Teaching Reading Is Rocket Science

WHAT EXPERT TEACHERS OF READING SHOULD KNOW AND BE ABLE TO DO

PAGE 4
Limiting Children to Books They Can Already Read
Why It Reduces Their Opportunity to Learn

By Timothy Shanahan

Fifty years ago, when I was becoming a teacher, reading instruction consisted of a ubiquitous classroom practice: placing students in instructional groups according to their reading level. These groups were sometimes known by various colors or animals. I distinctly remember the redbirds, the bluebirds, and the buzzards among the most popular appellations. Today, although groups are now labeled with letters (Level G, Level L, etc.) and ornithological monikers are out of fashion, assigning students to instructional groups according to their reading levels is still a common practice in classrooms across the country.

A recent survey aimed at identifying the most popular current programs used to teach reading found that one common feature of all the top sellers was that they organize their teaching around leveled books. Other recent surveys show that teaching reading with leveled books is on the increase and that teachers believe it is endorsed or supported by their state educational standards, though, in most cases, it is not.

But how effective is such teaching? Does it work? On the surface, those are easy questions. Leveled readers obviously work. Most American students are learning to read, at least at basic levels, and since most are being taught with leveled books, there must be some potency in the approach.

But the real question isn’t whether children can learn from leveled books, but whether such leveling confers any learning advantages. Might students do even better if taught with books they can’t already read so well? That’s the real question. In this article, I examine the research on leveled reading approaches and offer more effective ways that classroom teachers can ensure their students acquire the skills and knowledge they need to not only read a text but also comprehend it. First, I provide a brief history lesson in how we got here.

Teaching with Gradually More Difficult Texts

The idea of testing students to place them in different levels of text for instruction was first recommended more than 100 years ago, and an early survey indicated that 58 percent of primary-grade reading instruction was already being delivered in small ability-based groups and that 42 percent of the teachers were adjusting text levels to facilitate learning. They may not have referred to

Timothy Shanahan is a distinguished professor emeritus of urban education at the University of Illinois at Chicago and the founding director of its Center for Literacy. Previously, he was the director of reading for Chicago Public Schools. A past president of the International Literacy Association, he was the chair of the National Early Literacy Panel and a member of both the National Reading Panel and the English Language Arts Work Team for the Common Core State Standards Initiative. He writes about education at www.shanahanonliteracy.com.
Despite the current popularity of leveling, research evidence has not been especially supportive of the approach.

These texts as “leveled readers,” and no company that published basal readers had yet coined the term “guided reading,” but the practices of that time were markedly similar to those of today.

During the 20th century, research identified text features that correlate with reading comprehension, and publishers started to control these features to a degree previously unimagined. One reading program I remember from my childhood bragged that it never introduced more than one new word per page, and any word that was introduced was repeated 15 times over the following pages. That’s why those texts could be so mind-numbingly repetitive: “Oh. Oh. Oh. Look, Jane, look.” If learning to read meant learning words—and at least since the time of Horace Mann that’s been an idea held by many—then the accumulation of words gradually from selection to selection was how someone would best advance in learning to read.

But such tight readability controls made beginning reading texts so artificial that they eventually elicited an adverse reaction. The most remarkable of these reactions was the adoption of “whole language” policies by California in the 1980s. These reforms required that the texts used to teach reading not be designed for reading instruction, and severely limited the text revisions that could be made for pedagogical purposes. What this meant was that for a brief period of time, even the beginning reading materials got much harder, perhaps so hard for beginning readers that they represented a significant impediment to learning. If the old basal readers were easier than necessary, these new books were decidedly too hard for the beginners, and they provided teachers with little or no guidance on how to teach with texts that the older students couldn’t read successfully on their own. Exacerbating the effects of these harder books was that California’s policies simultaneously discouraged phonics and spelling—instruction that could have helped students to better read the challenging materials, and the term “whole language” often came to mean “whole class” instruction in many schools. Perhaps the thinking was, why group for instruction if nobody could read the books anyway?

It was in this environment that Irene Fountas and Gay Su Pinnell published their landmark book in 1996, Guided Reading: Good First Teaching for All Children. There was nothing terribly original in their presentation, but they rediscovered and championed a set of teaching procedures that in the not-too-distant past had been widely used to facilitate reading instruction. They recognized that texts varied in difficulty and that one could guide student progress successfully across a gradually harder progression of books. To do this successfully, they asserted that teachers would need to group children, matching different books to students based on their varied levels of reading. Their approach offered immediate relief for those beginning reading classes where easier books made sense, but even in the grades beyond, the shift was welcome because of the lack of any pedagogical support for teaching challenging books. Fountas and Pinnell’s approach, although reminiscent of earlier popular instructional practices, differed from them in one important regard: because of the burgeoning availability of high-quality children’s trade books, they could propose doing this without textbooks.

In the Fountas and Pinnell version of guided reading, teachers assess students to determine their reading levels and then assign them books that they can read with a high degree of accuracy and comprehension. Over time, if retesting shows improvement, the students are switched to more demanding books. When it doesn’t work so well, students may languish for long periods at their current levels. Such languishment has been enough of a problem that in the second edition of their book, they recommend moving students up sometimes even when the testing shows no evident improvement. (This seems to me like a judicious amendment to the original plan, but it raises the question about why these students can be expected to learn from the harder books when it is assumed that no one else would be able to.) Currently, this approach to reading, in which leveled books are matched to student reading levels for reading instruction, predominates in U.S. classrooms.

Determining Text Levels

There are basically two ways to determine how difficult texts may be and to set their levels: quantitative readability measures and qualitative judgments about texts. Although they approach the task differently, the purpose of both is to array texts on a continuum of difficulty.

The quantitative study of readability identifies text features that may affect comprehension and then tries to array these features in an algorithm that will allow accurate predictions of text difficulty. It turns out that accurate predictions can be obtained with only two text variables: vocabulary and sentence complexity. Such formulas are imperfect, but reasonably accurate. They aren’t able to make fine distinctions, and until recently, haven’t been able to measure beginning text levels very well. Nevertheless, quantitative readability algorithms are able to provide a largely reliable and accurate scientifically derived text gradient.

Still, it is important to remember that readability was not developed to match books to students in a way that would facilitate learning. Readability measures predict comprehension, not reading progress. The idea of using these kinds of measures to establish which books would best promote learning to read came later.

With the advent of computer technology, readability measurement has improved.11 The newer readability measures that have emerged are now widely used by researchers and publishers and were employed by the Common Core State Standards to specify
text level aspirations for the various grade levels. Despite all this, when teachers speak of “leveling” books, they are most likely referring to Fountas and Pinnell levels.

Researchers provide a useful history of the development of this qualitative leveling system.12 Basically, an early version of the approach was developed for use with Reading Recovery,13 a short-term reading intervention for first-graders who have difficulty learning to read, and Fountas and Pinnell refined and expanded this system to apply to texts from beginning readers through eighth-grade texts. Texts are evaluated by judges who place them on a multipoint continuum (from A to Z) based on 10 criteria: genre/forms, text structure, content, themes and ideas, language and literary features, sentence complexity, vocabulary, words, illustrations, and book and print features.14

No studies have evaluated the reliability of these judgments, but a couple of small studies suggest that the Fountas and Pinnell gradient correlates reasonably well with the better-validated quantitative readability measurements.15 Publishers have now leveled tens of thousands of books using this scheme. But given its complexity—the simultaneous qualitative evaluations of 10 factors with no explicit prioritization rules—it is unclear how accurate these levels may be (a point Fountas and Pinnell themselves make16). Clearly, this approach lacks the scientific rigor of the quantitative approaches and may result in varied book placements depending on who makes the judgments. But until more evidence is available, let’s at least for the sake of argument accept that these levels are sufficiently accurate to consider their use.

To sum up, there are two approaches to setting text levels—one based on a great deal of scientific evidence and one less well understood. Nevertheless, existing data suggest that both can place texts on a reasonable comprehensibility continuum, from easy to difficult. The problem is that research does not support the idea that either approach can identify from which texts students will learn best. The point of leveling is both to establish a text gradient and to place students in the appropriate text along that gradient. The latter is the issue to which we now turn.

**Book Levels That Promote Learning**

More than 70 years ago, Emmett Betts published an influential textbook on the teaching of reading.17 Betts claimed all readers have three reading levels: independent, instructional, and frustration. According to Betts, the independent level refers to texts that readers can handle on their own without assistance. Instructional-level texts are a bit harder, but not so hard that students can’t improve their reading from working with them under the guidance of a good teacher. And, frustration level? These books would be so difficult that learning would be unlikely even with supportive teaching.

Betts wrote that the way to determine these levels was to have students read from the books aloud and answer comprehension questions. Instructional-level texts, according to Betts, were those that could be read with 95-98 percent accuracy (in terms of word reading) and understood with 75–89 percent comprehension—the criteria that continue to be used today. Instructional-level texts generate small numbers of mistakes and misunderstandings, which can then presumably be addressed successfully through instruction and practice. Betts claimed that research supported the idea of matching books to students in this way to optimize learning. This instructional-level scheme is what is used today in most popular reading programs.

Instead of a steady diet of instructional-level texts, students should be reading a range of texts in their classrooms. It’s easy to understand why someone might propose (or adopt) such an approach. It is incredibly frustrating when students can’t read a text very well. At a time when teachers were limited to one grade-level text for reading, there would be plenty of students who wouldn’t be able to read it proficiently. Under those circumstances, teachers would gladly embrace the idea of teaching students at their instructional levels. But as gratifying as the idea of teaching students at their instructional levels may have been, there are legitimate questions about the degree of effectiveness of this approach. When there is so little to learn from a particular text, it is possible that progress will be needlessly slow moving.

Despite Betts’ original claims and the current popularity of leveling, research evidence has not been especially supportive of the approach. The study Betts referred to as the source of the instructional-level criteria was a doctoral dissertation of one of his students,18 and that study neither matched books to students for instruction nor evaluated learning. Betts’ doctoral student simply checked to see how many oral reading errors fourth-graders could make and still maintain 75–89 percent reading comprehension; that was the source of the 95–98 percent accu-
racy criterion. Years later, the researchers were questioned about the source of comprehension numbers, and they couldn’t remember from whence those had come.¹⁹ Not a very substantial basis for such a widely recommended instructional practice.

In the 1960s and 1970s, William Powell challenged Betts’ criteria, though he fully accepted the kind of evidence Betts had used to set them.²⁰ Powell thought Betts got the numbers wrong. To that end, he conducted studies in which children from grades 1–8 were evaluated in much the same way as in Betts’ doctoral student’s study. Powell found a couple of interesting things. He reported different instructional levels for different grades; that is, some children could tolerate more disfluency and still comprehend what they were reading. He also reported that some students could tolerate quite a bit of disfluency, suggesting Betts was placing students in books that were too easy.

Later, another study tested how well second-graders could read the books they were taught with, and then measured how much they learned. The researcher found that texts that could be read with about 85 percent accuracy and less than 50 percent comprehension led to the biggest learning gains. In other words, students learned more from books that were at their “frustration levels.”

Over the past few decades, there have been several direct tests of the instructional level, and these have all ended with one of two outcomes. Instructional-level texts either have provided no learning advantages or have done harm. One example of the latter is another study with second-graders.²¹ This study was the first randomized control trial of this practice. Students were tested and, using Betts’ criteria, randomly assigned to one of three treatments. One group worked with texts at their instructional levels, one worked with texts two grades above this, and the third worked with books four grades above. Students read in pairs, practicing reading fluency with a partner. At the end of the school year, the students placed in books above their instructional level had made significantly bigger learning gains than those placed in the books supposed to facilitate their learning. This study was later replicated with third-graders.²² Other studies again found big learning advantages from working with books at the children’s grade levels rather than reading levels.²³ Even students with learning disabilities have been found to obtain no benefit from these text placements.²⁴

Betts saw a problem—students being taught from books that many couldn’t read—and he proposed a solution, moving students to books that they could. Another solution, one he apparently didn’t entertain, was that teachers could adjust their instruction in particular ways to facilitate students’ interactions with these hard-to-read books. As a recent study found—this one with high school students—most students who were asked to read grade-level materials were able to learn more than those placed in the easier books.²⁵

Basically, what this research reveals is that limiting students to texts they can already read well reduces their opportunity to learn—by limiting their exposure to sophisticated vocabulary, rich content, and complex language. With knowledge of the research on effective reading instruction, skilled teachers can facilitate students’ productive interactions with harder text.

But what has happened since states started requiring that students be taught to read more challenging text? In 2010, the majority of states adopted the Common Core State Standards. These standards, for the first time ever, set text levels that students were supposed to be able to read by the time they reached particular grade levels. The levels were set high to enable students to reach levels of proficiency that would ensure later life success. States may have thought they had accomplished something pretty big by adopting those standards, and likewise district administrators may have thought they had dealt successfully with the complex text requirements when they purchased new textbooks matched to these new requirements. However, according to national surveys,²⁶ all that has happened is that teachers, seeing that more of their students are now struggling with these newer texts, have increasingly relied on the idea of instructional-level teaching, and more and more are placing students in below-grade texts for reading instruction.

A More Effective Approach

As a teacher, I always taught with leveled books and worked hard to match texts to students in the ways described here. However, as I’ve learned of the research, I’ve gone a very different way—except with beginning readers. I know of no studies with kindergartners or first-graders showing that they should be trying to read particularly demanding texts (in contrast, there is a benefit to teachers reading aloud demanding texts to build young children’s knowledge and vocabulary).
With research on effective reading instruction, teachers can facilitate students’ productive interactions with harder text.

I’ve come to think of reading as the ability to make sense of the ideas presented in text, by taking advantage of the affordances and overcoming the barriers included in the text. Learning to read means becoming aware of these text features—and learning how to deal with them. Instructional-level texts are usually too easy to provide students with the opportunities to confront text features that they cannot already manage.

Affordances or barriers—and these are basically the same things—are features that authors build into their texts to facilitate communication. A particular text feature serves as an affordance if it does that, and—for some readers—it may serve as a barrier to understanding. For example, an author might aim for clarity and accuracy through apt diction, and for readers who know the meanings of the words so chosen, this can be a powerful text affordance. But for readers with more limited vocabularies? That potential affordance may become an unfortunate barrier for them.

It’s not, however, as simple an equation as instructional-level theory makes out. It is not that some students have better vocabularies, so we should let them work with the relatively difficult books (the ones with the rich content and complex language), and that the other students—who who know fewer words—should be segregated into easier and more limited texts. That approach can have some unfortunate implications for students who are minorities and those from low socioeconomic backgrounds.27

What if, instead of segregating them into what some students call the “stupid books,” we placed them in books with demanding vocabulary and taught dictionary skills, use of context, and morphology? What if we taught them when it was essential to figure out an unknown word meaning and when they might be able to soldier on successfully without doing that?

And, of course, vocabulary is just one of many such text features. Studies have long shown that teaching students how to disentangle the grammar of some sentences,28 how to take advantage of the cohesive links across a passage,29 and how to identify and use a text’s organizational structure30 all can improve reading comprehension. Teaching students to negotiate these features of a text only makes sense if students are to be confronted by challenging texts, and none of them have value for students reading, what for them, are easy books.

If we are serious about raising reading achievement, we must think hard about whether it makes sense to continue teaching students to read books they can already understand so well. These easier books make learning unnecessary and, without adequate challenge, may even drain the fun out of learning. That doesn’t mean that every selection used for reading instruction must significantly challenge students, only that grade-level texts should be part of the instructional mix.

Instead of a steady diet of instructional-level texts, students should be reading a range of texts in their classrooms. Some proponents of leveled reading claim they too support this idea, but they propose that instructional-level texts should be the focus of small-group teaching. I recommend just the opposite, having students reading really demanding texts when the teacher is close by and ready to help, and less demanding ones when on their own or when a teacher just isn’t going to be available.

But this is not just an avenue to higher achievement (though research suggests that it could be), it is also an issue of equity. If fourth-graders are taught from a second-grade book, when will they have the opportunity to confront the language and ideas of fourth-grade books? This is a cruel math problem that tells students they are best served by books that don’t match their interests, their curiosity, or their social aspirations. Leveled reading emphasizes students’ current limitations, rather than increasing their possibilities, especially for the least advantaged of our students. We can do better.

Endnotes


16. Fountas and Pinnell, “Guided Reading.”


(Continued on page 39)
Teaching Reading Is Rocket Science
(Continued from page 9)


The Reading League
(Continued from page 12)


7. Blackman et al., “Intensive Reading Remediation in Grade 2 or 3.”


Building Knowledge
(Continued from page 21)
that students living in poverty forget the mental health and support services they need, and engaging families in their children’s learning. But if we don’t also give students access to knowledge of the world during elementary school, we’ll never achieve the result we want: a system that equips all students to lead productive and fulfilling lives and carry out their responsibilities as members of a democratic society.

Endnotes


11. D. L. Kirp, “Ending the Curse of Remediial Math,” New York Times, June 10, 2017. More than two-thirds of community college students need remedial math or English classes, which cost them money but don’t earn them any credit. Only 15 percent of these students get their degrees on time.


