Liben, David. 2020. "Text Complexity." In *The SAT® Suite and Classroom Practice: English Language Arts/Literacy*, edited by Jim Patterson, 9–29. New York: College Board.

CHAPTER 1

Text Complexity

By David Liben

David Liben is a national literacy expert who led schools and taught for over three decades before turning to support publishers and teachers in the equity work that undergirds college and career readiness standards. He helped synthesize the research behind the Common Core State Standards (CCSS) as well as the text complexity measurement work published in appendix A of the CCSS.

Introduction

Every U.S. state and dominion now has college and career readiness standards requiring that students be given access to grade-appropriate complex text, an emphasis that began with the Common Core State Standards in 2010. Despite this requirement, the majority of instruction in our classrooms fails to provide all students with the opportunity, as part of the instructional mix, to work regularly and productively with text at appropriately challenging levels of complexity within successive grade bands (TNTP 2018). The regrettable outcome is that over the thirteen-year span of K–12 schooling, too few students climb the staircase of increasing text complexity that ends with them becoming skilled, independent readers of the kinds of texts typically required in first-year, entry-level college courses and in workforce training programs (Schak et al. 2017).

The primary reason for this failure is the well-intended but mistaken belief that if students work only or primarily with texts they can read with fairly minimal support (that is, within their "instructional" level or "zone of proximal development"), no matter how far below grade level this is, students will still progress toward grade level (Allington 2013). Intuitively, this notion of meeting students where they are and moving them up from there makes some sense, and it certainly feels kind to avoid texts that might lead to (short-term) frustration or discouragement,

yet this approach has no support in the research literature beyond the earliest grades and hasn't been borne out in practice (Shanahan 2011). Compounding the problem of inadequate exposure to complex text is the fact that our K–12 system hasn't provided all students with the opportunity to engage regularly in the volume of reading necessary to grow the vocabulary and knowledge base needed to comprehend text at the college and career readiness level by no later than the end of twelfth grade (TNTP 2018; Landauer and Dumais 1997; Cervetti, Wright, and Hwang 2016). The majority of students who don't reach this level by the end of high school are from low-income families, making both access to complex text and expectations for amount of reading significant equity issues (Schak et al. 2017).

This chapter will explore text complexity in terms of both research and practice. We begin with an overview of the research on text complexity and the related concepts of *standard of coherence* and volume and range of reading. We then turn to how text complexity is measured quantitatively and qualitatively. Finally, we look at how to provide all students with access to both the volume and range of reading they need to engage in and the tools they need to comprehend complex text with steadily increasing facility.

Key Definitions

Before proceeding further, we should examine briefly what we mean by the terms *text complexity*, *complex text*, and *grade-level complex text*. Later sections of this chapter will explore the factors that contribute to complexity (or ease) of text and how complexity is assessed, so for now we'll focus on defining our terminology. We use the term *text complexity* to refer to the inherent difficulty (or ease) of text, distinct from reader and task considerations. Aspects of text contributing to text complexity (or ease) include both word- and sentence-level factors, such as word frequency and sentence length, and broader considerations, such as text structure, density of information and ideas, and demands on the reader's knowledge (i.e., the extent to which the text assumes readers have relevant prior knowledge). Text complexity conceived this way can be measured quantitatively and qualitatively, and, as we'll see, various tools take some or all the above elements (and others) into consideration.

By *complex text*, we mean text that's at a level of complexity typically assigned to students as part of course work in common entry-level, credit-bearing courses in college and workforce training programs. In this sense, *complex text* represents a destination. As defined by college and career readiness standards, students must be able to read complex text proficiently and independently by no later than the end of high school in order to be ready for the kinds of texts they'll encounter in postsecondary education. In other words, by no later than high school graduation,

students must be able to read and comprehend text at the college and career readiness level with little or no scaffolding and support.

Text that's *grade-level complex*, by partial contrast, is material that's appropriately challenging relative to students' level of schooling. *Grade-level complex text* is thus a relative concept: what's grade-level complex for, say, fourth and fifth graders isn't the same as what's grade-level complex for first- and second-year high school students. We can think of grade-level complex text as steps on the journey to the destination of complex text, as we defined the latter above. In this conception, in order for students to reach college and career readiness proficiency in reading—that is, have the ability to read complex text independently—by the end of high school (if not earlier), they need to be asked to read appropriately challenging texts—that is, grade-level complex texts—all along the path of K–12 education. Practically speaking, this means that as students move through the grades, the average level of text complexity should increase, with students in the upper division of high school routinely being expected to read text at the college and career readiness level. (Thus, in the upper grades of high school, *grade-level complex texts* and *complex texts* are the same.) To be sure, students in the various grades should be given a range of texts, some easier and some perhaps more challenging than the "average," and they should receive appropriate scaffolding and support as they tackle the harder ones, but the general movement should be toward mastery of higher levels of complexity as students advance through school.

Text Complexity, Standard of Coherence, and Volume and Range of Reading

Now that we've established something of a framework for discussing text complexity conceptually, let's look briefly at some of the early research in the area. We'll also examine the closely related ideas of *standard of coherence* and volume and range of reading.

Reading between the Lines (ACT 2006) was an early, illuminating study validating the role of complex text in reading and college readiness. Each year, the ACT test is given to roughly two million students. Based on test data, ACT had previously determined a benchmark score for the reading portion of the test that correlated with likelihood of success in college.[1] In the 2006 study, ACT sought to determine whether there were certain kinds of reading comprehension questions that students scoring at or above the benchmark were able to answer that students scoring below the benchmark weren't. From its analysis of student data from seven test forms, ACT found no statistically significant differences

[1] The benchmark score was associated with a 75 percent chance of earning a course grade of C or better and a 50 percent chance of earning a B or better in such credit-bearing college courses as psychology and U.S. history.

between students scoring at or above the benchmark and those scoring below it in terms of performance on literal and inferential questions and on questions testing five separate reading concepts (main ideas/author's approach, supporting details, relationships, meaning of words, and generalizations and conclusions). Nor was there any significant difference in performance with respect to whether passages were literary or informational. The *only* statistically significant difference ACT found was in performance in relation to the complexity of the reading passages themselves. On the most complex passages,[2] students below the benchmark performed at chance level, answering 25 percent of the four-option multiple-choice questions correctly. In other words, the single identifiable predictor of readiness for college-level reading found by ACT was the ability to read complex text. These findings held for male and female students, students from varying income levels, and students from all racial/ethnic groups sampled. *Reading between the Lines* was a major influence on the drive toward including text complexity in the college and career readiness standards that subsequently emerged. Nelson et al. (2012) later established the validity of particular measures of complexity and confirmed the findings of the ACT study.

In another line of research, several studies have investigated whether K–12 text complexity declined over the course of the twentieth century, as prior research (e.g., Adams 2009) has sometimes suggested, and whether, if true, this left a widening gap between K–12 (particularly high school) reading requirements and those for incoming postsecondary students. Work by Gamson, Lu, and Eckert (2013) most directly challenged this premise. While the researchers found that text complexity hadn't declined over a fifty-year window, their study has several important limitations. First, their work addressed only grades three and six. Second, when Gamson, Lu, and Eckert evaluated the central texts of elementary reading textbooks, commonly known as basal readers, they looked only at a single text in each chapter, leaving out the variety of supplementary texts that could conceivably occupy the bulk of student time weekly since basals are designed for daily small-group work. There's no way of knowing how much time teachers spent on the myriad texts in these programs nor how many ignored the central text in favor of working exclusively with texts at students' reading levels. Finally, the researchers did acknowledge that complexity had declined relative to that for texts from a period earlier than the fifty-year window they examined, and they didn't dispute an earlier finding of a four-year gap in complexity between twelfth-grade texts and entry-level college and career training texts (Williamson 2008).

[2] For the study, ACT (2006) defined *complex text* as text that explores subtle, involved, and deeply embedded relationships; has a high degree of richness; has an elaborate, sometimes unconventional structure; has an often intricate style; includes demanding, highly context-dependent vocabulary; and has an implicit, somewhat ambiguous purpose.

Hiebert and Mesmer (2013) also looked at the question of whether K–12 text complexity had declined and what the implications for teaching and learning might be. Observing that it was in middle school and high school where text complexity had dropped over a fifty-year span, they argued that efforts to increase text complexity in school should wait till those grades rather than begin in second grade, as the Common Core State Standards had called for. It should be noted, however, that this approach would greatly shorten the window students have to achieve college and career readiness levels of independent proficiency with complex text and would place a heavy instructional burden on middle school and high school teachers, who generally have less training in and knowledge about reading instruction than do their elementary school counterparts.

Importantly, neither of these reviews seriously challenged the consensus that students' ability to navigate grade-level complex text is a major contributor to success in both the secondary grades and in postsecondary work, whether in college or workforce training. While there's some disagreement about instructional means and whether and to what extent K–12 text complexity has declined over time, there's general agreement that a gap exists between high school and postsecondary reading requirements, and both standards and research have attempted to help close that gap.

A pertinent body of work about how to support students' capacity to navigate complex text successfully concerns a concept known as *standard of coherence* (van den Broek et al. 2011). According to this work, proficient readers have a high standard of coherence in that they enter into the task of reading with an expectation that they'll be able to comprehend all the text has to offer; they're therefore willing to work to achieve understanding when their understanding breaks down. Research indicates that in order to support the development of this proclivity, students need to work regularly, with appropriate scaffolding and support, with texts at levels of complexity above what they can read on their own. (Chapter 2, on close reading, textual evidence, and source analysis, offers some practical suggestions for such instructional support.)

Another aspect of reading instruction that contributes to students' growing capacity to handle the demands of increasingly complex text is sufficient volume and range of reading material. Reading comprehension improves through practice, and the more varied and frequent those practice opportunities are, the greater the increase in capacity. This idea is explored in depth in chapter 3, which focuses on increasing students' troves of knowledge and vocabulary, but it's important to note here the complementary relationship between volume and range of reading and the ability to read complex text. It's from frequent, wide-ranging reading that vocabulary grows and knowledge is gained; these acquisitions, in turn, facilitate the development of reading ability. This

relationship is essential to growing students' capacity to read complex text independently and proficiently (Cervetti, Wright, and Hwang 2016; Cunningham and Stanovich 1998; Guthrie et al. 2009; Landauer and Dumais 1997; Nagy, Anderson, and Herman 1987).

Dimensions of Text Complexity

Earlier we introduced the concept of *text complexity* and offered a rough-and-ready definition of it as the inherent difficulty of a text irrespective of reader and task considerations. In this section, we delve more deeply into what factors contribute to text complexity (or ease) primarily by examining what elements various quantitative (algorithm-based) and qualitative (human judgmental) measures of text complexity attend to.

A text is made more or less complex by the features it contains and the demands it makes on readers. Some of these features and demands are straightforward and fairly intuitive. It makes sense, for example, that word frequency and sentence composition would have important bearing on text complexity. The more uncommon the words in a text and the longer the sentences, the more complex the text becomes. However, many other aspects of text also contribute to how complex (or simple) it is. How densely packed are the information and ideas in the text? How familiar is the structure of the text? How many allusions does the text make? How much knowledge does the author assume the reader already has about the subject? How transparent has that author made the purpose for writing? Is there, in fact, more than one purpose? Answers to these and other questions help determine how complex a given text is.

Below we explore the concept of text complexity in two parts: first, the quantitative dimensions of complexity (those a computer can readily analyze) and second, the qualitative dimensions of complexity (those best or only evaluated by human judgment).

QUANTITATIVE DIMENSIONS OF TEXT COMPLEXITY

Most quantitative measures of text complexity assess only two dimensions of text: word frequency, or how common the words in the text are, and sentence complexity, generally defined as sentence length. Importantly, these features happen to account for most of the variance in text complexity (Nelson et al. 2012). Put more simply, if you have to measure only two factors of text complexity, word frequency and sentence complexity are the ones to attend to, as they have a strong association with the difficulty (or ease) of text.

By far the most common commercial measure of complexity, the Lexile Framework for Reading (**https://lexile.com/**) (Mesmer 2008; White and Clement 2001; Stenner, Sanford-Moore, and Williamson 2012), examines only these two ingredients. The more uncommon the words and the longer the sentences in a text, the higher the Lexile score. Numerous

studies over the years have shown that the higher a text's Lexile score, the lower readers' average comprehension score associated with that text is (Nelson et al. 2012; Mesmer 2008; White and Clement 2001).

Measures that focus only on word frequency and sentence length have been around for decades and are commonly known as *readability measures*. Although Lexiles are used by the majority of publishers today to measure the complexity of their texts, other readability measures are available as well. Microsoft Word includes one of these measures, Flesch-Kincaid, among its spelling and grammar check options. All readability measures produce comparable complexity results, meaning that when these measures are applied to the same set of texts, there's no statistically significant difference in the grade levels they report the texts falling into (Nelson et al. 2012).

Readability measures can, however, mask or at least fail to reveal what actually makes a text more or less complex. Let's first consider informational text. The uncommon words and phrases in a particularly challenging informational text may reflect a generally high level of diction or may instead name specialized concepts, perhaps those associated with a domain of knowledge such as science. In the former case, the resultant complexity is largely a product of vocabulary level; in the latter case, the complexity is more closely associated with the knowledge demands the text places on readers than with vocabulary per se. What's more, while informational texts often have numerous uncommon words (e.g., *respiration*, *caucus*, *guild*, *membrane*), authors of such texts, aware of the complexity that the use of such vocabulary can introduce, take pains to support readers' understanding with glosses, repetition, and the like, thereby mitigating the challenge. Sentence length is also not perfectly correlated with complexity. Many students comprehend no better when a long sentence dense with information is broken up into shorter sentences carrying the same information (Hiebert 2002, 2012). Thus, density of information, rather than sentence length itself, may be the more significant contributor to a text's complexity.

To better understand the effects of informational density and sentence length on readability, consider the following sentence from the preceding paragraph:

> *What's more, while informational texts often have numerous uncommon words (e.g.,* respiration*,* caucus*,* guild*,* membrane*), authors of such texts, aware of the complexity that the use of such vocabulary can introduce, take pains to support readers' understanding with glosses, repetition, and the like, thereby mitigating the challenge.*

The sentence has a Flesch-Kincaid score of 24.5. A rewrite that broke up the content into multiple sentences but otherwise preserved the message might result in something like the following:

> *Informational texts often have numerous uncommon words. Examples of such uncommon words are* respiration*,* caucus*,* guild*, and* membrane*. Authors are aware of the complexity that the use of such vocabulary can introduce. Therefore, they take pains to support readers' understanding with glosses, repetition, and the like. These efforts mitigate the challenge of uncommon words in sentences.*

This cluster of sentences yields a Flesch-Kincaid score of 10.4. In both cases, a reader has to work fairly hard to understand the ideas the author's communicating, yet the multiple-sentence rewrite results in a substantially lower readability score even though the content hasn't appreciably changed.

In addition, the impact on readability of visual elements common to many types of informational texts can't be assessed by electronic means. Illustrations as well as informational graphics (tables, graphs, and the like) can influence complexity, yet their effect on complexity is a matter of qualitative judgment.

The complexity of literary text, too, isn't always adequately captured by some kinds of quantitative analysis. Literary texts often contain a great deal of dialogue, and most dialogue contains a high proportion of common words. Yet dialogue can be difficult for many students to follow. For instance, dialogue can be heavy with idioms (e.g., *fill his shoes*, *flash in the pan*), which, by definition, convey ideas that don't directly correspond to the meaning of the individual words composing them. If unfamiliar to students (including, but not limited to, some English learners), these idioms add to rather than lessen the difficulty of comprehending the text, something a quantitative measure reliant on word- and sentence-level factors would fail to account for. To take another example, authors of certain literary texts (e.g., some political speeches) may use long sentences that contain a great deal of repetition for emphasis or other effect, changing only a word or two between sentences. This repetition makes the text less complex, but the number of long sentences would contribute to a high readability score.

Despite the limits discussed above, Lexiles and other readability measures generally give a good, easily obtainable and interpretable *initial* sense of a text's complexity. However, as we've seen, they don't tell the whole story. Given the considerations discussed above, it might at first appear that text complexity dimensions other than word frequency and sentence complexity—to return to our earlier examples, information density and use of repetition—are beyond the capacity of computer programs to account for and that our only recourse for greater nuance is a human reader. As it turns out, however, some computer applications can attend to features in addition to word frequency and sentence length (though the latter two factors remain important). Information density, for example, can be measured by the frequency with which noun and

verb phrases appear in text, as these two elements are particularly likely to contain informational content. Greater relative density of noun and verb phrases in text is associated with higher complexity regardless of sentence length and word frequency.

Most of the additional factors measurable by computer application fall under the heading of *cohesion features*. Cohesion features are elements that help tie a text together in order to make comprehension easier. The greater the number of cohesion features, the more cohesive and less complex a text will be and vice versa. The best way to get an idea of the various types of cohesion features is to examine text excerpts that contain examples of the most common sorts.

Referential cohesion concerns how an author's word choice informs the connection between and among clauses and sentences to produce greater cohesion and thus lower text complexity. For the most part, referential cohesion is simply the result of the overlap of words or word stems from one clause or sentence to another. The greater the incidence of this overlap, the less complex the text and vice versa. Consider this excerpt (emphasis added) from a 1944 speech by Judge Learned Hand addressing newly naturalized citizens:

> What then is the **spirit of liberty**? I cannot define it; I can only tell you my own faith. The **spirit of liberty** is the **spirit** which is not too sure that it is right; the **spirit of liberty** is the **spirit** which seeks to understand the minds of other men and women; the **spirit of liberty** is the **spirit** which weighs their interest alongside its own without bias . . .

Hand's repetition of "spirit of liberty" and "spirit" throughout the excerpt makes it easier for the listener (or, in our case, reader) to absorb the meaning of each new example of the concept than would've been the case if he'd instead repeated the pronoun "it" throughout, used terms such as "this idea" or "this notion," or a combination of these approaches.

From the same text, consider the line "Some of us have chosen America as the land of our adoption; the rest have come from those who did the same." It takes some processing on the part of the listener (or reader) to understand that "the rest" are the children and grandchildren of those who chose America as the land of their adoption. If Hand had instead written something such as "Some of **us** have **chosen** America as the land of our adoption; the rest of **us** are the children and grandchildren of those who made this **choice**," the repeated use of the word "us" and the overlapping word forms "chosen" and "choice" would've made the content easier for the listener (or reader) to process. The overlap characteristic of referential cohesion can occur within sentences, between adjacent sentences, between sentences farther apart, or all of the above. The closer the overlapping portions are to each other, the greater the enhancement to cohesion.

Global cohesion concerns how an author's word choice informs the connections among the events, ideas, concepts, and information in different parts of a text to produce greater cohesion and thus lower complexity. These connections are established by several different types of words and phrases. Some are relatively straightforward. These include time connectives such as *after*, *earlier*, *before*, *during*, *while*, and *later*; sequential connectives such as *first*, *second*, *next*, and *from here on*; causal connectives such as *because*, *consequently*, and *thus*; and additive connectives such as *additionally*, *furthermore*, *moreover*, *what's more*, and *both*. Adversative connectives, which include *but*, *yet*, *however*, *although*, and *nevertheless*, are a little trickier for student readers. Adversatives connect two notions that on some level conflict with each other—for example, "My favorite sport is baseball; **however**, I watch more football" and "Whales aren't fish, **yet** they spend their lives in the water." All these connectives help to tie the events, ideas, concepts, and information in a text together for the reader. The greater the number of these connectives, the more cohesive and less complex the text. A smaller number of such connectives results in a more complex text, one in which the reader has to infer more of the implicit connections.

Some other aspects of cohesion are *degree of narrativity*, *word concreteness*, and *sentence similarity*. Narrativity is simply how story-like the text is: the more story-like (narrative) it is, the more cohesive and therefore less complex it is. Concrete words (e.g., *mask*, *spoon*, *ammunition*) are easier to integrate into the meaning of a text than are more abstract words (e.g., *democracy*, *appear*, *vary*, *joy*); therefore, greater degrees of concreteness and abstraction are associated with greater ease and difficulty, respectively. The presence of sentences with similar structures, regardless of the length of those sentences, also increases cohesion.

A small number of online tools account for the cohesion features discussed above. The publicly available Coh-Metrix Web Tool (**http://tool.cohmetrix.com/**) yields scores on over one hundred different dimensions, including both readability indices and less traditional factors. The Coh-Metrix Text Easability Assessor (**http://tea.cohmetrix.com/**), also publicly available, channels aspects of the full results into easier-to-interpret percentile scores on a smaller range of dimensions (narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion) and also provides a grade-level readability estimate. TextEvaluator (**https://textevaluator.ets.org/TextEvaluator/**), a tool operated by ETS, provides both free and client-based text complexity evaluation services. TextEvaluator analyzes text on eight dimensions (academic vocabulary, word unfamiliarity, concreteness, syntactic complexity, lexical cohesion, level of argumentation, degree of narrativity, and interactive/conversational style) and also yields an indication of the text's grade level. While these tools' results provide deeper insight into the difficulty

CollegeBoard

SAT Suite Connections

Text complexity is a key consideration on the SAT® Suite Reading and SAT Suite Writing and Language Tests and on the optional SAT Essay.

Students taking the Reading Test and the Writing and Language Test are presented with passages of a consistent range of text complexity, as measured quantitatively and qualitatively. The SAT versions of the tests contain passages in the grades 9–10, grades 11–CCR (college and career readiness), and early postsecondary text complexity bands. (Passages in the last range have a complexity comparable to that of texts frequently assigned in common first-year, credit-bearing postsecondary courses.) The PSAT/NMSQT® and PSAT™ 10 Reading and Writing and Language Tests contain passages in the grades 9–10 and grades 11–CCR bands. The PSAT™ 8/9 tests include passages in the upper end of the grades 6–8 band as well as passages in the grades 9–10 band. The SAT Essay includes passages in the high school (grades 9–12) text complexity range.

Text complexity consistency in each testing program (SAT, PSAT/NMSQT and PSAT 10, and PSAT 8/9) ensures that students are presented with passages of comparable difficulty regardless of when or how they take the test. The increasing average complexity of passages as one moves from PSAT 8/9 to PSAT/NMSQT and PSAT 10 to SAT is consistent with the research outlined in this chapter as well as the general goal of assessing students' readiness or progress toward readiness for success in college and workforce training programs.

(or ease) of text, they do so at the cost of some additional interpretive complexity.

In any event, no quantitative measure alone will give a complete picture of how complex a text is. For that, human judgment is required. Quantitative assessments about text complexity, therefore, should always be used in conjunction with qualitative measures (as well as consideration of the reader and the purpose for reading), even if you're using one of the more sophisticated measures described above.

QUALITATIVE DIMENSIONS OF TEXT COMPLEXITY

A qualitative evaluation of text complexity involves bringing human judgment to bear on the task of ascertaining how easy or challenging a text is to read. The typical process of qualitative evaluation involves a reader using a rubric, ideally accompanied by level-setting exemplar texts, to rate a text's complexity on a number of different dimensions.

A text complexity rubric developed by staff at the College Board is included in this chapter's appendix. A number of other organizations have also developed rubrics as guides to help users determine text complexity by qualitative means. Of these, the most robust tools are available from Achieve the Core (**https://achievethecore.org/page/2725/text-complexity**), Teaching Tolerance (**https://www.tolerance.org/magazine/publications/reading-diversity**), and Achieve (**https://www.achieve.org/files/EQuIP-ELArubric-06-24-13-FINAL.pdf**).

To be most productive, qualitative evaluation should be used in conjunction, whenever possible, with one or more quantitative measures and should focus on those aspects of text not easily or directly assessed by computer. While, for instance, a person could count the number of words in each sentence of a text and determine how common each of the words in it is by using a frequency dictionary, this sort of drudgery should be left to a computer application; the human should instead attend to aspects of text less amenable to machine measurement. Though the specifics of qualitative rubrics vary, in nearly all cases they focus on *knowledge demands*; *language demands* pertaining to vocabulary and syntax; *content* and *theme*; and *structure*.

The *knowledge demands* of a text are the presumptions a text makes about the reader's prior understanding of a topic or situation. A text can make relatively few knowledge demands, supplying significant amounts of background information, or it can make many knowledge demands on the assumption that the reader is already well versed in the subject. Knowledge demands include assumptions about the reader's life experiences, such as those pertaining to family, travel, or work. (The impact of such demands on text complexity can be difficult to evaluate, as what's common knowledge for some readers won't be for others from different backgrounds, yet these sorts of demands appear in all kinds of

texts and shouldn't be ignored.) Knowledge demands can also include assumptions about the reader's prior knowledge in subject domains, such as natural science, social science, the arts, the humanities, and technology. Though it doesn't get the attention it deserves, knowledge of people—of personality types, socioeconomic classes, common motivations, and so on—can also be required for or at least helpful to an understanding of certain texts (e.g., many works of fiction). These categories of knowledge demands can easily overlap. For example, students who travel a good deal are likely to have experience with a wide variety of people. Quantitative measures may pick up on some knowledge demands through a text's use of less common vocabulary, but they can't account for the full extent of these demands.

It's important to understand that knowledge in any of these categories can be gained through wide reading; even life experiences distinct from one's own can be better understood through reading. It's equally important for educators to analyze texts for knowledge demands so that they can anticipate where in the text they may have to stop to check student understanding and perhaps provide some timely information that will clarify what otherwise might block comprehension.

As we've seen, *language demands* can be partially determined by quantitative means, since syntax and vocabulary load are the two variables all the measures we previously reviewed assess, but language demands should also be evaluated using qualitative rubrics and human judgment. The qualitative assessment of language demands addresses some of the weaknesses of readability measurement approaches discussed earlier, such as artificially low readability ratings for some literary narratives (e.g., those making extensive use of common vocabulary but nonetheless containing challenging content) and artificially high readability ratings for some informational texts (e.g., those making extensive use of uncommon vocabulary but providing significant scaffolding for those words and phrases).

Poetry exemplifies the value of qualitative assessment of text complexity over reliance on a machine-produced rating. Consider "The Red Wheelbarrow" (1923), William Carlos Williams's familiar poem:

so much depends
upon

a red wheel
barrow

glazed with rain
water

beside the white
chickens

The words are simple, the structure unusual, the punctuation and print conventions nonexistent. Flesch-Kincaid gives it a 6.8 grade level (when formatted as a conventional sentence), which isn't a fair reflection of its richness and subtlety. Only a human reader can wrestle with the challenge of assigning complexity to a poem.

Determining the challenges presented by *content* and *theme* is another area where qualitative analysis shines. Quantitative methods can't evaluate text for maturity of theme or sophistication of author's purpose, nor can they recognize when themes and purposes are multiple or subtle. This is important because many challenging works by canonical authors commonly read in high school classes (e.g., Morrison, Hemingway, Wright) use simple sentences, common vocabulary, and extensive dialogue. These factors can result in readability scores in the lower elementary range. Yet the themes and content in these works present demands that are perhaps inappropriate for or insurmountable by elementary-age students.

Text structure is another area in which qualitative analysis of complexity provides a critical complement to quantitative approaches. Structure does have a known relationship to complexity: the more the text conforms to the conventions of narrative, the less challenge it presents. This is because students are used to narrative structure from hearing stories and seeing movies before entering school, because narrative elements (e.g., setting, protagonist, problem, outcome/solution) have real-life analogs, and because narratives generally require less prior knowledge (and what knowledge they do call on is generally widely available from life experience) (McNamara, Graesser, and Louwerse 2012). Informational text structures, which include comparison-contrast, problem-solution, goal-action-outcome, chronology, and description, are comparatively more challenging because students don't frequently listen to or read texts with these structures, because these texts don't mimic structures found in life experience, because (unlike the largely unitary narrative structure) there are multiple informational structures to learn, and because informational texts can employ more than one structure. Textbooks essentially use every structure, often within a single chapter—which is likely one of the reasons so many students struggle with them. In most respects, quantitative methods can't account for text structure, though, as noted earlier, both Coh-Metrix and TextEvaluator assess the extent to which a text is story-like (and thus easier to comprehend).

As the above discussion suggests, the most valid analysis of text complexity involves a combination of quantitative and qualitative approaches. Quantitative measures can more efficiently account for most word- and sentence-level factors than can humans, and the more sophisticated tools can go well beyond that to factor in elements such as cohesion devices and some aspects of text structure. There remains,

> **"The most valid analysis of text complexity involves a combination of quantitative and qualitative approaches. There remains a vital place for the skilled, knowledgeable human evaluator in assessments of text complexity."**

however, a vital place for the skilled, knowledgeable human evaluator in assessments of text complexity. No machine has yet to replace the insight that a thoughtful human reader can bring to bear on many of the more substantive aspects contributing to the complexity (or ease) of text.

Besides the fact that there are specific dimensions of text complexity unknowable to even the most advanced algorithm, the synergistic nature of text argues strongly for human involvement. The ease or challenge of any one aspect of text is influenced by that of others. Complex syntax with simple vocabulary is less difficult to process than the same syntax with commensurately challenging vocabulary. High information density presents fewer demands if readers have extensive background knowledge on the topic. Richard Dawkins (1996) once compared the genome to a game of cat's cradle: tugging on any one string affects every other. Text complexity, with all its ingredients and features, is just as tightly interwoven.

Still, too much can be made of the abstruseness of text complexity. The large number of features that one could consider in judging a text's difficulty can be misleading in one crucial sense: the features aren't equal in their contribution to complexity. As previously noted, word- and sentence-level factors account for much of what makes a text easier or harder to read. Moreover, a great deal of work over nearly a century has established the key role of knowledge in determining complexity. (Vocabulary and knowledge are the particular subject of chapter 3.)

Implementation Advice and Sample Activities

Before teachers can help their students understand what makes text complex, they must understand text complexity themselves. The best way to do that is to practice analyzing text qualitatively for complexity. At the time of this writing, the most comprehensive set of tools and training materials can be found at Achieve the Core (**https://achievethecore.org/page/2725/text-complexity**). After getting a feel for using such tools, teachers then can work with their students to analyze sections of texts for complexity by breaking down the text's features and come to shared conclusions about the complexity of parts of the text in relation to the whole. Teachers and students can even make decisions together about where to concentrate classroom time and attention in close reading. For example, characters in novels sometimes digress philosophically; these stretches are generally more complex than surrounding text. Similarly, the opening chapters of informational texts often introduce the topic or purpose as well as provide a picture of what's to come. The information density, knowledge demands, and vocabulary load of such sections can present greater challenge than that of the work as a whole and often merit disproportionate amounts of class time.

In addition to giving students a "meta-awareness" of features contributing to complexity (or ease), this purposeful exploration of text helps students see the whole as greater than the sum of its parts, a reflection of the previously discussed synergistic nature of texts. For example, uncommon vocabulary becomes less challenging if syntax is simple, text structure is narrative, and knowledge demands are minimal. In other words, as noted, every feature of a text tugs on every other feature; an understanding of the resulting whole likely gives students a better sense of how to recognize complexity in texts and the features that contribute to difficulty and ease. Such recognition will help comprehension. Students will be aware of these elements as always present in some form in any text they'll encounter, and as they internalize that reality they'll start to evaluate these features for themselves. This will cultivate much more sophisticated reading and self-monitoring of comprehension and heighten their expectation that text can be made to make sense (i.e., raise their standard of coherence).

Examination of text complexity also plays out at the sentence level, with equally valuable results. Analyzing challenging syntax supports students' ability to process the varied and sophisticated sentence structures they'll encounter in complex text. Students can undertake a range of activities that will prepare them to work with complex syntax encountered in sophisticated texts. These activities include parsing longer sentences (i.e., breaking them down into their simplest elements in order to lay bare the underlying structure and logic) and combining short sentences into longer ones while preserving the content of each original sentence.[3] (For a fuller treatment of syntax, see chapter 4, on the conventions of standardized English.)

Analyzing texts to determine complexity, working with challenging sentences, and reading challenging text regularly all contribute to students' ability to comprehend grade-level material proficiently. These activities have to be accompanied by students engaging in lots of diverse reading on their own in order for them to experience a volume and range of reading sufficient to promote the development of a robust vocabulary and to gain access to the knowledge that will allow them to connect with complex information and ideas. This work should culminate, by no later than high school, in students' attainment of the ability to read independently and proficiently at the college and career readiness level. Many students, however, are below or far below grade level and will need support to accomplish this goal. An optimal mode of

[3] Two good resources for sentence combining are Richard Nordquist, "An Introduction to Sentence Combining," ThoughtCo., updated October 22, 2018, **https://www.thoughtco.com/an-introduction-to-sentence-combining-1692421**, and Johnson County (KS) Community College, sentence combining handout, accessed February 8, 2019, **http://www.jccc.edu/student-resources/resource-centers-tutoring/writing-center/files/sentence-combining.pdf**. Quill.org (**https://www.quill.org/tools/connect**) also offers a number of sentence-combining activities accompanied by electronic feedback.

support is in the form of text sets, groups of readings organized around a topic appropriate to a subject of study. Such cohesive sets have the advantage of helping students efficiently develop vocabulary and build knowledge on a given subject. Chapter 2 discusses the ways that close reading, with support, facilitates the reading of grade-level complex text by all students. Chapter 3, on vocabulary and knowledge development, provides additional information on supporting students' learning through text sets and adequate volume of reading.

Conclusion

This chapter has established the importance of students being able to comprehend complex text to be ready to succeed in college and workforce training. We know that far too many students need to improve their reading capacity before they're able to take common entry-level, credit-bearing postsecondary courses with a high chance of success (Cromley and Azevedo 2007; Oakes and Guiton 1995; Slavin 1990; Stanovich 1986). We know, too, the features that determine text complexity and techniques that can support all students in working with the complex text typically encountered in the last years of high school and in common first-year, entry-level postsecondary courses. And we know that reading widely and voluminously leads to higher levels of comprehension. Why, then, aren't more students sufficiently skilled in reading to be prepared for the challenges they'll face in K–12 and after high school?

Students who fall behind in the early grades stay behind or even fall farther behind as they move through the grades (Stanovich 1986). As these students proceed from one grade to another, they're unable to read text at grade-level complexity or, in many cases, even close to grade-level complexity. Confronted with this situation, we have too often restricted these students to reading texts at their "level" despite the lack of research attesting to the efficacy of this approach beyond the very earliest grades (Shanahan 2011). When these students reach high school, they're invariably put in lower-track classes where, once again, they read texts of lower complexity levels and read less than their higher-tracked peers. National Assessment of Educational Progress data (Schak et al. 2017) consistently show that the majority of these students are from lower-income families, are children of color, are English learners, or are some combination of the three. Access to complex text and the advantages that access bestows are therefore equity issues as well as academic ones. Understanding text complexity, the importance of facility with complex text to future success, and how to support all students attaining such facility are essential to changing this trajectory for students we need to help the most.

Appendix

Dimension	Complexity Band				
	Grades 4–5	**Grades 6–8 (PSAT 8/9)**	**Grades 9–10 (SAT, PSAT/ NMSQT, PSAT 10, PSAT 8/9)**	**Grades 11–CCR (college and career readiness) (SAT, PSAT/ NMSQT, PSAT 10)**	**Grades 13–14 (early postsecondary) (SAT)**
	Basic	**Somewhat challenging**	**Moderately challenging**	**Complex**	**Highly complex**
Purpose ***Chiefly informational texts***	Single Clear and direct	Single Generally clear and direct	Single Relatively straightforward	Single or multiple Relatively subtle or complex Possibly hidden or intentionally obscured	Multiple Subtle or complex Possibly hidden or intentionally obscured
Level(s) of Meaning ***Chiefly literary texts***	One or multiple; if multiple, text can be understood/ enjoyed on a literal level	One or multiple; if multiple, text can be understood/ enjoyed on a literal level	One or multiple; if multiple, useful to a full understanding of the text	Multiple and important to a full understanding of the text	Multiple and necessary to a full understanding of the text
Central Idea(s) and Theme(s)	Explicit Straightforward	Explicit or implicit; if implicit, easy to infer Relatively straightforward	Explicit or implicit; if implicit, relatively easy to infer Relatively subtle	Explicit or implicit; if implicit, relatively challenging to infer Relatively subtle or complex	Explicit or implicit; if implicit, challenging to infer Subtle or complex
Information, Ideas, and Relationships	Straightforward; connections are explicit and clear	Somewhat challenging; connections may be implicit but easy to infer	Moderately challenging; connections may be implicit but relatively easy to infer	Challenging; connections are often implicit and relatively challenging to infer	Highly challenging; connections are frequently implicit and challenging to infer
Accessibility of Experiences and Ideas	Common or easily relatable	Sometimes unfamiliar	Sometimes unfamiliar	Often unfamiliar	Frequently unfamiliar
Abstraction	Concrete	Generally concrete	Sometimes abstract or theoretical	Often abstract or theoretical	Frequently abstract or theoretical
Density and Pace	Low to moderately low Slow to fairly slow	Moderately low Fairly slow	Moderate Fairly rapid	Moderately high to high Fairly rapid to rapid	High to very high Rapid to very rapid
Text Structure	Basic; easy to predict	Straightforward; generally easy to predict	Relatively straightforward	Relatively intricate or complex	Intricate or complex
Syntax	Basic Mostly simple sentences	Somewhat challenging Mostly simple and compound sentences	Moderately challenging Simple, compound, and complex sentences	Challenging Many complex sentences	Highly challenging Mostly complex sentences
Diction	Similar to everyday language; may be conversational in style and tone	Generally similar to everyday language; may be conversational in style and tone	Somewhat elevated and somewhat distinct from everyday language	Elevated and distinct from everyday language; ironic, ambiguous, or intentionally misleading language possible	Elevated and sharply distinct from everyday language; ironic, ambiguous, or intentionally misleading language possible

Appendix (*continued*)

Dimension	Complexity Band				
	Grades 4–5	**Grades 6–8 (PSAT 8/9)**	**Grades 9–10 (SAT, PSAT/ NMSQT, PSAT 10, PSAT 8/9)**	**Grades 11–CCR (college and career readiness) (SAT, PSAT/ NMSQT, PSAT 10)**	**Grades 13–14 (early postsecondary) (SAT)**
	Basic	**Somewhat challenging**	**Moderately challenging**	**Complex**	**Highly complex**
Vocabulary	Familiar Moderately low tier 2 and/or tier 3 demands; tier 3 words/phrases routinely glossed and foregrounded Archaic, foreign, and otherwise uncommon words/ phrases generally absent	Generally familiar Moderate tier 2 and/or tier 3 demands; tier 3 words/phrases routinely glossed and foregrounded Archaic, foreign, and otherwise uncommon words/phrases generally absent, discernible from context, or not central to meaning	Moderate High tier 2 and/or tier 3 demands; tier 3 words/phrases less explicitly foregrounded Archaic, foreign, and otherwise uncommon words/phrases more likely to appear and to be relevant to a full understanding of the text	Moderately high Very high tier 2 and/ or tier 3 demands; relatively little scaffolding for tier 3 words/phrases Archaic, foreign, and otherwise uncommon words/phrases much more likely to appear and to be important to a full understanding of the text	High Major tier 2 and/or tier 3 demands; little scaffolding for tier 3 words/phrases Archaic, foreign, and otherwise uncommon words/phrases highly likely to appear and to be central to a full understanding of the text
Knowledge Demands (World/Cultural, Subject Matter)	Moderately low	Moderate	Moderate to moderately high	Moderately high to high	High to very high
Intertextuality	Absent, low, or incidental to full understanding of the text	Absent, low, or incidental to full understanding of the text	Absent, low, or incidental to full understanding of the text	Low to moderate; may be important to full understanding	Moderate to high; may be central to full understanding of the text
Subject Matter Sensitivity	Little emotional or intellectual maturity specifically expected	Certain degree of emotional and intellectual maturity sometimes expected; recognition that one's viewpoint may differ from that in the text is required	Some degree of emotional and intellectual maturity and ability to distance oneself from text expected	Emotional and intellectual maturity and ability to distance oneself from text often expected	Emotional and intellectual maturity and ability to distance oneself from text routinely expected

References

ACT. 2006. *Reading between the Lines: What the ACT Reveals about College Readiness in Reading.* Iowa City, IA: ACT. **https://files.eric.ed.gov/fulltext/ED490828.pdf**.

Adams, Marilyn Jager. 2009. "The Challenge of Advanced Texts: The Interdependence of Reading and Learning." In *Reading More, Reading Better: Are American Students Reading Enough of the Right Stuff?*, edited by Elfrieda H. Hiebert, 163–89. New York: Guilford.

Allington, Richard L. 2013. "What Really Matters When Working with Struggling Readers." *Reading Teacher* 66, no. 7 (April): 520–30.

Cervetti, Gina N., Tanya S. Wright, and HyeJin Hwang. 2016. "Conceptual Coherence, Comprehension, and Vocabulary Acquisition: A Knowledge Effect?" *Reading and Writing: An Interdisciplinary Journal* 29, no. 4 (April): 761–79.

Cromley, Jennifer G., and Roger Azevedo. 2007. "Testing and Refining the Direct and Inferential Mediation Model of Reading Comprehension." *Journal of Educational Psychology* 99, no. 2 (May), 311–25.

Cunningham, Anne E., and Keith E. Stanovich. 1998. "What Reading Does for the Mind." *American Educator* 22, no. 1–2 (Spring–Summer): 8–15.

Dawkins, Richard. 1996. *The Blind Watchmaker: Why the Evidence of Evolution Reveals a Universe without Design.* New York: W. W. Norton.

Gamson, David A., Xiaofei Lu, and Sarah Anne Eckert. 2013. "Challenging the Research Base of the Common Core State Standards: A Historical Reanalysis of Text Complexity." *Educational Researcher* 42, no. 7 (October): 381–91.

Guthrie, John T., Angela McRae, Cassandra S. Coddington, Susan Lutz Klauda, Allan Wigfield, and Pedro Barbosa. 2009. "Impacts of Comprehensive Reading Instruction on Diverse Outcomes of Low- and High-Achieving Readers." *Journal of Learning Disabilities* 42, no. 3 (May–June): 195–214.

Hand, Learned. 1944. "The Spirit of Liberty" (speech). Digital History. University of Houston. **http://www.digitalhistory.uh.edu/disp_textbook.cfm?smtID=3&psid=1199**.

Hiebert, Elfrieda H. 2002. "Standards, Assessment, and Text Difficulty." In *What Research Has to Say about Reading Instruction*, 3rd ed., edited by Alan E. Farstrup and S. Jay Samuels, 337–69. Newark, DE: International Reading Association.

Hiebert, Elfrieda H. 2012. "The Common Core State Standards and Text Complexity: What Librarians Need to Know . . . and Do." *Teacher Librarian* 39, no. 5: 13–19.

Hiebert, Elfrieda H., and Heidi Anne E. Mesmer. 2013. "Upping the Ante of Text Complexity in the Common Core State Standards: Examining Its Potential Impact on Young Readers." *Educational Researcher* 42, no. 1 (January): 44–51. https://doi.org/10.3102/0013189X12459802.

Landauer, Thomas K., and Susan T. Dumais. 1997. "A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge." *Psychological Review* 104, no. 2 (April): 211–40.

McNamara, Danielle S., Art Graesser, and Max Louwerse. 2012. "Sources of Text Difficulty: Across Genres and Grades." In *Measuring Up: Advances in How to Assess Reading Ability*, edited by John P. Sabatini, Elizabeth R. Albro, and Tenaha O'Reilly, 89–116. Lanham, MD: Rowman and Littlefield Education.

Mesmer, Heidi Anne E. 2008. *Tools for Matching Readers to Texts: Research-Based Practices*. New York: Guilford.

Nagy, William E., Richard C. Anderson, and Patricia A. Herman. 1987. "Learning Word Meanings from Context during Normal Reading." *American Educational Research Journal* 24, no. 2 (Summer): 237–70.

Nelson, Jessica, Charles Perfetti, David Liben, and Meredith Liben. 2012. *Measures of Text Difficulty: Testing Their Predictive Value for Grade Levels and Student Performance.* Report to the Gates Foundation. Washington, DC: Council of Chief State School Officers. https://achievethecore.org/content/upload/nelson_perfetti_liben_measures_of_text_difficulty_research_ela.pdf.

Oakes, Jeannie, and Gretchen Guiton. 1995. "Matchmaking: The Dynamics of High School Tracking Decisions." *American Educational Research Journal* 32, no. 1 (Spring): 3–33.

Schak, Oliver, Ivan Metzger, Jared Bass, Clare McCann, and John English. 2017. *Developmental Education: Challenges and Strategies for Reform.* Washington, DC: U.S. Department of Education. https://www2.ed.gov/about/offices/list/opepd/education-strategies.pdf.

Shanahan, Timothy. 2011. "Rejecting Instructional Level Theory." *Shanahan on Literacy* (blog). August 21, 2011. http://shanahanonliteracy.com/blog/rejecting-instructional-level-theory#sthash.HEjW2UUM.8d0TpVhf.dpbs.

Slavin, Robert E. 1990. "Achievement Effects of Ability Grouping in Secondary Schools: A Best-Evidence Synthesis." *Review of Educational Research* 60, no. 3 (Autumn): 471–99.

Stanovich, Keith E. 1986. "Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy." *Reading Research Quarterly* 21, no. 4 (Fall): 360–407.

Stenner, A. Jackson, Eleanor Sanford-Moore, and Gary L. Williamson. 2012. *The Lexile Framework for Reading Quantifies the Reading Ability Needed for "College and Career Readiness."* Durham, NC: MetaMetrics. **https://metametricsinc.com/research-publications/lexile-framework-reading-quantifies-reading-ability-needed-college-career-readiness/**.

TNTP. 2018. *The Opportunity Myth: What Students Can Show Us about How School Is Letting Them Down—And How to Fix It.* New York: TNTP. **https://tntp.org/publications/view/student-experiences/the-opportunity-myth**.

van den Broek, Paul, Catherine M. Bohn-Gettler, Panayiota Kendeou, Sarah Carlson, and Mary Jane White. 2011. "When a Reader Meets a Text: The Role of Standards of Coherence in Reading Comprehension." In *Text Relevance and Learning from Text*, edited by Matthew T. McCrudden, Joseph P. Magliano, and Gregory Schraw, 123–39. Charlotte, NC: IAP Information Age Publishing.

White, Sheida, and John Clement. 2001. *Assessing the Lexile Framework: Results of a Panel Meeting.* NCES 2001-08. Washington, DC: U.S. Department of Education, National Center for Education Statistics. **https://nces.ed.gov/pubs2001/200108.pdf**.

Williams, William Carlos. 1923. "The Red Wheelbarrow." In *The Collected Poems of William Carlos Williams*, vol. 1 (1909–1939), edited by A. Walton Litz and Christopher MacGowan, 224–25. New York: New Directions, 1991.

Williamson, Gary L. 2008. "A Text Readability Continuum for Postsecondary Readiness." *Journal of Advanced Academics* 19, no. 4 (August): 602–32. **https://doi.org/10.4219/jaa-2008-832**.