

When Students Perform at the Below Basic Level on the NAEP: What Does It Mean and What Can Educators Do?

Elfrieda H. Hiebert

When the National Assessment of Educational Progress in Reading (NAEP) results are published biennially, journalists and policymakers focus on the approximately third of a fourth-grade cohort who fail to attain the basic standard in reading comprehension. A legitimate concern is that these students do not have the literacy levels required for full participation in the global digital world of the 21st century. However, the attributions and claims of their literacy levels go far beyond this concern, as illustrated in a segment on the National Public Radio website titled “Why Millions of Kids Can’t Read and What Better Teaching Can Do About It” (Hanford, 2019).

In 1992, the NAEP framework committee (National Assessment Governing Board, 1992) called for periodic studies of oral reading fluency (ORF) to establish whether students can recognize the words on the NAEP passages when reading orally. In the latest NAEP ORF study (White et al., 2021), oral reading assessments were administered to a sample of students who represented the 36% of fourth graders who performed below basic on silent reading comprehension on the 2017 assessment. The research team divided students who performed below the basic level into three groups: high, medium, and low. Averages for words correct per minute (WCPM) and accuracy levels for these three groups were as follows: high, 108 WCPM and 94% accuracy; medium, 95 WCPM and 92% accuracy; and low, 71 WCPM and 83% accuracy. These findings led the researchers to conclude that most students scoring below the basic level had problems with fluency, word reading, and phonological decoding and could benefit from support in these areas. They also recommended investigations into whether elementary schools are teaching accurate and efficient reading skills.

Reports of a strong correlation between ORF and silent reading comprehension (Reschly et al., 2009) have increased the emphasis on ORF in assessment, instruction, interventions, and policies (e.g., *No Child Left Behind*, 2002).

A correlation between two variables like rate of oral reading and silent reading comprehension, however, does not necessarily mean that one is the cause of the other. The profiles of students who do poorly on silent reading comprehension assessments vary considerably (Buly & Valencia, 2002). There are students who read slowly on oral reading tasks but do reasonably well on silent reading tasks, just as there are fast oral readers who do not comprehend well (Trainin et al., 2015). Furthermore, oral reading fluency interventions have shown relevantly lackluster results on silent reading comprehension (O’Keeffe et al., 2012). Additionally, even though oral fluency rates of students in grades one through four, including those in the bottom quartile, increased over the past 15 years (Hasbrouck & Tindal, 2017), the percentage of students who score below the basic level on the NAEP silent reading assessment over this period has remained stable (National Center for Education Statistics, 2019).

Before policymakers respond to the performances of students on the most recent NAEP/ORF study with mandates for fluency and word-level interventions at middle grades and beyond, available data from ORF assessments merit investigation. In this paper, I examine the performances on oral reading assessments of students who score at or below the 36 percentile to identify their strengths and challenges. I then use patterns from these analyses to address how instruction and interventions can better support students in developing the proficiencies required to perform successfully on silent reading comprehension tasks.

What Do Students Who Score Below the Basic Level Know?

Students who score below the basic level typically are viewed from a deficient perspective, but the numbers of words read correctly per minute by students in the three sub-groups—77–108 (White et al., 2021)—indicate that

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Elfrieda H. Hiebert the president and CEO of TextProject at TextProject.org, Santa Cruz, CA, USA; email hiebert@textproject.org.

students were able to recognize at least some words. What are the words these students know? Passages from the NAEP ORF are unavailable for analysis, but passages and norms for WCPM and accuracy are available for the ORF assessment of DIBELS (Gray et al., 2018). The DIBELS organization presents ORF data (both WCPM and accuracy) as percentiles rather than as averages (as in the NAEP/ORF report), which allows for a view of subgroup variation. Table 1 presents performances for the top and middle percentiles of three groups within the bottom 36 percentiles as well as for the 1st percentile.

Variation in Groups

The data in Table 1 indicate considerable variation among the students scoring below the basic level. The greatest variation is in the low group, where WCPM declines by 58 words and accuracy by 32%. Reading educators such as Clay (1989) identified accuracy levels of 90% as adequate for meaningful participation in fluency interventions. At the 90% level, students may recognize the meanings of words slowly but they are able to recognize the majority of the words in texts. The drop below a 90% accuracy level occurs at the 5th percentile, suggesting that students in the bottom 5th percentile require different solutions than their peers in the 36th to 6th percentile range. Research on the needs and instruction of students with dyslexia and severe reading disabilities is extensive (Gjessing & Karlsen, 2012).

In this paper, I attend to the profiles of and solutions for students from the 36th to 6th percentiles in the middle grades. Researchers have paid substantial attention to primary-grade solutions, which is both necessary and appropriate to ensure a strong literacy foundation; however, for students for whom primary-grade instruction

has not produced adequate literacy proficiency, the nature of instruction in the middle grades also requires attention. The solutions for beginning readers should not be assumed to be the same for students who are reading below expected levels in the middle grades and beyond.

Method of Predicting Known and Unknown Words

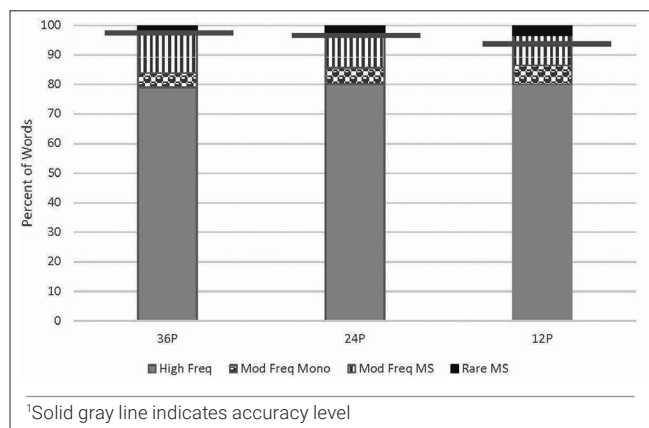
To predict the words known and unknown by students in the three groups formed of those scoring below the basic level, I used a digital tool (Hiebert, 2011) to analyze all words in DIBELS passages that were read by students at a particular percentile. For example, the analysis of 12th percentile students consisted of the first 82 words from the three end-of-year DIBELS-6 passages since this group of students read 77 words at 94% accuracy. Students at the 36th and 24th percentiles read more words than 12th percentile students, meaning the profiles of words in Figure 1 for different percentile groups are similar but not identical.

The variables used to analyze words were ones that have been shown to influence students' word knowledge in vocabulary studies (e.g., Hiebert et al., 2019; Lawrence et al., 2021). For comprehension to occur, students must recognize a word's meaning, not simply pronounce a word. The ability to decode a word is fundamental to this process, but simply decoding a word without recognition of its meaning is insufficient (Perfetti & Hart, 2001). Numerous variables can influence the recognition of a word's meaning, but the frequency with which a word occurs in written language consistently predicts knowledge of word meanings. A second variable that is highly influential in developing automatic recognition of word meaning is the number

Table 1
WCPM and Accuracy Levels on DIBELS Assessment by Fourth Graders at Different Percentiles (End of Year)

Performance Level	Percentile	WCPM	Accuracy
Performing below basic: High	36	110	98
	30	104	98
Performing below basic: Medium	24	97	97
	18	88	96
Performing below basic: Low	12	77	94
	6	59	91
	1	19	62

Figure 1
Profiles of Texts Read by Fourth Graders at Three Percentiles¹



of syllables in a word. These two variables are the basis of the four-part rubric for sorting words within DIBELS passages that appears in Table 2.

A third variable, age of acquisition (AoA), measures the likelihood that students hear or use a word in their oral language environments (Kuperman et al., 2012). Age of acquisition also has been shown to influence students' recognition of word meaning. Data on age of acquisition were applied when candidates for unknown words are numerous.

Predictions of Known and Unknown Words

Figure 1 shows the profiles of the words in DIBELS passages read by students at the three target percentile points. Precise errors cannot be identified but based on the profiles of the texts and findings from previous research, predictions can be made about the words that students in the three percentile groups likely knew and did not know. Table 3 presents these predictions for the three percentile

groups. The data in Table 3 indicate that students from the 36th to 6th percentiles can read the most frequent words in texts (i.e., the 1,000 most frequent word families). Reading high-frequency words is no small feat; numerous grapheme–phoneme relationships for vowels are present in this group of words, including multisyllabic words.

Students in the 36th and 24th groups are able to read a substantial number of moderately frequent words with both monosyllabic and multisyllabic patterns. Even when students at these percentile levels have not encountered words previously in texts, they appear able to recognize rare words with a single syllable. Students at the 36th and 24th are distinguished in how they recognize words with multiple syllables. While students in the former group are likely able to recognize the meanings of rare, two-syllable words, students in the latter group are challenged by these words.

Students' errors in the 12th percentile levels were more extensive than the number of rare, multisyllabic words in the texts they read; these students likely did not recognize at least a modicum of moderately frequent, multisyllabic words (e.g., *formal*) that may not be used extensively in their oral language environments. The texts also contained proper names with unusual orthographic patterns (e.g., *Niagara*) that were likely challenging to students in this percentile group.

What Are the Challenges for Students Who Score Below the Basic Level?

Fourth graders in the 36th to 6th percentiles can recognize numerous words, especially words they have seen frequently in texts; however, these students cannot be described as highly proficient oral readers. The degree of challenge is not precisely the same for all students in the 36th to 6th percentile range, but the nature of the challenge is similar.

Table 2
Rubric for Establishing Word Complexity

Category	Description	Examples
1 (Easiest)	Highly frequent (HF) words with frequencies of 100 or more per million words of text (/million) (Zeno et al., 1995), both monosyllabic (Mono) and multisyllabic (Multi)	HF, Mono: <i>for, when</i> HF, Multi: <i>many, other</i>
2	Moderately frequent (MF), Mono words with frequencies of 99 to 10/million	MF, Mono: <i>rolls, twelve</i>
3	MF, Multi words with frequencies of 99 to 10/million and Rare, Mono words with frequencies of 9 or fewer/million	MF, Multi: <i>vacation, details</i> Rare, Mono: <i>maid, wreathe</i>
4 (Hardest)	Rare, Multi words with frequencies of 9 or less/million	Rare, Multi: <i>astronomy, blistering</i>

Table 3
Predictions of Known and Unknown Words for 36 to 6 Percentiles (Fourth Grade)

Percentile Group	Total Words Read (per passage)	Accuracy (%)	Predicted Words Known	Predicted Errors
36	112	98	<ul style="list-style-type: none"> ■ HF (Mono and Multi) ■ MF Mono ■ MF Multi ■ Rare Mono 	<ul style="list-style-type: none"> ■ Rare three-syllable words with high AoA (e.g., <i>astronomy</i>) ■ Rare three-syllable proper names (e.g., <i>Niagara</i>)
24	100	97	<ul style="list-style-type: none"> ■ HF (Mono and Multi) ■ MF Mono ■ MF Multi ■ Rare Mono 	Same as above, plus <ul style="list-style-type: none"> ■ Rare two-syllable words with high AoA (e.g., <i>pinto</i>) ■ Rare two-syllable proper names (e.g., <i>Mitchell</i>)
12	82	94	<ul style="list-style-type: none"> ■ HF (Mono and Multi) ■ MF Mono ■ MF Multi ■ Rare Mono (most but not all) 	Same as above, plus <ul style="list-style-type: none"> ■ Some MF MS words with high AoA (e.g., <i>advance, educated</i>)

Automaticity

Automaticity in recognizing the meanings of words in texts is essential for comprehension. If readers' cognitive resources are consumed with identifying word meanings, word by word, then readers are unlikely to be able to comprehend a text's content (LaBerge & Samuels, 1974). According to DIBELS guidelines, the rates of fourth graders at the 36th to 6th percentile group would mean all are classified as at-risk (Gray et al., 2018).

Often, an at-risk designation has translated into interventions at the middle grades that stress decoding instruction and fluency training (California State Board of Education, 2015). As already noted, faster ORF rates among first through fourth graders in the United States (Hasbrouck & Tindal, 2017) have not resulted in higher performances on the silent reading assessment of the NAEP.

An explanation for the lack of growth in silent reading comprehension as a result of ORF interventions may lie in the differences between oral and silent reading processes and the contexts of assessment (Vorstius et al., 2014). In oral reading, either an adult or a digital device will prompt students if they stop reading. In silent reading contexts, students receive no prompting; they must monitor their recognition of word meanings and their construction of the text's meaning on their own.

Moreover, silent reading tasks are invariably longer than oral reading tasks. Whereas the texts of the NAEP's

ORF assessments averaged 162 words in length (White et al., 2021), the texts in the 2017 NAEP silent reading assessments averaged 750 words. An additional difference between the two contexts is that the task on which silent reading proficiency is based—answers to comprehension questions—occurs only after students have read the entire text of 750 words.

Comprehension—the outcome of the silent reading task—involves processes beyond the speed of decoding words orally. As a result, interventions that attend to only fluency in oral reading may not address the issues that underlie poor comprehension performances. For example, when students with low comprehension levels but fast reading rates on silent reading tasks in a digital context were asked to read similar texts on paper in a setting observed by an adult, they comprehended well (Hiebert & Daniel, 2019). The researchers concluded that poor comprehension in independent silent reading may be due to factors in addition to recognition of word meanings (such as the assessment context).

Multisyllabic Words

The evidence points to multisyllabic words, especially words with three or more syllables, as a challenging area for fourth graders in the 36th to 6th percentiles (see Table 2). Students in this group appear to perform adequately with monosyllabic words, although not necessarily with high levels of automaticity. But students who score below

the basic level seem to be especially challenged in recognizing the meanings of multisyllabic words.

This interpretation of the challenge of multisyllabic words requires a cautionary note; all multisyllabic words are not similar in their word meaning demands. In beginning texts dominated by short words, young children may attend to unusual, long words such *elephant* or *hippopotamus*. Furthermore, if students can recognize the 1,000 most frequent words, they have at least some facility with multisyllabic words (e.g., *important*, *children*).

It is multisyllabic words, however, that pose the greatest challenge for students who are at the beginning stages of reading and those who are performing below expected levels. Most of the words that are unknown by children in the bottom half of a first-grade distribution are multisyllabic (Hiebert et al., 2020). Multisyllabic words continue to challenge students designated as below expected levels in middle grades through high school. Wang et al. (2019) reported that students from 5th through 10th grade who did not attain a specified proficiency level on a decoding assessment (almost entirely of multisyllabic words) failed to progress in reading comprehension over the next 3 years.

Decoding instruction in the elementary grades often has concentrated on letter–sound patterns in monosyllabic words. Typical curricula introduce multisyllabic words toward the end of Grade 2, but many instructed strategies have limited applicability (Kearns, 2020). The treatment of multisyllabic words in reading instruction, both in the primary and middle grades, must be addressed for students who score below basic levels to reach proficient literacy levels.

What Have Been the Effects of Response to Intervention Projects?

Students' performance on the NAEP has been the incentive for congressional legislation, the most notable of which is the Individuals with Disabilities Education Improvement Act (2004). To meet all students' needs as mandated by this legislation, the response to intervention (RTI) model was initiated; under the RTI model, students with varying needs participate in different tiers of reading instruction (Fuchs & Fuchs, 2006).

Studies designed to provide evidence-based practices, especially for readers who are performing below expected levels, have a long history (e.g., Harris, 1967) and number in the thousands. But a particular form of intervention has proliferated since the Institute for Education Sciences mandated randomized controlled treatments (RCTs) as the gold standard for instructional efficacy (Whitehurst, 2005). Multiple components

(usually decoding, fluency, and comprehension) distinguish these studies, and the studies are conducted over a relatively extended period (typically at least 40 hours; Scammacca et al., 2016).

Since 1996, researchers have described via eight meta-analyses the effects of RCT intervention studies focusing on fluency, word recognition and decoding, and comprehension. In a narrative review of this group of meta-analyses, Scammacca et al. (2016) concluded that the most effective interventions focused on comprehension rather than decoding and fluency and that even the comprehension interventions typically had relatively low effects on standardized comprehension measures.

What Instructional Components Must Be Addressed for Students to Receive the Experience They Need?

At least one-third of a grade cohort is not attaining the literacy levels required for full participation in 21st-century careers and communities. This situation warrants addressing the content of current English/Language Arts (ELA) curriculum and instruction, including interventions. Middle-grade readers' challenges are complex and solutions are not simple. Implementing more phonics instruction or increasing oral fluency practice has not produced the desired results, which is not to say that experiences that promote and extend decoding and fluency should be left unaddressed. In particular, the manner in which students in the middle grades can be supported in recognizing the meanings of multisyllabic words require attention. Studies of morphological interventions have shown positive effects (Goodwin & Ahn, 2013) and results from these interventions have been modelled in classroom settings (Kieffer & Lesaux, 2007). Additionally, educators need to attend to the manner in which three components are addressed in ELA classrooms as well as content-area instruction for middle-grade students: mode of reading, text, and reading volume.

Mode of Reading

Oral reading provides an easily accessible context for examining students' automaticity in recognizing words. The mode of reading assessment seems to have influenced the mode of reading in elementary classrooms (Brenner et al., 2009) as well as in middle and high school ELA and social studies classrooms (Swanson et al., 2016). What is uncertain is the degree to which reading words faster in oral reading contexts transfers to the construction and monitoring of meaning, which is the goal of silent reading.

Research that addresses how to support automaticity in silent reading in the context of small group interventions is limited. One strategy that has proven effective in supporting students' silent reading comprehension in several interventions is to alternate between silent and oral reading of sections of text (Trainin et al., 2016; Vadasy & Sanders, 2008). In the oral reading phase of the reading cycle, teachers and tutors can provide the monitoring and feedback that characterize effective ORF interventions (Scammacca et al., 2016).

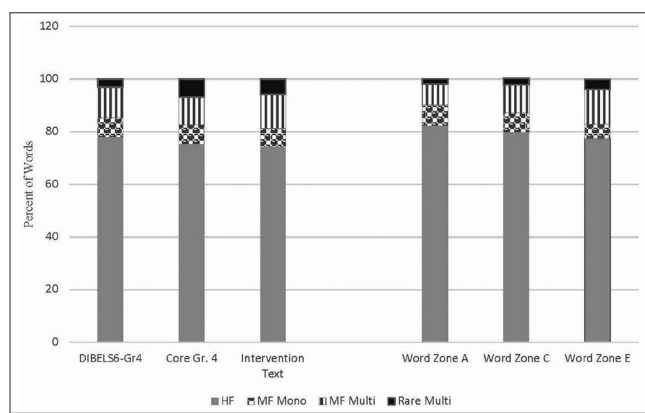
Approximately 1.5 million students are in the 36th to 6th percentiles of an American grade cohort. Resources for face-to-face interventions, even in small groups, may be simply insufficient to provide the intensive and extensive support to the critical mass of students who need help. Educators might explore the role of digital intervention programs that are designed to increase the amount of reading that students who perform below basic are doing. Although interventions that aim to increase volume of reading through digital devices are few, participation in digital programs has been shown to increase how much students read and to result in significant gains in standardized reading comprehension (e.g., Spichtig et al., 2019).

Text

Research in early reading shows the importance of connecting the content of word recognition lessons to the words in texts read by students (Juel & Roper/Schneider, 1985). In RTI interventions such as those reviewed by Scammacca et al. (2016), the same zeal has often not been paid to the selection of texts as to the words of instruction. Texts may be labeled as "grade-level" texts but the role of vocabulary in determining these designations is frequently not established, even though readers who are performing below expected levels often find vocabulary a challenge. In instructing children to read in English, current pedagogy relies on either a quantitative approach, where vocabulary serves as a less powerful factor than sentence length (Cunningham et al., 2018), or a qualitative approach, where the number of words in texts and sentence length are the best predictors of assigned text complexity levels (Cunningham et al., 2005).

By contrast, vocabulary drives text selection in the English as a second language (ESL) field, which is aimed at young adults preparing to attend English-speaking higher education institutions. ESL students move through an instructional progression of texts that is based on research that specifies the number of words that readers need to know to comprehend texts (Nation, 2014) and

Figure 2
Profiles of Four Types of Texts: DIBELS Assessment, Core Reading Program, Intensive Intervention Texts, and Word Zone Texts of Three Levels



on the percentages of words that account for portions of texts at different levels (Nation & Waring, 1997).

The word zone approach is a perspective on text that closely aligns to the ESL perspective (Hiebert, 2011). The word zones are based on the 2,500-word families (lead words and their morphological relatives) that account for the majority of words in texts (Hiebert et al., 2018). Approximately half of these 2,500 families appear in primary-grade texts, while the remaining word families become prominent in subsequent grade bands. If students fail to build automaticity in recognizing meanings of word families in early bands or zones, they are unlikely to be successful in becoming automatic with the word families added in subsequent zones.

An obstacle to students' automaticity with the highly and moderately frequent members of word families lies in the large numbers of rare words that occur in instructional texts, beginning in kindergarten and first grade (Fitzgerald et al., 2016). To illustrate, consider the texts in Figure 2 from fourth-grade, core reading programs and texts adopted for intensive interventions for Grades 4 through 8 (California State Board of Education, 2015). I analyzed these texts according to the same criteria as those used to determine words known and unknown by students (see Table 1). The core reading texts used in tier-one instruction have an average of 7% words of this type. When students go to tier-two or tier-three interventions, they receive texts with an average of 6% rare, multisyllabic words. If students who perform at below basic levels are reading slowly on DIBELS texts where rare words account for 3% of the total words, it is highly unlikely that their automaticity will be enhanced when reading texts with an average of 6%–7% rare words in intervention and instructional contexts.

Table 4
Predictions of Exposure to Middle-Grade Critical Words as a Function of Amount of Daily Classroom Reading

Minutes Read Daily	Words Read Daily	Total Words Read in Grade 4	Predicted Exposure to Word Zone Critical Words
8	738	132,840	2.7
12	1,104	198,720	4.0
16	1,472	264,960	5.3
20	1,840	331,200	6.6

Figure 2 also includes the profiles of a program of texts, available as open-access texts (Hiebert, 2022), based on the word zone approach. Compared to the other programs in Figure 2, highly and moderately frequent words account for a higher portion of the word zone texts. That is, students are having the opportunity to gain automaticity with the word families that account for the majority of the words in text (i.e., the 2,500-word families) without having to attend to large numbers of rare words, which are typically multisyllabic.

Several studies conducted with word zone texts have shown that, even in short interventions, students' gains are sufficient to influence standardized comprehension scores (Trainin et al., 2016; Vadasy & Sanders, 2008). Researchers have not investigated how long and intensive the exposure to word zone or ESL instruction texts needs to be for students to become highly automatic as readers in independent, silent reading contexts. Indeed, little is known about the amount of text reading that is needed to attain particular levels of reading proficiency over students' school careers, as is evident in the discussion that follows on reading volume.

Reading Volume

As in any domain of expertise, reading proficiency depends on extensive practice. A straightforward question for educators to ask is: What does the research say about how much text students need to read at different developmental levels to become automatic in recognizing the meanings of the majority of words that make up texts? From educators' perspectives, answers to this question will determine how they allocate time to reading in their classrooms. Unfortunately, answers to this question from research are few.

One useful way of establishing appropriate amounts of reading for students who perform at below basic levels is to apply data on predicted appearances of words in written texts. Consider, for example, the number of opportunities with the general academic words—a source of particular challenge for many readers (Nagy & Townsend, 2012)—that

students can be expected to have with different amounts of time spent reading. The 570 words on the Academic Word List (Coxhead, 2000) are each predicted to occur an average of 24 times per million words of text. General academic words such as *transfer* and *approximate* first appear with substantial frequency in texts at grades 4 and 5 and continue to increase in importance in subsequent grades.

The data in Table 4 illustrate how frequently a student at the 18th percentile (in the middle of the students who perform below basic levels) is likely to encounter the word *transfer* with different amounts of reading. The number of repetitions required to become automatic in recognizing word meanings is unknown, but the critical feature of word repetition may lie with the diversity of the contexts in which words appear (Hoffman et al., 2013). For example, *transfer* varies in its meanings, from a unit on heat transfer to a description of a person's enrollment in a different school. Only through extended reading opportunities in ELA and content-area classrooms are students likely to gain the automaticity with the meanings of words that make up the majority of words in texts.

Conclusion

In this paper, I have not addressed all aspects of the opportunities required for successful reading. For example, I did not cover the differentiation of experiences for students in different bands of those performing below the basic level. Such research has urgency when policymakers mandate specific interventions—often for all students. A growing number of states have mandated specific word recognition programs for all students in elementary grades (Schwartz, 2021), but evidence for these mandates (such as the need for all students to move through the same word recognition program at the same pace) is lacking (Stevens et al., 2021). Interventions and classroom instruction must address both the strengths and the challenges of students who are performing at below basic levels. Only then will students get the support they require and deserve.

Conflict of Interest

No conflict of interest to declare.

Funding

No external funding supported the research and writing of this article.

REFERENCES

- Brenner, D., Hiebert, E. H., & Tompkins, R. (2009). How much and what are third graders reading. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 118–140). Guilford.
- Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis, 24*(3), 219–239. <https://doi.org/10.3102/01623737024003219>
- California State Board of Education. (2015, November 5). *2015 English language arts/English language development instructional materials adoption (K-8)*. Sacramento, CA: Author. Retrieved from www.cde.ca.gov/ci/rl/im
- Clay, M. M. (1989). *The early detection of reading difficulties* (3rd ed.). Heinemann.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. <https://doi.org/10.2307/3587951>
- Cunningham, J. W., Hiebert, E. H., & Mesmer, H. A. (2018). Investigating the validity of two widely used quantitative text tools. *Reading and Writing, 31*(4), 813–833. <https://doi.org/10.1007/s11145-017-9815-4>
- Cunningham, J. W., Spadorcia, S. A., Erickson, K. A., Koppenhaver, D. A., Sturm, J. M., & Yoder, D. E. (2005). Investigating the instructional supportiveness of leveled texts. *Reading Research Quarterly, 40*(4), 410–427. <https://doi.org/10.1598/RRQ.40.4.2>
- Fitzgerald, J., Elmore, J., Relyea, J. E., Hiebert, E. H., & Stenner, A. J. (2016). Has first-grade core reading program text complexity changed across six decades? *Reading Research Quarterly, 51*(1), 7–28.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93–99. <https://doi.org/10.1598/RRQ.41.1.4>
- Gjessing, H. J., & Karlsen, B. (2012). *A longitudinal study of dyslexia: Bergen's multivariate study of children's learning disabilities*. Springer.
- Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading, 17*(4), 257–285. <https://doi.org/10.1080/10888438.2012.689791>
- Gray, J. S., Warnock, A. N., Kaminski, R. A., & Good, R. H. (2018). *Acadience™ Reading National Norms 2014-2015* (Technical Report No. 23). Dynamic Measurement Group.
- Hanford, E. (January 2, 2019). Why millions of kids can't read and what better teaching can do about it. National Public Radio website. Retrieved from <https://www.npr.org/2019/01/02/677722959/why-millions-of-kids-cant-read-and-what-better-teaching-can-do-about-it>
- Harris, A. J. (1967). *Five decades of remedial reading*. Paper presented at the meeting of the International Reading Association, Seattle, WA. Retrieved from <https://files.eric.ed.gov/fulltext/ED011830.pdf>
- Hasbrouck, J., & Tindal, G. (2017). An update to compiled ORF norms. *Behavioral Research and Teaching*.
- Hiebert, E. H. (2011). *The WordZone profiler*. TextProject.
- Hiebert, E. H. (2022). *TopicReads*. TextProject. Retrieved from <http://textproject.org/teachers/students/>
- Hiebert, E. H., & Daniel, M. (2019). Comprehension and rate during silent reading: Why do some students do poorly? *Reading and Writing, 32*(7), 1795–1818. <https://doi.org/10.1007/s11145-018-9917-7>
- Hiebert, E. H., Goodwin, A. P., & Cervetti, G. N. (2018). Core vocabulary: Its morphological content and presence in exemplar texts. *Reading Research Quarterly, 53*(1), 29–49. <https://doi.org/10.1002/rrq.183>
- Hiebert, E. H., Scott, J. A., Castaneda, R., & Spichtig, A. (2019). An analysis of the features of words that influence vocabulary difficulty. *Education Sciences, 9*(1), 8. <https://doi.org/10.3390/educs9010008>
- Hiebert, E. H., Toyama, Y., & Irey, R. (2020). Features of known and unknown words by first graders of different proficiency levels in winter and spring. *Education Sciences, 10*(12), 389.
- Hoffman, P., Ralph, M. A. L., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods, 45*(3), 718–730. <https://doi.org/10.3758/s13428-012-0278-x>
- Juel, C., & Roper/Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly, 134*–152. <https://doi.org/10.2307/747751>
- Individuals With Disabilities Education Improvement Act of 2004. Pub. L. No. 108–446, § 118, Stat. 2647 (2004).
- Kearns, D. M. (2020). Does English have useful syllable division patterns? *Reading Research Quarterly, 55*(S1), S145–S160. <https://doi.org/10.1002/rrq.342>
- Kieffer, M. J., & Lesaux, N. K. (2007). Breaking down words to build meaning: Morphology, vocabulary, and reading comprehension in the urban classroom. *The Reading Teacher, 61*(2), 134–144. <https://doi.org/10.1598/RT.61.2.3>
- Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. <https://doi.org/10.3758/s13428-012-0210-4>
- LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lawrence, J. F., Knoph, R., McIlraith, A., Kulesz, P. A., & Francis, D. J. (2021). Reading Comprehension and Academic Vocabulary: Exploring Relations of Item Features and Reading Proficiency. *Reading Research Quarterly*.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91–108. <https://doi.org/10.1002/RRQ.011>
- Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1–16.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.
- National Assessment Governing Board. (1992). *Reading framework for the 1992 National Assessment of Educational Progress*. U.S. Government Printing Office.
- National Center for Education Statistics. (2019). *Nation's report card: 2019 National Assessment of Educational Progress Reading Assessment*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Dept. of Education.
- No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).
- O'Keeffe, B. V., Slocum, T. A., Burlingame, C., Snyder, K., & Bundo, K. (2012). Comparing results of systematic reviews: Parallel reviews of research on repeated reading. *Education and Treatment of Children, 333*–366. <https://doi.org/10.1353/etc.2012.0006>
- Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). APA.

- 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*(6), 427–469. <https://doi.org/10.1016/j.jsp.2009.07.001>
- Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–12, 1914–2014. *Review of Educational Research, 86*(3), 756–800. <https://doi.org/10.3102/0034654316652942>
- Schwartz, S. (2021). More states are making the 'science of reading' a policy priority. *Education Week*. Retrieved from <https://www.edweek.org/teaching-learning/more-states-are-making-the-science-of-reading-a-policy-priority/2021/10>
- Spichtig, A. N., Gehsmann, K. M., Pascoe, J. P., & Ferrara, J. D. (2019). The impact of adaptive, web-based, scaffolded silent reading instruction on the reading achievement of students in grades 4 and 5. *The Elementary School Journal, 119*(3), 443–467. <https://doi.org/10.1086/701705>
- Stevens, E. A., Austin, C., Moore, C., Scammacca, N., Boucher, A. N., & Vaughn, S. (2021). Current state of the evidence: Examining the effects of Orton-Gillingham reading interventions for students with or at risk for word-level reading disabilities. *Exceptional Children, 111*(3), 387. <https://doi.org/10.1177/0014402921993406>
- Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., Fogarty, M., & Hairrell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly, 32*(3), 199–222. <https://doi.org/10.1080/10573569.2014.910718>
- Trainin, G., Hayden, H. E., Wilson, K., & Erickson, J. (2016). Examining the impact of QuickReads' technology and print formats on fluency, comprehension, and vocabulary development for elementary students. *Journal of Research on Educational Effectiveness, 9*(1), 93–116. <https://doi.org/10.1080/19345747.2016.1164778>
- Trainin, G., Hiebert, E. H., & Wilson, K. M. (2015). A comparison of reading rates, comprehension, and stamina in oral and silent reading of fourth-grade students. *Reading Psychology, 36*(7), 595–626. <https://doi.org/10.1080/02702711.2014.966183>
- Vadasy, P. F., & Sanders, E. A. (2008). Benefits of repeated reading intervention for low-achieving fourth-and fifth-grade students. *Remedial and Special Education, 29*(4), 235–249. <https://doi.org/10.1177/0741932507312013>
- Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition, 22*(3–4), 458–485. <https://doi.org/10.1080/13506285.2014.881445>
- Wang, Z., Sabatini, J., O'Reilly, T., & Weeks, J. (2019). Decoding and reading comprehension: A test of the decoding threshold hypothesis. *Journal of Educational Psychology, 111*(3), 387. <https://doi.org/10.1037/edu0000302>
- White, S., Sabatini, J., Park, B. J., Chen, J., Bernstein, J., and Li, M. (2021). *The 2018 NAEP Oral Reading Fluency Study* (NCES 2021-025). U.S. Department of Education. Washington, DC: Institute of Education Sciences, National Center for Education Statistics. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2021025>
- Whitehurst, G. J. (2005). *Biennial Report to Congress: Institute of Education Sciences*. Retrieved on August 14, 2021, from <http://ies.ed.gov/pdf/biennialrpt05.pdf>
- Zeno, S., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates.