Thinking through research and the science of reading

When it comes to learning to read, the research is clear about some things, but our knowledge is still growing, and questions remain about how to apply what we know in classrooms.

By Elfrieda H. Hiebert

The term "science of reading" has been prominent in educational publications and conversations over the past five years. During this time, I have read numerous research articles and books; listened to podcasts, webinars, and conference presentations; and interacted with colleagues with a range of perspectives. I also have explored several new, challenging, and previously unfamiliar lines of research in my own writing (Hiebert, Toyoma, & Irey, 2020).

I have gained several insights from these experiences. But to understand those ideas, we need to first consider the task of generating evidence-based reading practices from research.

Identifying what is evidence-based

The massive body of research on the learning and instruction of reading is a fairly recent phenomenon. To illustrate the increase in studies on reading over the past 120 years, I entered "phonics instruction" into Google Scholar and searched for findings in each of the 12 previous decades, starting with 1900-09 and ending with 2010-19. In the last three decades of the 20th century, the number of publications was almost five times the number in the first seven decades (17,940 versus 3,675). Articles published in the first two decades of the 21st century were almost double the amount for 1970 to 1999 — 34,300.

AT A GLANCE

- The science of reading has captured the attention of educators, policy makers, and the public.

- Research is clear that proficient reading is built on automatic recognition of consistent and common sound-letter patterns.

- Emerging research suggests that recognizing letter-sound patterns results from encountering those patterns in words rather than through rule memorization.

- Research calls into question common assumptions about the 4th-grade slump and what below-basic performance looks like.

- Understanding the sources of below-basic performances can lead to instruction that addresses students' needs.

ELFRIEDA H. HIEBERT (hiebert@textproject.org; @freddyreads) is president and CEO of TextProject. She is the author of *Teaching Words and How They Work: Small Changes for Big Vocabulary Results* (Teachers College Press, 2020).

Admittedly, this Google Scholar search has a few flaws. Were all the identified articles reports of empirical studies? What would the volume of studies be with keywords such as "phonemic awareness," "phonological decoding," "word recognition," "automaticity," "fluency," "comprehension," "prior knowledge," "spelling," and so on? However, the figures give us a general sense of the rapid pace at which new research has been published since federal investments in educational research began in the late 1960s.

As the research on reading acquisition and development has grown, national commissions under the aegis of the National Academy of Education (Anderson et al., 1985); the National Institute of Child Health and Human Development (National Reading Panel, 2000); or the National Research Council (Snow Burns, & Griffin, 1998) have vetted the research. These commissions were made up of scholars from different research perspectives who spent a substantial amount of time identifying and reviewing appropriate research and coming to a consensus on conclusions.

Like these previous efforts, the science of reading aims to increase the use of evidence-based research in reading instruction. The science of reading work is distinct from previous efforts, however, in that educators have no nationally vetted report to which they can refer. Without such documents, we must turn to a variety of sources for our understanding about what is scientifically valid. Some of these are research reviews in peer-reviewed journals (e.g., *Reading Research Quarterly*'s 2020 issues on the science of reading) and books that summarize a scholar's line of work (Seidenberg, 2017). At the same time, messages about what is scientific have proliferated in professional magazines, newspaper articles, blogs, and Facebook posts. Determining what is evidence-based can be challenging, especially because the science of reading label is placed on numerous products — sets of texts, workshops, books, and so on — that may or may not have scientific backing.

Through my experiences with research over the past five years, I have identified three categories of evidence: research that provides unequivocal conclusions, research that holds promise for solving enduring problems, and research that calls into question long-standing assumptions. I will illustrate each category with an aspect of reading development or instruction.

Research findings that are unequivocal

The orthography (i.e., spelling system) of English is quasi-regular (Plaut et al., 2020), which means that, although English writing is always alphabetic, the relationships between the 26 graphemes (i.e., letters) and 44 phonemes (i.e., sounds) can vary. This variation is especially significant for vowels, as shown by Edward Fry's (2004) identification of 106 combinations of sounds and letters for vowels. The variability in sound-letter relationships for vowels is illustrated with the phoneme /i/, typically referred to as the long e. Using Paul Hanna and colleagues' (1966) comprehensive database, Fry identified 16 different graphemes associated with the phoneme. Of the 16 graphemes, 12 occur in relatively few words, although some of the words occur often (e.g., eo in *people*). Four graphemes occur in numerous words (e.g., ea in *eat*). To add to the complexity, however, the consistent and common graphemes are often associated with other phonemes. For example, the grapheme ea is associated with /ɛ/ in *head*, /ɪ/ in *great* and, when followed by r, /ɜ/ in *learn*.

Despite this variability in English, research tells us that we should not avoid teaching consistent and common patterns. As far back as the First-Grade Studies in the 1960s (Bond & Dykstra, 1967), the research has been unequivocal: Learning the connections among letters and sounds is fundamental to proficient reading. The English orthography may not always be consistent, but proficiency in reading English will not occur without automaticity in connecting letters and sounds. This finding does not need to be reinvented or debated.

Research findings that hold promise for solving problems

Although research is clear that readers need to learn to connect letters to sounds, the quasi-regular orthography of English can seem like an insurmountable challenge for readers. Journalists, policy makers, and educators cite the below-basic performances of a significant portion of 4th graders on the National Assessment of Educational Progress (NAEP; National Center for Educational Statistics [NCES], 2023) as evidence that U.S. students are struggling to surmount this challenge.
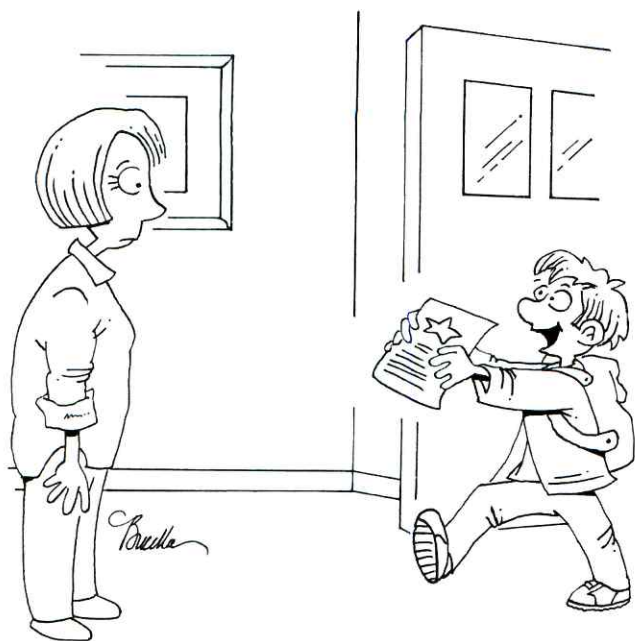
"I got a gold star for going the longest without looking at my phone in class."

Theory-building and research on statistical learning offers new perspectives on how more students can become proficient readers (Elleman, Steacy, & Compton, 2019). According to this model, students acquire and use statistical regularities in language to become more proficient readers (Seidenberg, 2017). Readers build connections between phonemes and graphemes through encounters with patterns in the words that they read. The more that students read, the more frequently they will encounter common and consistent patterns. This frequency, in turn, leads to increased automaticity in recognizing words with patterns when they encounter them in words.

This phenomenon explains the learning of both regular and irregular letter-sound relationships in English's quasi-regular orthography. Initially, readers may respond more slowly to words with regular or rule-based patterns (e.g., *bead, beat*) that have irregular neighbors (e.g., *head, great*) than to words without irregular neighbors (e.g., *must*). Mark Seidenberg (2005) described rule-governed and irregular words as points on a continuum of spelling-sound consistency. The frequency with which words appear in language and the volume of text students read work together to ensure successful learning of both rule-governed words and their irregular neighbors.

A critical aspect of the statistical learning position is that learning results from experiences with words rather than through rule memorization (e.g., when two vowels go walking, the first one does the talking). Lessons in which teachers review rules with students are insufficient. Students need to encounter examples of a pattern in actual words. Writing words with magnetic letters or on whiteboards can support connections, as can sorting words on cards (Georgiou et al., 2021). The words in students' text diets are also critical. In describing the process of statistical learning, Seidenberg (2005, 2017) did not mention decodable text (i.e., text containing the specific patterns students have learned), but he did describe the need for a large volume of reading to develop reading proficiency in English.

At present, the statistical learning model has not been applied to analyses of students' text diets and their effects on proficiency. This model offers insight, however, into why texts based on lesson-to-text-match (Stein, Johnson, & Gutlohn, 1999) — where decodability of texts is evaluated as percentages of words where all patterns have been taught in lessons — have not shown significant effects over other types of texts (Cheatham & Allor, 2012). At present, numerous questions have yet to be addressed in the statistical learning model, but the model holds promise for understanding the nature and volume of writing and reading required for students to become proficient readers of English.

Research findings that call assumptions into question

For several decades, an assumption underlying many policies and practices has been the existence of the 4th-grade slump (Chall, Jacobs, & Baldwin, 1990). According to this

The research has been unequivocal: Learning the connections among letters and sounds is fundamental to proficient reading.

view, 4th grade is when literacy demands increase, as students are required to read more informational text. The view is that, if students leave 3rd grade without strong foundational word recognition, they will do poorly as 4th graders and beyond (Chall, 1983).

Research points to the fallacy of the assumption that students first learn to read by grade 3 and, subsequently, read to learn (Pearson et al., 2020). Word recognition, vocabulary, and comprehension interact in multiple ways, from the earliest to most proficient stages of reading (Perfetti & Hart, 2001). Furthermore, literacy standards in the primary grades are higher today than in the early 1980s (Bassok, Latham, & Rorem, 2016), when the idea of the 4th-grade slump emerged (Chall et al., 1990). As evident in the Common Core State Standards (National Governors Association Center, 2010), the informational text that Jeanne Chall (1983) described as entering the curriculum at 4th grade is now part of the kindergarten experience.

Enactment of policies to increase foundational skills in the primary grades, such as retention of students in 3rd grade, does not appear to have appreciably changed the profile of American 4th graders. In fact, the most recent NAEP results (NCES, 2023) showed a decline in the percentage of 4th graders who attained the proficient standard and growth in the percentage who scored below basic since 2019. This recent decline likely reflects pandemic learning contexts but, in general, patterns on the NAEP have remained fairly consistent over the past 20 years with students' performances falling into three roughly equivalent groups of proficient, basic, and below-basic readers.

Journalists and policy makers often interpret these results to mean that students who perform at the below-basic level are unable to read (Hanford, 2019) or, at the very least, need to revisit phonological decoding and word recognition (White et al., 2021). But research examining the profiles of 4th graders scoring at the below-basic level shows it is not foundational word identification that challenges these students but rather vocabulary and meaning-making.

When Marsha Riddle Buly and Sheila Valencia (2002) studied the proficiencies of 4th graders who scored at the below-basic level on a standards-based assessment that was modeled after the NAEP, they found substantial variations in these students' profiles. A third of this group had higher performances on word identification and fluency measures

The depth and breadth of the research on reading and the complexities of life in the classroom, especially when the variation among students is substantial, means answers are neither simple nor easily translated to practice.

than in comprehension, while another 40% had relatively strong word-identification skills but read slowly, which compromised their meaning-making. Of the students scoring in the below-basic range, 17% were described as diligent readers who maintained attention to the text but whose word identification skills were not adequate for comprehension. Only 9% of students in the below-basic sample had sufficiently poor word-identification tasks to make sustained reading impossible.

In two decades since the Buly and Valencia (2002) findings, other researchers have confirmed the manner in which inadequate reading rates serve as an impediment to comprehension, especially when students are asked to read extended texts silently (Hayden, Hiebert, & Trainin, 2019). A pattern similar to that reported by Buly and Valencia has also been reported on the word-identification proficiencies of students scoring in the bottom third. For example, data from large samples of students on oral reading assessments, such as Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Center on Teaching and Learning, 2020), indicate that, while a portion of students in the bottom third (those in the 9th to 1st percentile) continue to have inadequate word-identification skills, students from the 35th to 10th percentile read with high levels of accuracy — 99% to 96%.

Furthermore, the words that challenge students scoring in the below-basic range are multisyllabic and rare in meaning (Hiebert, 2022), not the monosyllabic words that are typically the focus of phonics curricula. Because most English words (87%) are multisyllabic (Vousden, 2008), this aspect of orthography requires attention beyond the primary grades. The seemingly insurmountable task of reading words with multiple syllables is only made possible through English's morphological systems. Guiding students to recognize the root words in morphological families and anticipate that most words are members of morphological families will greatly benefit their reading development (Lee, Wolters, & Kim, 2022). Attention to the complexity of the English orthography and morphology and the extensive English lexicon must extend across the grades — not be limited to the primary ones.

Research that provides fine-grained analyses of the strengths and needs of students who perform at below-basic levels on state and national assessments challenges conclusions regarding NAEP results that "millions of kids can't read" (Hanford, 2019) and that the solution lies in revisiting decoding in subsequent grades (White et al., 2021). This research also points to productive directions for closing the reading gap. As Seidenberg (2017) states, "Readers become orthographic experts by absorbing a lot of data, which is one reason why the sheer amount and variety of texts that children read is important" (p. 92).

The value and limits of research

The science of reading perspective has engaged the attention of educators. This call for evidence-based practice has enriched my perspectives as a teacher, teacher educator, and researcher. However, the depth and breadth of the research on reading and the complexities of life in the classroom, especially when the variation among students is substantial, means answers are neither simple nor easily translated to practice. In some cases, answers are long-standing and clear, but, in others, research is in its early stages. In still other areas, long-standing assumptions can influence how we interpret research and how we talk about research findings. On numerous aspects of reading instruction, questions have yet to be addressed. An example of an especially urgent question is how much text constitutes the "sheer volume and variety" (p. 92) that Seidenberg (2017) describes as needed to become highly skilled and automatic in comprehending. Finally, as hypotheses are tested, scientific research can be expected to be refined and extended.

As is often the case in education, where choices of practices have consequences for students, educators will be bombarded with claims and counterclaims. It's important for educators to remain aware of what research does and does not reveal and to seek support from professional organizations as they engage their students in practices that support the high levels of literacy required in the digital global age. ∎

References

Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers.* National Academy of Education.

Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open, 2* (1).

Bond, G.L. & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2* (4), 5-142.

Buly, M. & Valencia, S.W. (2002). Below the bar: Profiles of students who fail state reading assessments. *Educational Evaluation and Policy Analysis, 24* (3), 219-239.

Center on Teaching and Learning. (2020). *DIBELS 8th Edition 2017-2019 Percentiles* (Technical Report 2001). University of Oregon.

Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Harvard University Press.

Chall, J.S. (1983). *Stages of reading development*. Harcourt Brace College Publishers.

Cheatham, J.P. & Allor, J.H. (2012). The influence of decodability in early reading text on reading achievement: A review of the evidence. *Reading and Writing, 25* (9), 2223-2246.

Elleman, A.M., Steacy, L.M., & Compton, D.L. (2019). The role of statistical learning in word reading and spelling development: More questions than answers. *Scientific Studies of Reading, 23* (1), 1-7.

Fry, E. (2004). Phonics: A large phoneme-grapheme frequency count revised. *Journal of Literacy Research, 36* (1), 85-98.

Georgiou, G.K., Savage, R., Dunn, K., Bowers, P., & Parrila, R. (2021). Examining the effects of Structured Word Inquiry on the reading and spelling skills of persistently poor grade 3 readers. *Journal of Research in Reading, 44* (1), 131-153.

Hanford, E. (2019, January 2). *Why millions of kids can't read and what better teaching can do about it*. NPR.

Hanna, P.R., Hodges, R.E., Hanna, J.S., & Rudorf, E.H., Jr. (1966). *Phoneme-grapheme correspondences as cues to spelling improvement*. U.S. Department of Health, Education, and Welfare.

Hayden, E., Hiebert, E.H., & Trainin, G. (2019). Patterns of silent reading rate and comprehension as a function of developmental status, genre, and text position. *Reading Psychology, 40* (8), 731-767.

Hiebert, E.H. (2022). When students perform at the below basic level on the NAEP: What does it mean and what can educators do? *The Reading Teacher, 75* (5), 631-639.

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.

Lee, J.W., Wolters, A., & Kim, Y.S. (2022). The relations of morphological awareness with language and literacy skills vary depending on orthographic depth and nature of morphological awareness. *Review of Educational Research, 93* (4), 528-558.

National Center for Education Statistics. (2023). *NAEP report card: Reading*. U.S. Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*.

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read*. National Institute of Child Health and Human Development.

Pearson, P.D., Palincsar, A.S., Biancarosa, G., & Berman, A. I. (2020). *Reaping the rewards of the Reading for Understanding Initiative*. National Academy of Education.

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). American Psychological Association.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (2020). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. In R. Ellis & G. Humphreys, *Connectionist psychology: A text with readings* (pp. 367-454). Psychology Press.

Seidenberg, M.S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science, 14* (5), 238-242.

Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it*. Basic Books.

Snow, C., Burns, M.S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. National Research Council.

Stein, M., Johnson, B., & Gutlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Remedial and Special Education, 20* (5), 275-287.

Vousden, J.I. (2008). Units of English spelling-to-sound mapping: A rational approach to reading instruction. *Applied Cognitive Psychology, 22* (2), 247-272.

White, S., Sabatini, J., Park, B. J., Chen, J., Bernstein, J., & Li, M. (2021). *The 2018 NAEP Oral Reading Fluency Study* (NCES 2021-025). U.S. Department of Education, National Center for Education Statistics.