

Why Children Should Be Taught to Read with More Challenging Texts

by Timothy Shanahan

A major instructional tenet of the past 70 years has been that students will make progress in reading only if they are taught with relatively easy texts—texts at their so-called “instructional levels.” This theory has been embraced by both reading educators and special educators and is widely honored in U.S. classrooms and in remedial interventions. The argument has been that learning will be disrupted if teachers try to teach using texts that elicit too many word-reading errors or that students may not fully understand from the start. However, this nearly universal assumption turns out to be completely unsupported by evidence.

Of course, the idea of placing students in texts in ways that would facilitate their learning to read has always been plagued by technical inadequacies (Klare, 1974–1975; Nilsson, 2013), though these problems evidently have not been enough to dissuade teachers trying to match children to text. For instance, even the most scientifically rigorous readability formulae have difficulty distinguishing text levels, such as determining whether a text is best for second- or third-grade readers. That’s why the most widely used readability schemes estimate grade placements in bands (e.g., grades 2–3, 4–5, 6–8) and why, even with this, there is so much overlap in the grades the texts are assigned to; many texts may be assigned to two or more of these ranges.

Likewise, the measures used to estimate a child’s reading levels have been dubious, as well. There have been controversies over what counts as errors and the ability of teachers to accurately make these judgments on the fly as they listen to children read. Basically, these text and student measures are able to provide no more than rough guestimates, neither being precise nor reliable enough for accurate individual decision making, and neither having been validated for the purpose of matching children to books in a way that would facilitate learning. There have been many articles exploring these measurement problems (Pikulski & Shanahan, 1982; Zamanian & Heydari, 2012). This article, instead, will consider the validity of the “instructional level” construct. That is, if we match texts to students in the ways that have been recommended, is learning actually facilitated?

Early Responses to Student Struggles with Texts

Throughout the first half of the 20th Century, retaining students at lower grade levels was seen as the solution to the age-old problem of students being unable to read their textbooks adequately. The way to protect against too great a mismatch between student and book was to prevent students from progressing up the grades (and, up the textbook levels) unless their growth in reading justified the advancement. Of course, it would have been possible to simply use texts that were, for example, at a third-grade level in a fifth-grade class. Educators at the time must have been discomfited by this alternative, as there are no contemporary references to that idea as

far as I can tell. The result of this practice of retention, according to various teacher memoirs of the time, was that increasingly older and larger students were using texts that were far below their maturity or interests, a situation that, not surprisingly, led to serious disciplinary problems and disaffection.

The result of the practice of retention, according to various teacher memoirs of the time, was that increasingly older and larger students were using texts that were far below their maturity or interests, a situation that led to serious disciplinary problems and disaffection.

During this period, psychologists were exploring the possibility of measuring the readability or comprehensibility of texts. Philosophers had long opined on the idea that texts varied in their depth or complexity, but until the 1920s and ’30s there was no objective or scientific way of teasing out these differences. That began to change with the development of readability formulae that allowed texts to be placed on a continuum of difficulty roughly corresponding to grade levels.

This innovation in the measurement of text difficulty opened the possibility of matching students to text scientifically. By the late 1930s, educators began to speculate that it would be possible to match text difficulties not just to grade levels, but to individual students’ reading levels. The idea that this practice could facilitate learning grew in popularity, though there was not yet any forcefully articulated theory or technology that could bring this notion to fruition.

Instructional Level Theory

That changed in 1946 with the publication of Emmett Betts’ *Foundations of Reading Instruction*, which was to become the major reading-education textbook for teachers in that era. Betts not only argued that learning was facilitated by placing students at their reading levels, but he also described it as a research-based approach and provided a set of operational criteria that could be used to match students with appropriate texts. As Betts stated, “Maximum development may be expected when the learner is challenged but not frustrated” (Betts, 1946, p. 448). Over time this idea gained adherents, and after the publication of Irene Fountas and Gay Su Pinnell’s *Guided Reading* in 1996, what is known as “leveled reading” became the dominant approach in U.S. classrooms (Fountas & Pinnell, 1996).

Continued on page __

Betts claimed that readers have three levels of performance, all linked to how closely the demands of a particular text correspond to their skills. The independent level referred to texts that readers would find easy enough to read and learn from on their own, with no teacher assistance needed. Texts at an instructional level would be a bit harder, requiring some teacher guidance if the student was to learn from them. And frustration level texts were those presumed to be too challenging from which to learn even with the support of a good teacher. (Against the background of a period in which Freud was king, avoiding frustration was a high priority for psychic health.)

Betts further asserted that these reading levels could be determined by examining students' accuracy in recognizing words and their degree of comprehension. He claimed that students learned best from texts in which their oral reading accuracy was in the range of 95 to 98% and their reading comprehension (on a cold read with no teacher assistance) was 75 to 89%. Betts cited as his source for these criteria the dissertation of one of his doctoral students, Patsy Aloysius Killgallon (1942).

How did Killgallon go about this research? One might assume Killgallon matched students with various texts and found that more learning occurred when students were working with texts at their instructional levels. Surprisingly, that was not at all what the study considered. Killgallon's dissertation, which has remained unpublished except for some brief excerpts (Shanahan, 1983), started from the premise that for students to learn from a text, they had to be able to read it with 75 to 89% comprehension. The source of this premise is unknown; when interviewed years later Betts and Killgallon could no longer remember from whence this criterion had come (Beldin, 1970). The study merely found that children who read with less than 95% accuracy usually failed to accomplish the required 75 to 89% comprehension outcome; which is the source of the widely used standard for accuracy. Killgallon did not explore the impact of different reading "levels" on learning but simply correlated the number of oral reading errors with a target comprehension level, based on results from a small number of fourth graders.

Killgallon did not explore the impact of different reading "levels" on learning but simply correlated the number of oral reading errors with a target comprehension level, based on results from a small number of fourth graders.

Research into Matching Texts to Students' Reading Levels

Throughout the 1950s, "instructional level" was widely recognized as more theory than proven fact. An early empirical attempt to determine the effects of instructional level placements on children's learning was carried out by J. Louis Cooper

in 1952. This ambitious study pre- and post-tested more than 800 students in grades 2 to 6 from eight different schools, using two reading achievement tests to determine each student's instructional level. Cooper then monitored the texts that students were actually taught in, hoping to determine which student-book matches resulted in the greatest learning (Cooper, 1952).

Unfortunately, in practice student ability was totally confounded with book placement; that is, teachers placed the best readers (who also had the highest IQ scores) in books at their independent levels and assigned the lowest readers to what were, for them, the most difficult books relative to their abilities. Essentially, Cooper found that the children who made the biggest learning gains were the ones who could, from the start, already read their instructional books perfectly (in other words, there would be nothing to learn in these books). He himself concluded that this was meaningless. Nevertheless, this study illustrates why teachers might conclude that particularly easy book placements would lead to the most learning: The best readers are most likely to be placed in relatively easy texts and to make the best learning progress, too. This relationship is obviously not a causal one, but the pattern may encourage an assumption of causation.

Since that first failed attempt at validation of the instructional level construct, there have been several additional attempts to evaluate whether such text placements facilitate learning. But even replications of the original Killgallon study, which merely linked oral reading performance and text comprehension, have not been particularly reassuring. For example, Powell replicated the Killgallon study and concluded that students could often comprehend text well despite evidencing many more oral reading errors than Betts' criteria prescribed. That would suggest that Betts' benchmarks were placing students in texts that were too easy. And, correlational studies have not been encouraging either, reporting that frustration-level book placements were more associated with learning than instructional level ones (Powell & Dunkeld, 1971).

The randomized controlled trial is the gold standard for validating the effectiveness of an instructional practice, and in those kinds of studies instructional-level theory has fared poorly, as well. For example, in a study of second graders, children's instructional levels were identified using Betts' criteria, with children randomly assigned to one of three treatment groups (Morgan, Wilcox, & Eldredge, 2000). One group worked at their instructional level and two others were placed in texts that were either two or four grade levels above their instructional levels. Student learning was then monitored across the school year to determine if these placements provided any learning advantages. Both frustration-level groups outperformed the students who were taught at their instructional levels. This study has been replicated with third graders as well (Brown, Mohr, Wilcox, & Barrett, 2017). Other experimental studies—conducted with learning disabled students with Individualized Education Programs (O'Connor, Swanson, & Geraghty, 2010), and with second, third and sixth graders (Homan, Hines, &

Kromrey, 1993; Kuhn et al., 2006)—have either concluded that the instructional level offered no advantage or that it actually resulted in lesser amounts of student learning.

Reading is the ability to make sense of text, and that means being able to negotiate any barriers to understanding that texts may include.

Learning from More Challenging Texts

Why such consistently negative results? Reading is the ability to make sense of text, and that means being able to negotiate any barriers to understanding that texts may include. Accordingly, if students are working with texts that they can already read quite well—a description that certainly applies to instructional level texts—there is little opportunity for learning since the students can already negotiate the vocabulary and other features of that text. Students taught from a steady diet of relatively easy texts may make some progress, but not as much as would be possible with more complex texts, since the easier texts would provide fewer opportunities for dealing with sophisticated vocabulary, morphology, complex syntax, subtle cohesive links, complicated structures, and richer and deeper content.

Of course, reading comprehension entails the use of prior knowledge, the knowledge that readers already have prior to reading with a particular text. The more a reader already knows about the information presented in a text, the better the comprehension is likely to be. Instructional level placements, since they emphasize relatively high initial reading comprehension, discourage the use of texts that present much information that students do not already know. This both reduces the opportunity to learn new content and also limits what students can be taught effectively about how best to make sense of texts that present much unfamiliar information.

Still, it would be foolish to conclude that facilitation of learning requires no more than that we place students in more demanding texts. That is a necessary condition, but possibly not a sufficient one. Several of the experimental studies already cited that placed students in markedly harder text for instruction also engaged the students in extensive fluency work—reading the texts multiple times orally with guidance. In other words, though the students may have initially had difficulty reading these instructional texts, by the time they finished, their performance levels with these texts had advanced markedly.

But fluency practice is just one of many scaffolds or supports that teachers can provide to students to help them to gain understanding of complex texts. One frequent barrier to text comprehension is that readers may lack the background information or content knowledge that would allow them to gain full understanding of a text. Authors make assumptions about what their readers will know about a topic or event, but these assumptions do not always match with the actual knowledge that readers may bring to the text. Schools could avoid the possibility of this kind of mismatch by providing students with

texts that relate to knowledge they have previously acquired through the curriculum. Sometimes young readers have relevant background knowledge, but they fail to apply it when they read. Scaffolding in such situations entails encouraging students to think about the related knowledge before and during the reading. However, students need to learn to make sense of texts even when they do not have a lot of specific background knowledge, and teachers can introduce them to strategies that can help in those situations, too—for instance, drawing on analogous situations they are familiar with, or seeking additional information from outside the text.

Additionally, scaffolds may help students with unfamiliar vocabulary and support them in making sense of the linguistic or conceptual demands of a text. Strengthening students' abilities to parse sentences, make accurate cohesive links, and analyze the organizational plan or structure of a text can boost comprehension. Instruction in comprehension strategies such as summarizing, self-questioning, monitoring, or visualizing can help, too, as long as the strategies are attached to understanding the specific content of a text and not pursued as ends in themselves.

Any text feature or characteristic used by an author to communicate information can stymie some readers and, thus, can become the focus of potentially useful instructional scaffolding or support. Of course, the actual supports provided by a teacher in a given instance will depend upon the specifics of the text and whether those features are actually disrupting a student's comprehension. Table 1 provides a partial list of some of the possible categories of scaffolds and supports that can be provided to readers to allow them to gain a more complete understanding of a text.

TABLE 1. Categories of Scaffolds or Supports that Teachers Can Provide to Readers

- Decoding/fluency supports
- Enhancements of prior knowledge
- Vocabulary supports
- Syntax guidance
- Coherence links
- Genre guidance
- Text structure/organization supports
- Author's tone supports
- Literary device assistance (e.g., metaphors, symbols, allusions)
- Text features assistance (e.g., italics, bolding, bullets)
- Data-presentation device (e.g., tables, charts) assistance
- Comprehension strategy guidance
- Motivational encouragement

The Role of Instructional Support

From the research conducted so far, it is impossible to recommend a particular degree of text difficulty with which students should be dealing within instruction. Obviously, the

Continued on page

harder a text is for a student, the more there is to learn, which is a positive thing. But it is also clear that the harder a text is relative to the current reading abilities of the students asked to read it, the greater the instructional support needed for success. The appropriateness of text challenge levels is probably less a matter of “how hard is the book for the student” and more an equation that would consider both this gap between student and text and the degree of support that teachers are willing and able to provide. The harder the book is for the student, the greater the instructional support needed for success.

The harder a text is for a student, the more there is to learn, which is a positive thing. But it is also clear that the harder a text is relative to the current reading abilities of the students asked to read it, the greater the instructional support needed for success.

In one study, the successful students *ended up* with what we have traditionally called an instructional level (Ehri, Dreyer, Flugman, & Gross, 2007). That is, students were taught from frustration level books, but by the time they finished working with one, they could read it with 98% accuracy and very high comprehension (these were first-graders so the major challenges were with the decoding). There is not enough research of this kind to mandate such an instructional approach, but I find it provocative. With this approach, both teachers and students could easily see the difference between where the students started with a text and how they ended up, something almost impossible to discern when students are placed in relatively easy materials. As they gain greater knowledge of the content and vocabulary that they are reading about, work through the confusing or complicated linguistic or textual demands, and develop fluency with the particular decoding requirements of the text, students should be able to read that text with high proficiency, at which point it would be time to move on to another text.

Another concern is whether it makes sense to place beginning readers in difficult texts. Various theories suggest that it might be wiser to start beginners out more gradually, lest they become overwhelmed. I suspect the issue is not so much the degree of challenge as what aspect of the text is challenging. Beginning readers struggle mostly with issues of decoding, and the texts used to teach them are often constructed to provide decoding support in a plethora of ways. The options include decodable texts, which ensure that a high percentage of the words can be decoded using the skills mastered to that point; texts with controlled vocabulary, providing children with a severely limited but gradually increasing collection of words; and texts with orthographies and printing techniques that give

young readers cues to pronunciation, through pronunciation keys such as those provided in dictionaries or the assignment of different colors to the letters associated with particular letter sounds. For beginners, more challenging text would usually use vocabulary with less repetition and a greater multiplicity of spelling patterns, which may slow these beginners’ development of proficiency.

Accordingly, no state has established text complexity standards for kindergarten or first-grade readers, and no instructional programs, to my knowledge, have ramped up text difficulties at these levels. I think this caution is prudent. Make sure that children have solid foundational skills in decoding—say through a high first-grade or a beginning second-grade level—before increasing the complexity of the texts used to teach reading. Several studies have shown that second grade is not too early for students to deal with more complex text successfully, a level by which those basic decoding skills should be well in hand.

The caution given here for young readers also would make sense for older readers who still decode like kindergartners or first graders. Don’t worry about taking students beyond “instructional level” texts until they are able to decode as well as a successful first-grader. I am sometimes told that this is still too early for learning-disabled readers because it might lead them to guess at rather than decode words in the harder texts. However, one study showed that learning-disabled children in grades 3 through 5 who read at a beginning grade 2 level or lower gain no advantages from being limited to books at their instructional level (O’Connor, Swanson, & Geraghty, 2010). Restricting students to easier materials usually means preventing them from dealing with content at their age- or maturity-levels and may serve to isolate these children from their social peers. These students are also aware that they are being relegated to the “dumb books,” with serious consequences for their self-esteem. And while the theory is that students will gradually make their way up the ladder of text complexity by reading voluminously at their own levels, the fact is that many children who begin at lower levels remain permanently behind. Without access to the more sophisticated concepts and complex vocabulary that their peers are being exposed to, they have no opportunity to acquire the knowledge and skills that could enable them to catch up.

A final concern oft expressed by teachers and parents is that teaching students from supposed frustration-level texts will be harmful to their motivation. Even the designation “frustration level” suggests that the damage might be less to learning and more to engagement. It is certainly possible that students would be discouraged by consistent placements in texts they will struggle to read. However, there are countervailing possibilities as well. For instance, research finds challenge itself to be motivational (Killeen, 1994). Also, studies have failed to link text complexity with lowered motivation or misbehavior (Fulmer & Tulis, 2013; Gambrell, Wilson, & Gantt, 1981), and it is evident that even good readers frequently choose harder texts when reading independently (Donovan, Smolkin, & Lomax,

2000). The motivational field of reading consists not just of the text, but also the explanations, instruction, support, and scaffolding provided by teachers—and most importantly, the provision of the knowledge required to understand the text. These may be sufficient to offset any sense of being overwhelmed that might occur if students were reading on their own.

It is certainly possible that students would be discouraged by consistent placements in texts they will struggle to read. However, there are countervailing possibilities as well. For instance, research finds challenge itself to be motivational.

Balancing Challenges with Supports

Contrary to long-standing assumptions, research has not supported the idea that there is a particular level of text with which students should be taught. Learning to read means learning to overcome the barriers and to exploit the possibilities of written language, and texts that students cannot already read well provide the greatest opportunity for helping them achieve that goal. But while it is reasonable to teach students with challenging texts, it is also essential that they develop proficiency with and derive knowledge from each text they work with. For that to happen, teachers need to provide students with scaffolding and support by—among other things—building the knowledge, vocabulary, and facility with features such as syntax and structure that are required to glean meaning from text.

References

- Beldin, H. O., (1970). Informal reading testing: Historical review and review of the research. In W. K. Durr (Ed.), *Reading difficulties: Diagnosis, correction, and remediation*, (pp. 67–84). Newark, DE: International Reading Association.
- Betts, E. A. (1946). *Foundations of reading instruction*. New York, NY: American Book Company.
- Brown, L. T., Mohr, K. A. J., Wilcox, B. R., & Barrett, T. S. (2017). The effects of dyad reading and text difficulty on third-graders' reading achievement. *Journal of Educational Research*, 111(5), 541–553.
- Cooper, J. L. (1952). *The effect of adjustment of basal reader materials on reading achievement*. Unpublished doctoral dissertation, Boston University.
- Donovan, C. A., Smolkin, L. B., & Lomax, R. G. (2000). Beyond the independent-level text: Considering the reader-text match in first graders' self-selections during recreational reading. *Reading Psychology*, 21, 309–333.

- Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44, 414–448.
- Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.
- Fulmer, S. M., & Tulis, M. (2013). Changes in interest and affect during a difficult reading task: Relationships with perceived difficulty and reading fluency. *Learning and Instruction*, 27, 11–20.
- Gambrell, L. B., Wilson, R. M., & Gantt, W. N. (1981). Classroom observations of task-attending behaviors of good and poor readers. *Journal of Educational Research*, 74, 400–404.
- Homan, S. P., Hines, C. V., & Kromrey, J. D. (1993). An investigation of varying reading level placement on achievement of chapter 1 students. *Reading Research and Instruction*, 33(1), 29–38.
- Killeen, P. R. (1994). Frustration: Theory and practice. *Psychonomic Bulletin & Review*, 1, 323–326.
- Klare, G. R. (1974–1975). Assessing readability. *Reading Research Quarterly*, 10, 62–102.
- Kuhn, M. R., Schwanenflugel, P. J., Morris, R. D., Morrow, L. M., Woo, D. G., Meisinger, E. B., Sevcik, R. A., Bradley, B. A., & Stahl, S. A. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research*, 38, 357–387.
- Morgan, A., Wilcox, B. R., & Eldredge, J. L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *Journal of Educational Research*, 94, 113–119.
- Nilsson, N. L. (2013). The reliability of informal reading inventories: What has changed? *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 29(3), 208–230.
- O'Connor, R. E., Swanson, L. H., & Geraghty, C. (2010). Improvement in reading rate under independent and difficult text levels: Influences on word and comprehension skills. *Journal of Educational Psychology*, 102, 1–19.
- Pikulski, J. J., & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J. J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading* (pp. 94–116). Newark, DE: International Reading Association.
- Powell, W. R., & Dunkeld, C. G. (1971). Validity of the IRI reading levels. *Elementary English*, 48, 637–642.
- Shanahan, T. (1983). A critique of P. A. Killgallon's study: A study of relationships among certain pupil adjustments in reading situations. In L. Gentile, M. L. Kamil, & J. Blanchard (Eds.), *Reading research revisited*, (pp. 577–582). Columbus, OH: Merrill.
- Zamanian, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43–53.

Timothy Shanahan, Ph.D., is Distinguished Professor Emeritus at the University of Illinois at Chicago and Founding Director of the UIC Center for Literacy. A former Director of Reading for the Chicago Public Schools, Shanahan is author/editor of more than 200 publications on literacy education. His research emphasizes the connections between reading and writing, literacy in the disciplines, and improvement of reading achievement. His blog, www.shanahanonliteracy.com, is widely read.