

Text Complexity

By

Dr. Timothy Shanahan

University of Illinois at Chicago

Member, National Reading Panel; President, International Reading Association, 2006

Chair, National Literacy Panel and National Early Literacy Panel

All texts are not equal. Some texts are harder to read and comprehend. Just as there are individual differences in children, there are individual differences in the texts that we ask children to read. Some children read better than others, and there are a variety of reasons for these differences in reading abilities. The same is true of texts. Some texts are easier or harder, and there are several reasons for these differences.

The purpose of this essay is to describe what factors cause texts to differ and to explore the relationship of text difficulty and children's learning.

What makes text complex?

Since the 1920s, there has been interest in measuring text difficulty or readability (Lively & Pressey, 1923). The basic purpose of these measures has been to try to place texts along a continuum or scale, from easiest to hardest. The most straightforward way to do this would be to have a group of people read the various texts and complete comprehension tests based on each one. The texts that most people understood would be rated as easy, and the ones that fewer people comprehended or that had fewer questions answered correctly would be rated as more difficult. Readability estimates, however, are not quite this direct.



It would be too costly and inconvenient to perform the above-explained process for every text, of course, so a second fundamental premise of readability measurement is that the measure must specifically summarize only text features or characteristics, and not how easily read the text has been in the past. On the basis of this enumeration of text features, the formula has successfully predicted reader comprehension for texts.

Readability measures have evolved since they first appeared. Now, most readability formulas only measure two factors: word complexity and sentence complexity (Klare, 1984). Thus, measuring the readability or complexity of a text involves an evaluation of the words: counting the average number of letters or syllables, checking the frequency of the words (common words are usually easier than rare words), or considering how abstract or concrete the words may be. Similarly, average sentence or clause lengths, as well as other measures of sentence or grammatical complexity, are included. These word and sentence measures are then combined into mathematical formulas that are used to predict actual reading comprehension. If a formula does well with some set of texts, then it is assumed that it will do equally well with other texts that have not been actually tested with readers.

Text Complexity

Readability measures often try to express student performance in terms of grade levels. (Measures that do not use grades usually provide some kind of conversion chart to show how to translate the readability score to a grade level.) The idea is that the readability formula is revealing the reading level at which readers would be able to understand the text. Thus, a fifth-grade text would be readable or comprehensible to the average fifth-grade reader.

Early on, these formulas improved our ability to predict text difficulty, but not by very much (they were able to anticipate about 35 percent of the variation in actual reading comprehension performance). Over the years, readability measures improved, and most of them are now able to explain about 50 to 60 percent of the comprehension variance. One recent approach, the implementation of Lexiles, is able to explain about 90 percent of reading comprehension differences (Smith, 2012).

Why aren't these measures able to do better than that? Why don't they predict 100 percent of the variation in reading performance? There are two basic reasons for this. The first has to do with the nature of the measures of themselves. Remember, only word and sentence difficulty are being considered. There are other text features that may impact complexity, but these are not accounted for directly in these formulas. These measures do well because other unmeasured text features tend to be correlated with the word and sentence challenges which are measured in the formulas. Thus, texts that have simple words and sentences will usually be about simpler concepts and will use simpler or more straightforward organizations. It is possible, however, that some of these other features will make a text either easier or harder for readers to understand than predicted. Accuracy is important, but so is efficiency (adding other variables that are harder to measure only improves accuracy a small bit, and these other features are much harder and more time-consuming to measure).

What are some of the other text complexity factors—

beyond sentence complexity and vocabulary sophistication—that are likely to influence reading comprehension? One such factor is the complexity of the ideas themselves. Some topics (rocket science, brain surgery, the infield fly rule), are more complicated, sophisticated, or subtle than others. Also, ideas connect with each other across a text through a feature called coherence. Readers have to interpret pronouns,

synonyms, and other connectors to make sense of a text. The greater the distance across ideas, the more varied the connections, and the more competition for a link (if the pronoun “he” is mentioned, how many male characters could be the topic of discussion?), the harder the text will be. Text structure or organization matters, too. Ideas can be arranged across text in many ways (e.g., time sequence, causation, lists, comparison). How straightforward the structure is will help to determine

how complicated the text will be to readers. Some texts may be more-or-less complete, including sufficient background information and context to help readers to interpret the ideas; other, more difficult texts may require that readers bring a lot of this kind of information to the text. Finally, texts may use various literary devices (e.g., irony, repetition, metaphor) or data presentation tools (e.g., tables and charts), adding to the interpretive burden.

Another reason for the imperfect measurement of readability formulas has to do with what we are trying to predict: reading comprehension. Readers differ in what they bring to a text in terms of background knowledge, motivation, abilities to cope with ambiguity, and so on. Even if two students have the same “reading level,” they may perform differently with a particular text, which would reduce the accuracy of the text-level prediction. Reading comprehension will always be the product of two factors: the text and the reader. Readability only measures the variations in the text; reader variations will present variability that the readability measure cannot capture specifically.

Readability only measures the variations in the text; reader variations will present variability that the readability measure cannot capture specifically.



Education

Text Complexity

Relationship of Text Complexity and Learning

Readability formulas examine text complexity features to predict reading comprehension. What connection does text complexity have to learning to read? Scholars have long claimed that the match of text difficulty and a reader's ability to read will be an important determinant of learning (Betts, 1946).

The original idea of this reader/text match concept is that it is important to place students in texts that are not too easy (if the students can read a text easily without help from a teacher or someone else, then there would be little opportunity to learn from that text) or not too hard (if the students struggle too much to make sense of a text, there would be a lot to learn, but they might be overwhelmed or frustrated, or the teacher might find it difficult to provide sufficient support to promote learning).

From this basic conception came the idea that there are independent, instructional, and frustration levels of performance.

Independent level refers to texts that students can read successfully without help; frustration level refers to those texts that are so hard for the students to comprehend that it would be difficult for them to learn; and instructional level would be the optimum level that would provide an opportunity for learning, but without being too hard or too frustrating.

As convincing and reasonable as this basic notion of an instructional level may be, actually determining which texts students would learn the most from requires some knotty measurements. The most straightforward approach to the problem would be to test children in terms of their ability to read texts at particular readability levels and then to randomly assign those children to reading instruction that employs texts written at different levels. Thus, some groups would spend their time being taught from easy (that is, easy for them) texts, and others would be taught from relatively harder materials. The results of such experiments would

reveal the student/text match that leads to the greatest amount of learning.

Unfortunately, there have been few such studies, and instructional practice has usually relied upon assertion rather than empirical evidence. The most influential claims have indicated that students needed to be taught from texts they could read with 95 to 98 percent oral reading accuracy and with a reading comprehension of 75 to 89 percent (Betts, 1946). More recent claims for optimum text complexity levels are based on minor adjustments to Betts' original assertions, even though we now know that Betts made his numbers up without evaluating them against student learning (Shanahan, 1983).

The few studies that do exist (Morgan, Wilcox, & Eldredge, 2000) suggest that students do better—that is, learn more—with texts that are somewhat harder than Betts claimed, and there is clear evidence that

The most straightforward approach to the problem would be to test children in terms of their ability to read texts at particular readability levels and then to randomly assign those children to reading instruction that employed texts written at different levels.

school texts have been getting easier and easier over a period of about seventy years for grades 3 to 12 (Chall, Conrad, & Harris 1977; Hayes, Wolfer, & Wolfe, 1996). Past efforts to place students in texts have striven to protect students from confronting too much difficulty, and yet, such protections also place an important limitation on opportunities to learn. "Just as it's impossible to build muscle without weight or resistance, it's impossible to build robust reading skills without reading

challenging text" (Shanahan, Fisher, & Frey, 2012, p. 58).

That suggests that schools should be striving to place students in more complex texts than in the recent past. There certainly is some justification for such an approach, but it would be wise to be cautious in this regard. First, it is important to remember that there is a difference between beginning reading and later reading. Initially, children have to figure out the decoding system. Part of this is learned from explicit teaching in phonological awareness and phonics, but part is learned from abstracting the alphabetic principle and how it works from what we read



Education

Text Complexity

(Venezky & Johnson, 1973). If early reading texts are very complicated, then it is likely that student progress in beginning reading will be slowed down. Readability measures used with beginning reading materials are not especially informative either, and unlike for older children, there is no evidence that such materials have been made easier. Thus, none of the arguments for ramping up text difficulty hold with beginning reading materials (though by a high first-grade or early second-grade reading level, it is reasonable to start shifting to harder texts).

Second, these theories of how text complexity affects learning are somewhat simplistic. They assume that the relative difficulty of the text will determine what the student can learn from the text. But there are other factors that need to be considered. At its extremes, the theory makes great sense—there likely are some texts that are too easy to learn much from and likewise some that are so hard they would not be worth using for instruction. But what of the vast expanse of text levels that lie in between those extremes?

Students probably could learn from any and all of those texts in between, but how much they learn will depend on the amount of scaffolding, support, explanation, rereading, and motivation that the teacher provides. If a text poses few challenges, then students may be able to learn something from it on their own, albeit something relatively small and limited (since the challenges would be small). If the reading material posed greater challenge, then learning would be determined by how much instructional support was there to help guide the students to successfully figure out the text.

Just shifting students into reading complex texts will not be sufficient to advance their learning. Shifting to the use of more complex text is a necessary, but insufficient, step for increasing student reading levels. Teachers will need to become proficient in recognizing text complexity (both by learning to estimate text levels using readability measurement tools, and then by recognizing which features might be contributing to the difficulty of these texts). They also must know the kinds of instructional supports that can be provided if students are to make maximum gains from materials that they struggle with initially.

References

- Betts, E. A. (1946). *Foundations of Reading Instruction*. New York: American Book Company.
- Chall, J. S., Conrad, S., & Harris, S. (1977). *An Analysis of Textbooks in Relation to Declining SAT Scores*. Princeton, NJ: College Entrance Examination Board.
- Hayes, D. P., Wolfer, L. T., & Wolfe, M. F. (1996). Sourcebook Simplification and Its Relation to the Decline of SAT-verbal Scores. *American Educational Research Journal*, 33, 489-508.
- Klare, G. R. (1984). Readability. In P. D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 681-744). New York: Longman.
- Lively, B. A., & Pressey, S. L. (1923). A Method for Measuring the 'Vocabulary Burden' of Textbooks. *Educational Administration and Supervision*, 9, 389-398.
- 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.
- Shanahan, T. (1983). The Informal Reading Inventory and the Instructional Level: The Study That Never Took Place. In L. Gentile, M. L. Kamil, & J. Blanchard (Eds.), *Reading Research Revisited* (pp. 577-580). Columbus OH: Merrill.
- Shanahan; T., Fisher, D., & Frey, N. (2012). The Challenge of Challenging Text. *Educational Leadership*, 69 (6), 58-62.
- Smith, M. (2012). *Not So Common: Comparing Lexile Measures With the Standards' Other Text Complexity Tools*. Raleigh, NC: MetaMetrics.
- Venezky, R. L., & Johnson, D. (1973). Development of Two-Letter Sound Patterns in Grades One Through Three. *Journal of Educational Psychology*, 64 (1), 109-115.