2009). Although the results of Jenkins et al. (2003a, 2003b) may suggest that word list reading rate may not be informative beyond CBM-R data, these results may be limited by the format of the word lists employed. The word lists used in both studies were developed by randomizing the words within the CBM-R passage administered to students. Although this procedure allowed for a more direct comparison of differences between word identification and oral reading rate, it opens the potential for practice as the words that made up the first task that students completed were also part of the second task they completed (i.e., word list or CBM-R passage). Furthermore, there may be external validity issues because word lists employed by schools are generally made up of high-frequency and/or decodable words. Use of words from within a passage introduces the possibility that the words that make up the word identification task are neither decodable nor words that should be part of students' sight word vocabulary.

In the current study, we examined students' performance on two types of word identification lists, with one made up of high-frequency words and the other consisting of decodable words. Although recent research suggests greater concurrent and predictive validity for students on high-frequency word lists as compared to nonsense word fluency (Fuchs, Fuchs, & Compton, 2004), no research has evaluated both high-frequency and decodable word lists. It is possible that these two types of probes differentially predict students' reading fluency and comprehension as they assess two different skills. Specifically, sight word lists containing high-frequency sight words do not require most students to use their phonetic skills because of repeated exposure. In contrast, decodable word lists require students to use their decoding skills much like a nonsense word fluency measure, but with real words. Given the extensive research demonstrating the importance of phonics to students' development of successful reading skills (National Institute of Child Health and Human Development, 2000), it is possible that word lists developed to evaluate phonemic skills will explain variance in students' reading comprehension and text reading not explained by a high-frequency word list. Thus, a third purpose of the current study was to extend research by examining differences in the extent to which high-frequency and decodable word list probe types explain variance in students' reading comprehension.

A final purpose was to conduct exploratory analyses to examine if results would generalize to extremely low-achieving students. Although we hypothesized that students' CBM-R passage would predict comprehension above and beyond their performance on word lists, word list performance may be a better predictor of comprehension for students with poor rates of reading CBM-R passages.

Method

Participants and Settings

Participants were first- (n = 143) and second-grade (n = 147) students enrolled in one of two elementary schools (School A and B) in the southeast region of the United States. The majority of first-grade participants were

male (58%) and parents identified the students' race/ethnicity as White (79%). A small percentage of parents reported their child's race/ethnicity as Hispanic or Latino (8%), Black or African American (7%), or nonidentified (6%). The mean age of the first-graders was 7.1 years (SD = 0.38, range = 6.5–8.3). A total of 5 first-grade students were eligible to receive special education services, and 8 students were designated as English language learners. Of the second-grade sample, the majority were male (62%) and self-identified their race as White (62%). In comparison to the first-grade sample, a larger percentage of second grade students self-identified their race/ethnicity as Hispanic or Latino (21%); however, a similar percentage self-identified their race as Black or African American (8%) or nonidentified (8%). The average age of the second grade students was 8.0 years (SD = 0.44, range = 6.9–10.0). Eight second grade students were eligible for special education services and 28 students were identified as English language learners.

The participants were enrolled in a total of 10 first-grade classrooms (School A = 4; School B = 6) and 9 second-grade classrooms (School A = 4; School B = 5). School A served 540 students in prekindergarten to fifth grade. A majority of students (61%) in School A received free or reduced-price lunch. School B served 450 students in kindergarten through fifth grade with 18.9% of the students receiving free or reduced-price lunch. All participating students were administered the ITBS within classroom and were administered all other measures in the hallway adjacent to their classroom.

Materials

All participating students were administered Form A of the ITBS, two CBM-R passages, a decodable sight word list (decodable SWL) and a high-frequency sight word list (high frequency SWL). Other measures were administered to students during the course of the study, but these measures were not relevant to the study and therefore will not be described.

ITBS. Classroom teachers administered the appropriate sections of the ITBS to the participants, which included Level 7 (first-grade students) and Level 8 (second-grade students). At the levels assessed, the ITBS comprehension subtest requires students to engage in a variety of reading tasks by responding to questions that reference (a) pictures which each tell a story, (b) individual sentences, and (c) short passages. The subtests were administered by teachers who followed scripted instructions provided by the test publishers. Those students absent on the day of teacher administration were individually administered the corresponding level by the authors.

In an attempt to compare the present study with results from other studies examining the shared and unique contributions of students' reading rate and word rate to reading comprehension (Jenkins et al., 2003a, 2003b), we transformed the Reading Comprehension (ITBS-RC) subtest into normal curve equivalents. This permitted the averaging of scores because normal curve equivalents are based on an equal-interval scale (Mertler, 2002). The ITBS comprehension subtests administered have a Kuder Richardson-20 internal consistency of .91 for first grade and .89 for second grade (University of Iowa, 2005).

CBM-R. Two CBM-R passages were employed for assessing students' oral reading rate in context. First-grade students received one preprimer passage, developed by the researchers, and a first-grade CBM-R passage selected from the EasyCBM series (http://easycbm.com/). Second-grade students received a first- and second-grade EasyCBM passage. Evidence exists to suggest that EasyCBM passages have adequate reliability and validity similar to other CBM-R passages (Jamgochian et al., 2010; Lai et al., 2010).

The passages differed in regard to the percentage of unique words presented in each passage as well as the percentage of unique words that were identified as Spache and/or Dolch first-grade sight words. Descriptive data regarding the composition of the passages are provided in Table 1. First-grade participants were administered the preprimer and first-

Table 1
Descriptive Data for CBM-R Passages

	No. of Words	No. (%) Unique Words*	No. (%) of unique words that are sight words
Preprimer CBM-R	258	88 (34%)	50 (57%)
First-grade CBM-R	247	119 (48%)	49 (41%)
Second-grade CBM-R	254	136 (55%)	58 (43%)

Note. CBM-R = curriculum-based measures of reading. Number of unique words was determined by deleting all repetitions of a word. Thus, the preprimer passages comprised only 88 unique words. Repetition of these words resulted in the passage being made up of a total of 258 words.

grade passages and second-grade participants were administered the first- and second-grade passages.

Decodable SWL. Both decodable SWLs consisted of 304 phonetically consistent real words selected from various word banks (e.g., *The Systematic Instructional Management Strategies Reading Program Manual*; Blackburn, Dickson, Millam, & Nelson, 1977). Words within the word banks were coded as falling within 1 of the 8 decodability criterion levels outlined by Menon and Hiebert (1999; see also Compton, Appleton, & Hosp, 2004). These eight decodability levels represent increasing word difficulty based on linguistic decoding patterns. The first three levels represent word constructions that are relatively easy to decode with increasing difficulty introduced at each subsequent level by including words with silent —*e,* double vowels, —*r* or —*l* endings, or diphthongs (Menon & Hiebert, 1999). Words for the first-grade decodability list were selected from existing word lists if they matched the linguistic patterns outlined in decodability Levels 1–5. Of the 304 words used in the second-grade word list, 159 words followed the linguistic patterns outlined in Levels 1–5, while the remaining 145 words followed the linguistic patterns outlined in Levels 6–8. The second-grade decodable word list consisted of words from all eight levels as opposed to only the first four levels so as to increase the level of difficulty of the second-grade probe as compared to the first-grade probe. We expected that increasing the difficulty would decrease the possibility of ceiling effects.

High-frequency SWL. The high-frequency SWLs were developed using words taken from the 315-item Dolch word list (Dolch, 1936; Johns, 1971) and from the New Instant word list (Fry, 1980). Both word lists use words that occur frequently in written English materials or were added as complementary supplements (Fry, 1957). For each grade level, 304 words were selected for use. This number assured that readers of advanced reading ability would have enough words to read in the time provided. Although both the first- and second-grade SWLs included a selection of words from the Dolch word list and the New Instant word list, a greater number of less frequent words from the New Instant word list were used in the second-grade list to reflect the change in grade level.

Procedure

Students were administered the Reading Comprehension portion of the ITBS during either the end of the 6th (School A) or beginning of the 7th month (School B) of the academic year. Standardized procedures provided in the ITBS examiners' manual were followed in the administration of the ITBS.

Within a week following ITBS administration, trained experimenters administered CBM-R and word list probes to students in a

counterbalanced order. First-grade students were administered the preprimer CBM-R passage, first-grade CBM-R passage, first-grade decodable SWL, and first-grade high-frequency SWL. Second-grade students were administered the first-grade CBM-R passage, second-grade CBM-R passage, second-grade decodable SWL, and the second-grade high-frequency SWL.

Prior to administration of any probe, examiners read scripted directions to each student. The directions informed students to do their best reading, that they would be reading multiple types of probes, and that some probes may contain nonsense words. Students were given 2 min to read each SWL, while examiners recorded errors and provided words on which students hesitated for 3 s. The SWL outcomes were divided by 2 to provide a direct comparison with the 1-min CBM-R outcomes.

Standard CBM-R administration and scoring procedures were used when students were presented with the CBM-R passages (Shinn, 1998). Students were allowed to read each passage for 1 min, examiners recorded words read incorrectly, and provided words to students on which they hesitated for 3 s. Words misread included words read incorrectly given their context, skipped words, and substitutions. Insertions and self-corrections made within 3 s were not counted as errors.

Interrater Agreement and Procedural Integrity

Examiners were undergraduate and graduate students trained in the scoring and administration procedures during an hour-long training session completed by the first author. In addition, prior to examiners administering procedures, they were required to observe one of the authors conduct an administration, then they conducted an administration while an author observed them, and finally they were provided with feedback regarding their administration. Examiners were required to attend to a procedural checklist, which included verbatim instructions to be read to students, to ensure procedural integrity. All sessions were audio recorded so that independent raters could assess interrater agreement and procedural integrity.

Procedural integrity was evaluated on 15% of sessions and was calculated by dividing the number of steps completed accurately by the total number of steps (40). The average procedural integrity was 98% (range = 83%–100%). Interrater agreement was assessed on 15% of sessions from each school. Interrater agreement was determined by dividing the number of words in agreement by the number of agreements plus disagreements. Overall, the average interrater agreement was 96.5% (range = 93%–99%).

Data Analysis

Data screening procedures revealed minimal problems with the distribution of the variables and no outliers were identified. Evaluation of missing data indicated that less than 1% of the sample (.6%; 12/2030) had missing data on one or more of the key study outcomes. Given the relatively few cases of missing data and that the missing data appeared to represent a random subsample of data, the missing cases were deleted in subsequent analyses. This appeared to be the most appropriate manner to address the missing data given that other procedures for handling missing values yields similar results if the missing data reflect 5% or less of the data set (Tabachnick & Fidell, 2007).

Hierarchical regression analyses were used to examine the degree of shared and unique variance associated with the predictors of reading comprehension and oral reading rate of words in context (i.e., average WRCM from CBM-R performances). In an attempt to compare the present study with results from other studies examining the shared and unique contributions of students' reading rate and word rate to reading comprehension (Jenkins et al., 2003a, 2003b), we conducted the regression analyses in pairs, reversing the order of the predictors so that the unique contribution of each predictor could be determined after controlling for the other predictor. For the present study, 12 a priori models were specified for each grade level. The order of predic-

Table 2
Means, Standard Deviations, and Intercorrelations Among Variables for First- (n = 143) and Second-Grade (n = 147) Students

	First Grade		Second Grade								
Variable	Mean	(*SD*)	Mean	(*SD*)	1	2	3	4	5	6	7
1. ITBS Reading Comprehension	64.04	(28.11)	58.74	(25.75)	—	.73	.75	.76	.68	.66	.68
2. CBM-R Lower Level[a]	67.05	(39.21)	100.87	(38.77)	.81	—	.88	.96	.80	.83	.81
3. CBM-R Upper Level[b]	58.27	(42.34)	93.30	(39.23)	.81	.95	—	.97	.87	.88	.88
4. Average CBM-R	63.13	(40.25)	97.10	(37.96)	.82	.99	.98	—	.86	.88	.88
5. High Frequency SWL	36.08	(25.98)	48.33	(24.72)	.82	.92	.91	.93	—	.91	.98
6. Decodable SWL	22.30	(18.87)	31.95	(20.60)	.73	.89	.90	.90	.93	—	.97
7. Average SWL	22.52	(12.88)	40.92	(20.59)	.80	.92	.92	.93	.99	.97	—

Note: ITBS = Iowa Test of Best Skills Reading Comprehension Normal Curve Equivalent; CBM-R = curriculum-based measurement for reading; SWL = sight word list; CBM-R scores are words read correct in 1 min. All word lists scores are words correct in 1 min. The intercorrelations among variables for the first-grade students are presented below the diagonal and the intercorrelations among variables for the second-grade students are presented above the diagonal. All intercorrelations were statistically significant, $p < .001$.
[a] CBM-R Lower Level = preprimer passage for first-grade students and first-grade passage for second-grade students.
[b] CBM-R Upper Level = first grade passages for first-grade students and second-grade passage for second-grade students.

tors was reversed for each analysis so that the unique contribution of reading comprehension and word reading rate in context could be determined, after controlling for the other predictors. In an attempt to control for Type I error across the 12 models, the Bonferroni multiple comparison method was applied, and an alpha level of .004 was used for all hierarchical regression analyses.

Results

Descriptive Statistics

Preliminary analyses were conducted to examine means, standard deviations, and intercorrelations among variables associated with the students in first and second grades (Table 2). Correlational analyses revealed statistically significant correlations on all outcome variables. First, the ITBS-RC NCE was highly correlated with first- and second-grade students' CBM-R and SWLs outcomes, with correlations ranging in magnitude from .66 to .82. Second, the CBM-R scores of the first- and second-grade students on varying grade-level material were highly correlated, with correlations ranging from .88 to .95. Students' average CBM-R outcomes were used throughout further analyses. Third, the SWLs outcomes (i.e., high-frequency SWL, decodable SWL) of the first- and second-grade students were highly correlated, with correlations being .93 and.91, respectively. Fourth, CBM-R outcomes on individual passages and SWLs outcomes were highly correlated, although the magnitude of correlations was slightly higher for the first-grade students (range = .89–.92) than for the second-grade students (range = 0.80–0.88). Finally, a similar pattern of results was observed when comparing average CBM-R and average SWL scores, with a slightly higher correlation observed for first-grade students (.93) than for second-grade students (.88). Statistically significant grade-level differences were observed among outcome measures, wherein the correlations among the first-grade sample were significantly greater than the second-grade sample for CBM-R and high-frequency SWL, $Z = 3.08$, $p < .001$. No statistically significant grade-level differences were observed between the first- and second-grade samples for

CBM-R and decodable SWL, $Z = 0.81$, $p = .21$.

Differences in Reading Rate Between Words Read In and Out of Context

Paired-samples *t* tests were conducted to evaluate the differences in reading rate between words read in and out of context among the first- and second-grade students. Across both grade levels, the results indicated a statistically significant difference between students' average CBM-R and high-frequency SWL for first, $t(143) = 17.08$, $p < .001$, and second-grade students, $t(146) = 24.02$, $p < .001$, with higher rates of reading observed for CBM-R (first-grade students, $M = 63.13$; second-grade students, $M = 97.10$) than high-frequency SWL (first-grade students, $M = 36.08$; second-grade students, $M = 25.98$). In addition, across both grade levels, the results indicated a statistically significant difference between students' average CBM-R and decodable SWL for first-graders, $t(143) = 19.67$, $p < .001$, and second-graders, $t(146) = 35.57$, $p < .001$, with higher rates of reading observed for CBM-R (first-grade students, $M = 63.13$; second-grade students, $M = 97.10$) than decodable SWL (first-grade students, $M = 22.30$; second-grade students, $M = 31.95$). Further, across both grade levels, the results indicated a statistically significant difference between students' average high-frequency SWL and decodable SWL for first-graders, $t(143) = 14.67$, $p < .001$, and second-graders, $t(146) = 14.91$, $p < .001$, with higher rates of reading observed for high-frequency SWL (first-grade students, $M = 36.08$; second-grade students, $M = 48.33$) than decodable SWL (first-grade students, $M = 22.30$; second-grade students, $M = 31.95$).

Word Reading Rate in Context and Word List Reading Rate as Constituents of Reading Comprehension Among First- and Second-Grade Students

Hierarchical regression analyses were conducted to evaluate the unique contribution of first-grade students' oral reading rate of words in context (CBM-R) and of word list reading rate (SWL) on their ITBS-RC scores. In these analyses, reading comprehension was predicted by using students' average CBM-R performance and outcomes from the two SWLs (Table 3). Although all of the predictors shared significant variance with reading comprehension (range = 53%–69%), CBM-R and high-frequency SWL scores accounted for more unique variance in reading comprehension. High-frequency SWL scores explained 14% of the variance in reading comprehension scores after controlling for decodable SWL scores. Likewise, CBM-R explained 14% of the variance in reading comprehension scores after controlling for decodable SWL scores. However, CBM-R explained only 3% of the variance in reading comprehension when controlling for high-frequency SWL.

Identical hierarchical regression analyses were conducted with the second-grade students. Similar to the first-grade results, all predictors shared significant variance with reading comprehension (range = 43%–60%). However, unlike the results for the first-grade students, neither SWL explained variance in students' reading comprehension above and beyond the other SWL. Conversely, CBM-R accounted for more unique variance in students' reading comprehension than SWL outcomes (Table 3). After controlling for high-frequency SWL, 13% of the variance in reading comprehension scores was explained by CBM-R. After controlling for decodable SWL in the next regression analysis, CBM-R explained 15% of the variance in students' reading comprehension. Among second-grade students, CBM-R accounted for more than 10 times the unique variance in reading comprehension than did the SWL outcomes.

Reading Comprehension and Word List Reading Rate as Constituents of Oral Reading Rate in Context (CBM-R) Among First- and Second-Grade Students

Additional hierarchical regression analyses were conducted to evaluate the unique contribution of reading comprehension and SWLs outcomes on first-grade students' oral reading rate in context (i.e., average CBM-R).

Table 3
Summary of Hierarchical Regression Analyses Using Students' Average CBM-R Performance and Decodable and High-Frequency SWL Performance to Predict Reading Comprehension Among First- and Second-Grade Students

	First-Grade Students		Second-Grade Students	
Regression and Steps	R^2	ΔR^2	R^2	ΔR^2
A. 1. Decodable SWL	.53*		.43*	
2. High-frequency SWL	.67*	.14*	.47*	.04*
B. 1. High-frequency SWL	.67*		.46*	
2. Decodable SWL	.67	.004	.47	.008
C. 1. Average CBM-R	.67*		.60*	
2. High-frequency SWL	.69*	.02*	.60	.001
D. 1. High-frequency SWL	.67*		.46*	
2. Average CBM-R	.69*	.03*	.60*	.13*
E. 1. Average CBM-R	.67*		.58*	
2. Decodable SWL	.67*	.001	.58	.001
F. 1. Decodable SWL	.53*		.43*	
2. Average CBM-R	.67*	.14*	.58*	.15*

Note. CBM-R = curriculum-based measurement of reading; SWL = sight word list.
* $p < .004$.

Average CBM-R was predicted by using scores on the ITBS-RC and SWLs outcomes (Table 4). Although all of the predictors shared significant variance with CBM-R (range = 67%–86%), SWLs outcomes accounted for more unique variance in students' CBM-R performance than did reading comprehension (1% and 6%, respectively). After controlling for reading comprehension, high-frequency SWL reading rate explained 20% of the variance in average CBM-R. Similarly, after controlling for reading comprehension, decodable SWL reading rate explained 20% of the variance in average CBM-R.

Among the second-grade students, hierarchical regression analyses identical to those described earlier were conducted. All three predictors shared significant variance with average CBM-R (range = 58%–78%). Specifically, similar to the first-grade results, after controlling for reading comprehension, both high-frequency and decodable SWL outcomes accounted for a large percentage of unique variance (21% and 26%, respectively; Table 4). However, reading comprehension accounted for slightly more unique variance (6%) than previously reported for the first-grade students.

Reading Performance Levels of Low-Achieving First-Grade Students

In a set of exploratory analyses, a subset of low-achieving first-grade students was identified ($n = 31$). These students demonstrated significant difficulties in their oral reading fluency, which was defined as achieving a percentile rank of 20 or less (i.e., average score of 1–20 WRC on CBM-R outcome) in comparison to the overall first-grade sample of students. By creating a clearly differentiated subset of first-grade students, we were able to examine whether these students demonstrated different patterns of reading performance from those previously reported. Only first-grade students were selected for this exploratory analysis; there were not a sufficient number of second-grade students who met the criterion.

Table 4
Summary of Hierarchical Regression Analyses Using Students' Performance on Decodable and High-Frequency SWL and Reading Comprehension Outcomes to Predict Oral Reading Rate in Context Among First- and Second-Grade Students

	First-Grade Students		Second-Grade Students	
Regression and Steps	R^2	ΔR^2	R^2	ΔR^2
A. 1. Decodable SWL	.81*		.78*	
2. High-frequency SWL	.87*	.06*	.80*	.02*
B. 1. High-frequency SWL	.86*		.74*	
2. Decodable SWL	.87*	.01*	.80*	.06*
C. 1. Reading comprehension	.67*		.59*	
2. High-frequency SWL	.87*	.20*	.80*	.21*
D. 1. High-frequency SWL	.86*		.74*	
2. Reading comprehension	.87*	.01*	.80*	.06*
E. 1. Reading comprehension	.67*		.58*	
2. Decodable SWL	.87*	.20*	.83*	.26*
F. 1. Decodable SWL	.81*		.77*	
2. Reading comprehension	.86*	.06*	.83*	.06*

Note. SWL = sight word list.
* $p < .004$.

Table 5 presents descriptive statistics on key outcome measures for these low-achieving students.

To further examine the reading performance differences of these first-grade students, hierarchical regression analyses were conducted to evaluate the unique contribution of average CBM-R and SWL outcomes on their reading comprehension performance as well as the unique contribution of reading comprehension and SWL on students' average CBM-R outcomes. The resulting findings are described in the following paragraphs.

Reading comprehension results. In the first analysis, reading comprehension was predicted by using the scores of the low-achieving students' SWL and average CBM-R (Table 6). Unlike our previous findings, only average CBM-R shared significant variance with reading comprehension. Analyses revealed that none of the three variables shared statistically significant unique variance with reading comprehension. Observed power for the nonsignificant findings ranged from .19 to .36, with an average observed power of .29.

Oral reading rate in context results. In the second analysis, the low-achieving first-grade students' average CBM-R was predicted by using the students' scores on the reading comprehension and word list reading rate outcomes (Table 7). Unlike the analyses predicting average CBM-R of first-grade students, only high-frequency SWL (69%) and ITBS Reading Comprehension (24%) shared significant variance with average CBM-R. In comparison to the analyses including first-grade students, there was also substantial difference in unique variance. Decodable SWL failed to explain significant variance beyond either high-frequency SWL or reading comprehension. High-frequency SWL, however, explained 64% and 48% of the unique variance in average CBM-R outcomes of low-achieving first-grade students after controlling for decod-

Table 5
Means and Standard Deviations of Variables for Low-Achieving First-Grade Students (n = 31)

Variable	*M*	(*SD*)
ITBS reading comprehension	31.47	(18.00)
Average CBM-R	13.98	(7.61)
High-frequency SWL	8.56	(4.70)
Decodable SWL	6.45	(6.59)

Note. ITBS = Iowa Test of Best Skills Reading Comprehension Normal Curve Equivalent; CBM-R = curriculum-based measurement for reading (scores are words read correct in 1 min). SWL = sight word list (all SWL scores are words correct in 1 min).

Table 6
Summary of Regression Analyses for Variables Predicting Reading Comprehension Among Low-Achieving First-Grade Students

Regression and Steps	R^2	ΔR^2
A. 1. Decodable SWL	.04	
2. High-frequency SWL	.18	.14
B. 1. High-frequency SWL	.17	
2. Decodable SWL	.18	.003
C. 1. Average CBM-R	.24*	
2. High-frequency SWL word rate	.24	<.001
D. 1. High-frequency SWL rate	.12	
2. Average CBM-R	.24	.13
E. 1. Average CBM-R	.24*	
2. Decodable SWL	.25	.006
F. 1. Decodable SWL	.04	
2. Average CBM-R	.25	.21

Note. SWL = sight word list; CBM-R = curriculum-based measurement of reading.
* $p < .004$.

able SWL and reading comprehension. Reading comprehension explained 21% of the unique variance in the average CBM-R of low-achieving first-grade students when controlling for their decodable SWL reading rate. Reading comprehension did not explain significant unique variance after controlling for high-frequency SWL. Observed power for the nonsignificant findings ranged from .20 to .39, with an average observed power of .31.

Discussion

Extensive research exists demonstrating a strong relationship between students' CBM-R performance as measured by WRCM and their comprehension of text. This relationship can be explained by Perfetti's (1985, 1992) verbal efficiency theory, which suggests that with greater proficiency/rate in reading words, students are able to attend more to the meaning of a passage. A primary purpose of this study was to evaluate among early readers the extent to which CBM-R reflects more than word reading rate. Specifically, the study extended research by Jenkins and colleagues (2003a, 2003b), which demonstrated not only that CBM-R explained greater variance in reading comprehension than students' reading of a word list, but also that students' (a) CBM-R rate was greater than their word reading rate on a word list; (b) word list reading rate did not explain variance above and beyond CBM-R; and (c) comprehension explained more variance, and twice as much unique variance in students' CBM-R performance than word list reading rates. The findings of the current study provide further support for Perfetti's verbal efficiency theory and further validate that CBM-R measures more than students' word reading skills. Results do not, however, exactly mimic those of previous research, which is likely a function of methodological differences coupled with the lower developmental reading level of the participants.

Similar to previous research (Jenkins et al., 2003a; Nicholson, 1991), participants' oral reading rate of words in context was significantly greater than their rate of reading words out of context. These results applied across grades for both high-frequency words and decodable words. Thus, students' reading rate on the high-frequency word lists, which consisted of words that were all likely part of their sight vocabulary, was slower than their reading rate

Table 7
Summary of Regression Analyses for Variables Predicting Oral Reading Rate Among Low-Achieving First-Grade Students

Regression and Steps	R^2	ΔR^2
A. 1. Decodable sight word list (SWL)	.05	
2. High-frequency SWL	.70*	.64*
B. 1. High-frequency SWL	.69*	
2. Decodable SWL	.70	.002
C. 1. Reading comprehension	.24	
2. High-frequency SWL	.72*	.48*
D. 1. High-frequency SWL	.69*	
2. Reading comprehension	.72	.03
E. 1. Reading comprehension	.24	
2. Decodable SWL	.26	.02
F. 1. Decodable SWL	.05	
2. Reading comprehension	.26	.21

Note. SWL = sight word list.
* $p < .004$.

when reading words presented in context despite the fact that some of the words in context were low-frequency and not decodable words (Table 1 provides percentages of words within passages that were sight words). In extending previous research, these data provide evidence that reading words in context facilitates students' word reading, suggesting that the administration of a CBM-R probe results in a score that represents more than students' reading of individual words (i.e., word lists).

Analyses of students' rate of reading words also indicated that students' rate of reading the high-frequency SWL was significantly greater than their rate of reading the decodable SWL. These data suggest that the high-frequency SWLs were at a minimum composed of more high-frequency words than the decodable SWL despite the fact that all of the words across the SWLs were in fact sight words. Moreover, not all sight words are the same and thus not all SWLs will produce equivalent results. In a review of the literature, Pratt, Martin, White, and Christ (2011) reported that definitions of sight words ranged from words that were recognized immediately, to words seen often within text, to words that did not require preparation by students to read. The variability in the types of SWLs that are available is a primary reason why two SWLs were administered in the current study, with one containing words that students were expected to know based upon frequency within text and another containing highly decodable words.

The fact that the two word lists resulted in differences in students' reading rates provides evidence that effort must be placed into the development of SWLs, if the desire is to develop multiple probes that are equivalent in level of difficulty and thus useful for universal screening and/or progress monitoring. Similar to recent evidence examining the equivalence of CBM-R probes within probe sets, evidence of the reliability and validity of probes is not sufficient to establish them as appropriate for progress monitoring (Ardoin & Christ, 2008; Ardoin & Christ, 2009; Ardoin, Christ, Morena, Cormier, & Klingbeil, 2013). Furthermore, the current data were consistent with previous research that found that readability formulas were not appropriate measures of CBM-R probe difficulty (Ardoin, Carfolite, Christ, Roof, & Klubnick, 2010; Ardoin, Suldo, Witt, Aldrich, & McDonald, 2005; Betts, Pickart, & Heistad, 2009), as many of these formulas treat all sight words as equivalent and evaluate probe difficulty based upon the number of words that are not sight words.

A second purpose of the current study was to examine the utility of administering word lists in addition to CBM-R with first- and second-graders. The hierarchical regression analyses closely resembled the findings from Jenkins et al. (2003a, 2003b) because CBM-R explained significant unique variance in reading comprehension after controlling for the decodable SWL and after controlling for the high-frequency SWL. Also in an identical fashion to Jenkins et al. (2003), neither SWL accounted for significant unique variance in reading comprehension for the second-graders, beyond that accounted for by CBM-R.

Although this was not the case for first-graders, the unique variance explained by the high-frequency SWL was relatively small as compared to the variance explained by CBM-R. Data from the second-graders more closely aligned with previous research in regard to the magnitude of variance explained by the second-graders' oral reading rate. Specifically, CBM-R accounted for 13% and 15% of unique variance in reading comprehension after accounting for high-frequency and decodable SWLs for second-grade students, whereas CBM-R explained only 3% and 14% of unique variance in reading comprehension for first-grade students. In regard to the magnitude by which CBM-R explained variance in reading comprehension above SWLs, results for both first- and second-grade students is substantially less than results reported by Jenkins et al. (2003b), in which the unique variance explained in fourth-grade students' reading comprehension by CBM-R after controlling for word list reading was 41%.

One potential explanation for CBM-R accounting for less variance in students' reading comprehension is that the less skilled a reader is, the less that context might affect oral reading rate of words in context versus word list reading. Perfetti's (1985, 1992) theory suggests that when students' attention is dedicated primarily to decoding words, their attention to the meaning of a story is diminished. If students are not attending to the meaning of a story, then it is unlikely that the context of the story will affect their reading rate to the same extent that it might for students who are able to devote attention to the meaning of a passage. Although there is debate regarding the extent to which students use context to read unknown words based upon their reading skills (Jenkins et al., 2003a, 2003b), with some suggesting that poor readers use context more than skilled readers (Nicholson, 1991; Zinar, 2000), it may be that some basic level of word reading rate and accuracy is required prior to students being able to use context. Thus, context may play less of a role and their sight word vocabulary may play a larger role. Further evidence suggesting that context may not affect the word reading rate for some readers is provided by the significantly stronger relationship between first-grade students' CBM-R and SWL outcomes as compared to second-grade students' CBM-R and SWL outcomes. The correlation between first-grade students' average CBM-R and the two SWLs ($r = .91$ and $r = .90$) was nearly identical to that of the relationship between the two SWLs.

Another interesting finding, which differs from that of previous research, was the amount of variance explained in students' CBM-R performance by their SWL performance and ITBS-RC (Table 4). Similar to Jenkins et al. (2003b), in the present study students' reading comprehension explained significant unique variance in their CBM-R performance. However, in contrast to Jenkins et al., the magnitude by which SWL explained unique variance in CBM-R was consistently greater than the unique variance explained by students' reading comprehension performance. Likewise, the magnitude of unique variance in CBM-R performance explained by SWL for both first- and second-grade students in this study exceeds that reported by Jenkins et al. (2003b). These data again provide some evidence that context may affect the performance of first- and second-grade students to a lesser extent than it does for students in upper elementary grades. At a bare minimum, these data suggest that variance in first- and second-grade students' word identification skills affects their CBM-R performance to a greater extent than it does for older students. Potentially, differences may be a function of less variability in the performance of upper elementary grade students in word identification skills and/or differences in the relationship between word decoding, reading fluency, vocabulary, and comprehension among students in different developmental stages of reading (Tilstra et al., 2009).

The final set of analyses was exploratory and only included the data from a small sample of low-achieving first-grade students. Although the data should be interpreted with extreme caution, the results do provide further support for CBM-R as a global measure of reading achievement. Results also provide fur-

ther evidence that context may play less of a role in the reading of CBM-R passages and/or sight word skills play a larger role in the reading of CBM-R passages for less skilled students. In support of CBM-R as a global measure of reading achievement, results indicated that even with only a small sample of students and a restricted range of WRCM (0–20 WRCM), CBM-R shared significant variance in students' reading comprehension outcomes. As evidence that context plays less of a role in students' reading of CBM-R passages for low-achieving students, high-frequency SWL outcomes explained unique variance in CBM-R performance after accounting for variance in reading comprehension. However, as was reported in the results, reading comprehension outcomes did not explain unique variance in CBM-R after accounting for the low-achieving students' high-frequency SWL performance. These findings provide preliminary support that it would be beneficial for school personnel to administer SWL to low-achieving students in order to better differentiate among these types of students.

Limitations

The primary limitation associated with this study is the number of regression analyses conducted and thus the increased probability of committing Type 1 errors. However, most findings were consistent with past research and those that were not consistent differed in that observed shared/unique variances were less than expected based upon past research with older students. Although some of these observed variances were recognized as significant, we acknowledge that they were lower than that observed in past research.

Another limitation to be considered is the relatively small sample size of low-achieving students and the fact that these students were identified as low achieving based upon a measure used within the analyses conducted. It is therefore essential that these analyses be treated as exploratory. Future studies are needed to evaluate the benefit of administering a SWL in addition to CBM-R passages to low-achieving first-grade students. It may be that when conducting universal screenings, school professionals should consider administering a SWL to first-grade students performing below a prespecified CBM-R level. This would possibly allow schools to further distinguish the reading skills of low-achieving students.

Other limitations to consider when interpreting the findings from this study include (a) the use of only one word list of each SWL type, (b) the administration of all measures during a single time frame (winter administration), and (c) the inclusion of only two schools. These limitations decrease the generalizability of the findings. Finally, on occasions, comparisons were made to findings of previous research, especially that of Jenkins et al. (2003a, 2003b). Although these studies also used the ITBS as a criterion measure, it is important to remember that their word list differed considerably from those employed in this study. Whereas the high-frequency and decodable SWLs employed in this study closely approximated those that might be used in the schools for universal screening purposes, the word list used by Jenkins et al. (2003a, 2003b) was developed using the words contained within the CBM-R probe administered to students. To appropriately make comparisons between lower and upper elementary students, all students should be administered materials developed using similar procedures.

Implications

Results of this study provide further support for the utility of CBM-R in evaluating students' reading achievement. Not only was a strong relationship once again found between students' CBM-R performance and a standardized measure of reading comprehension, but the current findings suggest that students' performance on CBM-R passages explains variance above and beyond their reading of word lists. Thus, these data further dispel suggestions made that CBM-R simply measures the rate at which students are able to read words. For both first- and second-grade students, context does facilitate their reading of words within CBM-R passages. However, results also provide

evidence that context may influence first- and second-grade students' CBM-R performance to a lesser degree than students in upper elementary grades and that context affects first-grade students' CBM-R performance to a lesser degree than it does for second-grade students. Thus, although CBM-R procedures do not directly measure students' comprehension skills, students' comprehension skills seem to affect their performance as does their ability to read sight words and high-frequency words. The interpretation of CBM-R as a general outcome measure seems appropriate.

Although word list reading may not explain variance in comprehension beyond that explained by CBM-R, word list reading appears to explain variance in CBM-R performance beyond that of reading comprehension for beginning readers. The variance in reading comprehension explained by CBM-R likely accounts for (a) differences in oral reading rate of words in context between readers who comprehend at least some of what they read to nonreaders who are unable to comprehend anything they read as their cognitive resources are entirely allocated to decoding; and (b) variability among students who are able to comprehend what they read but their comprehension is affected by the variability in the extent to which they dedicate cognitive resources to decoding and not the comprehension of passages. However, there remains variance to explain in the reading achievement of students who are unable to comprehend what they read, which can potentially be explained by word lists. Additional research is needed to both replicate these findings and to determine if word lists can be used to help further distinguish struggling beginning readers.

References

Adolf, S. M., Catts, H. W., & Little, T. D. (2006). Should the simple view of reading include a fluency component? *Reading and Writing, 19,* 933–958. doi:10.1007/s11145-006-9024-z

Ardoin, S. P., Carfolite, J., Christ, T. J., Roof, C. M., & Klubnik, C. (2010). Examining readability estimates' predictions of students' oral reading rate: Spache, Lexile, and Forcast. *School Psychology Review, 39,* 277–285.

Ardoin, S. P., & Christ, T. J. (2008). Evaluating curriculum-based measurement slope estimates using data from tri-annual universal screenings. *School Psychology Review, 37,* 109–125.

Ardoin, S. P., & Christ, T. J. (2009). Curriculum-based measurement of oral reading: Standard errors associated with progress monitoring outcomes from DIBELS, AIMSweb, and an experimental passage set. *School Psychology Review, 38,* 266–283.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51,* 1–18. doi:http://dx.doi.org/10.1016/j.jsp.2012.09.004

Ardoin, S. P., Suldo, S. M., Witt, J. C., Aldrich, S., & McDonald, E. (2005). Accuracy of readability estimates' predictions of CBM performance. *School Psychology Quarterly, 20,* 1–22.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connel, J. E., Koenig, J. L., Resetar, J. L., & Slider, N. J. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probe when conducting universal screening. *School Psychology Review, 33,* 218–233.

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

Blackburn, M., Dickson, P., Millam, W., & Nelson, K. (1977). *Systematic instructional management strategies (SIMS) reading program manual.* Minneapolis, MN: Minneapolis Public Schools.

Clemens, N. H., Shapiro, E. S., & Thoemmes, F. (2011). Improving the efficacy of first grade reading screening: An investigation of word identification fluency with other early literacy indicators. *School Psychology Quarterly, 26,* 231–244. doi:10.1037/a0025173

Compton, D. L., Appleton, A. C., & Hosp, M. K. (2004). Exploring the relationship between text-leveling systems and reading accuracy and fluency in second grade students who are average and poor decoders. Learning Disabilities Research & Practice, 19, 176–184.

Compton, D. L., Fuchs, D., Fuchs, L. S., & Bryant, J. D. (2006). Selecting at-risk readers in first grade for early intervention: A two-year longitudinal study of decision rules and procedures. *Journal of Educational Psychology, 98,* 394–409.

Compton, D. L., Fuchs, D., Fuchs, L. S., Bouton, B., Gilbert, J. K., Barquero, L. A., . . . , Crouch, R. C. (2010). Selecting at-risk first grade readers for early intervention: Eliminating false positives and exploring the promise of a two-stage gated screening process. *Journal of Educational Psychology, 102,* 327–340. doi:10.1037/a0018448

Compton, D. L., Olinghouse, N. G., Elleman, A., Vining, J., Appleton, A. C., Vail, J., & Summers, M. (2005). Putting transfer back on trial: Modeling individual differences in the transfer of decoding-skill gains to other aspects of reading acquisition. *Journal of Educational Psychology, 97,* 55–69. doi:10.1037/0022-0663.97.1.55

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive

skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10,* 277–299.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219–232.

Dolch, E. W. (1936). A basic sight word vocabulary. *The Elementary School Journal, 26,* 456–460.

Fry, E. (1957). Developing a word list. *Elementary English, 34,* 456–458.

Fry, E. (1980). The new instant word list. *The Reading Teacher, 34,* 284–289.

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

Good, R. H., III. (2004). *Using the Dynamic Indicators of Basic Early Literacy (DIBELS).* Retrieved May 10th, 2004, from http://dibels.uoregon.edu/index.php

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32,* 228–240.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing, 2,* 127–160. doi: 10.1007/bf00401799

The Iowa Test of Basic Skills. (2001). Rolling Meadows, IL: Riverside.

Jamgochian, E. M., Park, B. J., Nese, J. F. T., Lai, C. F., Sáez. L., Anderson, D., …, Tindal, G. (2010). *Technical adequacy of the EasyCBM grade 2 reading measures* (Technical Report No. 1004). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

January, S.-A. A., & Ardoin, S. P. (2012). The impact of context and word type on students' maze task accuracy. *School Psychology Review, 41,* 262–271.

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

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

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children, 59,* 421–432.

Johns, J. L. (1971). The Dolch basic word list—then and now. *Journal of Reading Behavior, 3,* 35–40.

Kranzler, J. H., Miller, M. D., & Jordan, L. (1999). Examination of racial/ethnic and gender bias on curriculum-based measurement of reading. *School Psychology Quarterly, 14,* 327–342.

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

Lai, C. F., Nese, J. F. T., Jamgochian, E. M., Kamata, A., Anderson, D., Park, B. J., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the EasyCBM primary-level reading measures (Grades K–1), 2009–2010 version* (Technical Report No. 1003). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Menon, S., & Hiebert, E.H. (1999). *Literature anthologies: The task for first grade readers* (Ciera Report No. 1–009). Ann Arbor, MI: Center for Improvement of Early Reading Achievement.

Mertler, C. A. (2002). *Using standardized test data to guide instruction and intervention.* College Park, MD: ERIC Clearinghouse on Assessment and Evaluation. (ERIC Document Reproduction Service No. ED470589).

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No 00–4769). Washington, DC: U.S. Government Printing Office.

Nicholson, T. (1991). Do children read words better in context or in list? A classic study revisited. *Journal of Educational Psychology, 83,* 444–450.

Nicholson, T., Bailey, J., & McArthur, J. (1991). Context cues in reading: The gap between research and popular opinion. *Journal of Reading, Writing, & Learning Disabilities International, 7,* 33–41.

Olinghouse, N. G., Lambert, W., & Compton, D. L. (2006). Monitoring children with reading disabilities' response to phonics intervention: Are there differences between intervention aligned and general skill progress monitoring assessments? *Exceptional Children, 73,* 90–106.

Perfetti, C. A. (1985). *Reading ability.* New York, NY: Oxford University Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. Gough, L. Ehri, & R. Trieman (Eds.), *Reading acquisition* (pp. 145–174). Mahwah, NJ: Erlbaum.

Pratt, K., Martin, M., White, M. J., & Christ, T. J. (2011). *Analysis of word lists for FAIP-R* (Technical Report No. 1). Minneapolis, MN: University of Minnesota.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47,* 427–469. doi:10.1016/j.jsp.2009.07.001

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement.* New York, NY: Guilford Press.

Spear-Swerling, L. (2006). Children's reading comprehension and oral reading fluency in easy text. *Reading and Writing, 19,* 199–220.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Pearson.

Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32,* 383–401.

University of Iowa. (2005). *Iowa Test of Basic Skills, Form A.* Rolling Meadows, IL: Riverside Publishing.

Zinar, S. (2000). The relative contributions of word identification skill and comprehension-monitoring behavior to reading comprehension ability. *Contemporary Educational Psychology, 25,* 363–377. doi:10.1006/ceps.2000.1024

Date Received: February 8, 2012
Date Accepted: December 11, 2012
Action Editor: Joseph Betts ■

Scott P. Ardoin, PhD, received his PhD from Syracuse University and is now an associate professor of school psychology within the College of Education at the University of Georgia. His research activities focus on applying the principals of behavioral analysis to improving the quality of assessment and interventions provided to students in elementary schools.

Tanya L. Eckert, PhD, received her PhD in school psychology from Lehigh University in 1996 and is an associate professor of psychology at Syracuse University. Her research interests include examining procedures for assessing academic and behavioral problems, developing classroom-based interventions, and measuring the acceptability of assessment and intervention procedures.

Theodore J. Christ, PhD, is an associate professor in the Educational Psychology Department at the University of Minnesota. He is also the director of the Formative Assessment System for Teachers (FAST; http://fast.cehd.umn.edu) and codirector of the Research Institute for Problem Solving (RIPS). He is engaged in research to develop, evaluate, and improve assessments that are used to enhance data-based decisions, problem solving, and response to intervention. He received the 2008 Lightner Witmer Early Career Award for outstanding early career scholarship from Division 16 of the American Psychological Association.

Laura Morena, MA, is currently a doctoral candidate in the school psychology program at the University of Georgia. Her research interests include reading assessment, reading intervention, behavior management, and educational strategies for English language learners.

Stacy-Ann A. January, MA, is currently a doctoral candidate in the school psychology program at the University of Georgia. Her research interests are in the areas of reading and curriculum-based measurement.

Jeffrey F. Hine, MEd, BCBA, received his MEd in special education from Vanderbilt University in 2005 and is currently a doctoral candidate in the school psychology program at the University of Georgia. His research interests include applied behavior analysis, behavioral consultation to schools, parent training and support, autism spectrum disorders, and the uses of technology in supporting skill acquisition and behavior management.