Advancing Our Students' Language and Literacy
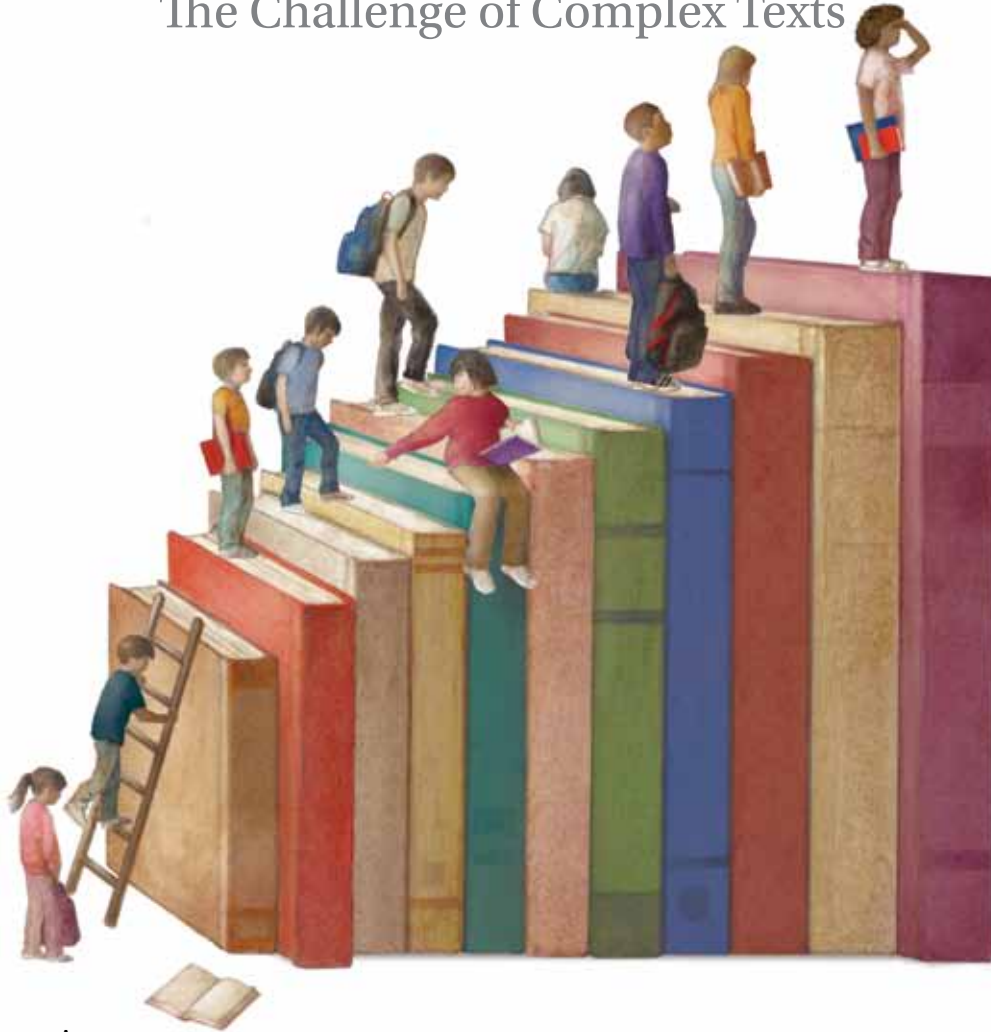
The Challenge of Complex Texts

By Marilyn Jager Adams

"Few Changes on SAT Posted by Class of 2010."[1] "Scores on SAT College Entrance Test Hold Steady."[2] "Class of 2008 Matches '07 on the SAT."[3] Year by year, point by point, it is hard to see the real news in these headlines. The real news is not that the SAT scores have held steady. The real news is that the SAT scores haven't increased. The SAT scores of our college-bound students

---

Marilyn Jager Adams is a research professor in the Cognitive, Linguistic, and Psychological Sciences Department of Brown University and former chief scientist for Soliloquy Learning Inc. She is the author of numerous scholarly papers and several books, including two landmark works: Beginning to Read: Thinking and Learning about Print *and* Phonemic Awareness in Young Children. *This article is adapted with permission of the Guilford Press from "The Challenge of Advanced Texts: The Interdependence of Reading and Learning," which Adams wrote for* Reading More, Reading Better, *edited by Elfrieda H. Hiebert, copyright 2009 by Guilford Press.*

have been languishing not for one or two years, but for a long time. Several decades ago, scores were much higher.

The SAT score decline began in 1962, nearly 50 years ago. From 1962 to 1980, math scores fell 36 points to 492 while verbal scores fell 54 points to 502. Since 1980, the math scores have been gradually climbing back and are now at 516. Fluctuations aside, the verbal scores remain unchanged, even today stuck at 502.

If I were writing the headline for the next newspaper story on the SATs, here's what you'd see: "Seniors and Their SAT Scores Sabotaged by Low-Level Textbooks." And if the copyeditor would let me, I'd add an exclamation point! The literacy level of our secondary students is languishing because the kids are not reading what they need to be reading. This is a strong claim. Let me lay out the evidence and argument so you can judge for yourself.

Not Just the SAT Scores

To be sure, whether scores on the SAT exams truly reflect relevant or important intellectual or academic proficiencies remains a topic of discussion.[4] Yet, the SATs are not the only indication that

the literacy growth of our secondary students has fallen behind.

Between 1994 and 1998, the United States joined 19 other developed countries in an international evaluation of adult literacy levels.[5] As compared with their peers in the other countries, the literacy scores of older U.S. adults (36 years old and up) were quite high, ranking in the top five. In contrast, the scores for younger U.S. adults (35 years old or less) ranked in the bottom half of the distribution by every measure. Among young adults with a high school diploma or less, those from the United States fell at the bottom of the pile, ranking 19th out of 20. Even among participants who had completed four or more years of postsecondary education, the scores of our young adults were below the average for same-aged and like-educated peers in the other countries. The young adults in this study would have graduated from high school

> An analysis of 800 schoolbooks published between 1919 and 1991 found that the difficulty of the text had been significantly reduced.

between 1974 and 1998, during the period when the verbal SAT scores were bottoming out.

In international assessments of schoolchildren, the performance of our fourth-graders is above average. However, the performance of our high school students is average, at best.[6] The results of our own National Assessment of Educational Progress (NAEP) show a similar contrast: while the reading of younger students has been improving over time, that of older students has not. NAEP's analysis of changes in reading performance between 1971 and 2008 shows that average scores of 9-year-olds increased by 12 points. Those of 13-year-olds increased by 4 points. But the average scores of 17-year-olds have not changed.[7] The lack of progress among 17-year-olds is especially jarring when factoring in our dropout problem. Roughly 25 percent of eighth-graders nationwide drop out of school before completing high school;[8] presumably, those who stay in school, and therefore participate in NAEP as 17-year-olds, disproportionately include the more successful and motivated students. One can't help but wonder whether they were trying hard when they took the tests, since there is no personal consequence for doing well or poorly on the international trials or on NAEP.

On the other hand, college entrance examinations are voluntary, and performing well on them is the very point of taking them. ACT (known until 1996 as the American College Testing Program) tracked the literacy scores of eighth-, tenth-, and twelfth-graders on ACT college readiness and entrance exams.[9] For each of the cohorts examined (and regardless of gender, race/ethnicity, or household income), the students were collectively on track in the eighth and tenth grades for better scores than they ultimately obtained in the twelfth grade. ACT's report concludes that there

is a specific problem at the secondary school level.*

Taking a closer look at the poor performance of students on its college entrance exam, ACT determined that the major stumbling block for students is complex texts. The maximum score on the reading component of the ACT college entrance exam is 36; scores of less than 21 predict reading difficulties in college coursework and also in the workplace. Among students who took the ACT exam in 2005, the scores of 51 percent—more than half—fell below 21. And among that 51 percent, average performance on the complex texts was at chance levels (i.e., random guessing would produce the same scores).

SAT Decline Prompts Investigation

Back in 1977, having watched SAT scores fall for 15 years, the College Board, which developed and administers the SAT, engaged a panel to try to identify the underlying causes of the decline.[11] A first hypothesis to be checked was whether the test had somehow become more demanding. But, no, to the contrary, indications were that scoring had become more *lenient*.[12] A second prominent hypothesis was that the decline was due to changes in the demographics of the test takers. Analyses showed this hypothesis to be largely correct, but only for a brief while. Over the early 1960s, changes in the composition of the tested population accounted for as much as three-quarters of the test score decline—and, no wonder, for during this period the number of students taking the SAT tripled. Over the 1970s, however, though the test-taking population stabilized, the scores did not. Instead, the decline continued, even steeper than before, while the extent to which it could be ascribed to demographic shifts shrank to 30 percent at most.[13] Furthermore, *the scores that dropped most were those of the strongest students, the students in the top 10 percent of their class*; the scores of students toward the bottom of the distribution held steady or even increased.[14]

Another hypothesis examined by the College Board's panel was that the reading selections on the tests had somehow become too hard for the students. Reading researcher Jeanne Chall and her colleagues tested this hypothesis by sampling passages from SAT tests administered between 1947 and 1975, and using readability analyses to compare their difficulty.[15] The data indicated that the SAT passages had actually become *easier* over this period—so scores should have been going up. Further, between 1963 and 1975, during the years of the score decline, the average difficulty of the test passages lay at the eleventh-grade level, which should have been solidly in range for twelfth-grade college-bound students. Yet scores were going down.

Chall thought there had to be some reason why the twelfth-graders were not able to read eleventh-grade texts. With this in mind, she and her colleagues evaluated popular eleventh-grade textbooks in history, literature, grammar, and composition. *The average difficulty of the textbooks lay between the ninth- and tenth-grade levels*.

Could this discrepancy between the reading level of the SAT and that of the textbooks explain the score decline? If students had neither practiced nor been instructed with reading materials as hard as the SAT passages, then one could hardly expect them

---

*The same conclusion was drawn by the College Entrance Examination Board in the mid-1970s and again in the mid-1980s.[10]

to read the latter with competence and confidence.

By the early 1990s, SAT scores appeared to have plateaued. The College Board decided to "recenter" the scale by adding about 80 points to the verbal scores (and about 25 points to the math scores) so as to return the mean of each test to a value close to 500 points.[†] Beleaguered, the College Board also changed the name of the test from the Scholastic Aptitude Test to simply the SAT, with the letters standing for nothing.

A Closer Look at Textbooks

In the 1980s and 1990s, another team of researchers, led by Donald P. Hayes, returned to Chall's hypothesis, extending her work with a revealing series of studies. In one of the most extensive, they analyzed the difficulty of 800 elementary, middle, and high school books published between 1919 and 1991.[16] Their results indicated that the difficulty of the text in these books had been significantly reduced and, further, that the years over which this reduction occurred were temporally aligned with the SAT score decline.

As one indication of this trend, the average length of the sentences in books published between 1963 and 1991 was shorter than that of books published between 1946 and 1962. In the seventh- and eighth-grade textbooks, for example, the mean length of sentences decreased from 20 words to 14 words—"the equivalent of dropping one or two clauses from every sentence."[17] Meanwhile, the sophistication of the books' wording also declined. The wording of schoolbooks published for eighth-graders from 1963 forward was as simple as that in books used by fifth-graders before 1963. Worse, among literature texts required in English classes, the wording of twelfth-grade texts published after 1963 was simpler than the wording of seventh-grade texts published prior to 1963.

Continuing their investigation, the researchers found that it was especially schoolbooks for students in grades 4 and up that were simplified in the years after 1962. Moreover, although the wording of schoolbooks for children generally increased across grades 1 through 8, the same was not true of high school books. *Across grades 9 through 12 (including texts for Advanced Placement courses), the difficulty levels of the literature books were shown to differ little from one another or from the grade 7 and grade 8 offerings*. One bright spot was high school students' science texts, which were significantly more difficult than their

English books. However, even among science texts, only those designated for Advanced Placement coursework evidenced difficulty levels comparable to that of the average daily newspaper for adults.

Such a disparity between the students' schoolbooks and the passages on the SAT might well explain the decline in SAT scores. More significantly, failing to provide instruction or experience with "grown-up" text levels seems a risky course toward preparing students for the reading demands of college and life.

To wit, while the analyses of Hayes and his colleagues showed that textbooks had become progressively easier over the century, they also indicated that the difficulty of English language newspapers had remained nearly constant.[18] Could this disparity be a factor in the declining circulation of newspapers? Similarly, they found the level of the wording of scientific magazines, whether aimed at professionals or laypersons, had increased dramatically from 1930 to 1990.[19] If it is a national goal to inspire more students to become engineers and scientists, then shouldn't the difficulty of our schoolbooks have increased alongside? If a goal is to ensure that our students will be able to stay sufficiently informed about scientific progress to conduct business, reflect on policy, and manage their family's health and education, then at a minimum, shouldn't the difficulty of our schoolbooks keep pace with the difficulty of scientific publications aimed at the general public?

The Vocabulary of Written Language

Reading educators have long appreciated that there is a very strong relationship between vocabulary and reading comprehension. But what exactly is it about the wording of texts that underlies this relation? Part of the answer is that written texts draw upon many more words than normally arise in oral language situations.[20]

To gain insight into this phenomenon, Hayes and colleagues compared spoken language with texts.[21] For this study, they focused on trade publications rather than school materials, and the texts they used included preschool books, children's books, comic books, adult books, magazines, newspapers, and abstracts from scientific magazines. For comparison, they compiled and analyzed a variety of oral language samples, including language from prime-time adult television shows, children's television shows, mothers' speech to children ranging in age from infancy to adolescence, conversations among college-educated adults (including from the Oval Office), and adults providing expert witness testimony for legal cases. Regardless of the source or situation and without exception, the richness and complexity of the words used in the oral language samples paled in comparison with the written texts. Indeed, of all the oral language samples evaluated, the only one that exceeded even preschool books in lexical range was expert witness testimony.

This difference between the wording of oral and written language must lie at the crux of the advanced literacy challenge, as it points to a profound dilemma. On the one hand, the extent of this disparity implies that the great majority of words needed for understanding written language is likely to only be encountered—and thus can only be learned—through experience with written text. On the other hand, research has taught us that written text is

accessible—and thus permits learning—only if the reader or listener already knows the vast majority of words from which it is constructed. Indeed, research indicates that reading with comprehension depends on understanding at least 95 percent of the words of a text.[22]

How Many New Words Do Readers Need to Learn?

So roughly how many words do kids need to learn in order to be proficient readers? This question raises the second key part of the vocabulary problem.

Suppose you counted the number of times each different word in this article occurred. What you would find is that there are a few words that I have used quite a number of times, and many, many others that I used only once or twice. This distribution of word counts or frequencies is an example of what is known as Zipf's law.[23]

According to Zipf's law, every natural language sample is made up of relatively few words that recur over and over again, and many, many words that arise very infrequently. The type of natural language sample does not matter and, provided that it is not too short, neither does its size. That is, whether you counted all the words in a casual conversation, a lecture, a newspaper article, a whole book, or even a whole library's worth of books, you would find the same thing: of all the different words in your sample, a small number would occur over and over again, while many, many others would occur only once.

Zipf's law may feel intuitively obvious. Less obvious, however, are its implications with respect to the vocabulary challenge.

An example may vivify the issue. Counting words that appear in relevant text is a common approach to making dictionaries. For example, if you wanted to make a dictionary for geologists, you might begin by gathering a sample of the kind of articles about geology that you think your customers would like to read and then counting the number of occurrences of all the different words within them. The goal is to make sure your dictionary contains all the words that your customers will want to look up most.

Similarly, as part of creating *The American Heritage School Dictionary*,[24] John Carroll and his colleagues were asked to figure out which words should be included by examining children's reading materials. To do this, the team gathered texts that had been written especially for children in grades 3 through 8, taking care that the collection as a whole captured the range of different kinds of text and topics that the children might read in amounts that were proportionate to how often they could be expected to read them. From across these materials, the team then extracted 10,000 excerpts, totaling 5 million words of text in all, which, after sorting, turned out to include 86,741 different words. Their job was then to figure out which of these 86,741 words arose sufficiently often to warrant inclusion in the dictionary.[25]

Enter Zipf's law. Just 109 very frequent words accounted for fully half of the vast sample of children's reading material that Carroll and colleagues had put together. Indeed, 90 percent of the sample was accounted for by just 5,000 relatively common words. At the other extreme, more than half of the words appeared only once. Still worse: the team estimated that the actual number of different words in the children's reading materials—that is, the number of different words that would have turned up if they had counted such texts exhaustively rather than just working with excerpts—would have totaled 609,606. Due to Zipf's law, a sample of 5 million words was just plain too small even to identify—much less to judge the relative frequency of—the vast majority of words that might well have belonged in the dictionary.

But hold it. We are talking about materials that are specifically written for and meant to be understood by schoolchildren in grades 3 through 8. How can they possibly be expected to know more than 600,000 different words?

In fact, many of these words are cousins of each other. For example, if a child knows the word *shoe*, then she or he is unlikely to experience difficulty with *shoes*. Similarly, a child probably won't have trouble with word families like *walk*, *walked*, and *walk-*

> Making textbooks easier ultimately denies students the very language, information, and modes of thought they need most to move up and on.

ing. Pushing this reasoning further, vocabulary researchers Bill Nagy and Richard Anderson[26] have argued that students shouldn't have problems with any sort of prefixing, suffixing, or compounding of a word, provided that the meaning of the word's base is preserved. As examples, they suggested that if children know the word *elf*, they should have little problem with *elfin* or with pairs such as *cow/cowhand*, *know/knowledge*, *therapy/therapeutic*, and *represent/misrepresent*. Eliminating all such "closely related" words from the word count that Carroll and colleagues had done for the dictionary, and keeping only base words plus affixed or compound words whose meanings are harder to figure out from their base words (such as *vice/vicious*, *well/farewell*, *shift/shiftless*, *fix/prefix*), Nagy and Anderson estimated that the actual number of separate words that children need be taught is closer to 100,000. If Nagy and Anderson's elimination rules were too aggressive given children's word sense, then the actual number might be double or triple their estimate. And, of course, if we extend concern from grade-school materials to advanced texts, the actual number must be larger still.

Developing Students' Vocabulary: Examining the Options

So, what is the best way to help students master the many, many words they must know to understand advanced texts? In broad terms, there appear to be only two options: (1) to endeavor to teach students the words they will need to know, and (2) to expect students to learn new words through reading.

Is direct vocabulary instruction worthwhile? Based on a highly regarded meta-analysis, the answer seems to be a resounding "yes."[27] Across studies involving a variety of students, instructional specifics, and outcome measures, the meta-analysis showed that

direct vocabulary instruction significantly increases knowledge of words that are taught. Just as importantly, students who received vocabulary instruction were found to perform significantly better on global nonspecific vocabulary measures such as standardized tests, indicating that such instruction promotes learning of words beyond those that have been explicitly taught (e.g., being taught a word like *aquarium* helps with indirectly learning words like *aquatic*, *aqueduct*, and *aqueous*).

However, we must bear in mind that, by its very nature, direct vocabulary instruction admits coverage of precious few words relative to the magnitude of the challenge. Even if, beginning in grade 1 and continuing through grade 12, teachers consistently taught—and students perfectly retained—20 vocabulary words each and every week, the gain in vocabulary would total only 8,640 words in all (20 words × 36 weeks of school × 12 years), many times fewer than what is required.

Such considerations have led some scholars to argue that the only feasible means by which students might acquire an adequate reading vocabulary is through the process of inferring the meaning of each new word from its context in the course of reading.[28] Indeed, research shows that the probability that students understand and retain any given new word that they encounter in print is 0.05.[29]

So how far will this get them? Researchers have (generously) estimated that median, middle-class, fifth-grade students read close to 1,000,000 words of text per year, in school and out.[30] Based on Carroll and colleagues' research, we can expect a million words of reading to include roughly 17,200 different words. If we suppose that the students already know one-quarter of the words in their texts, then the number of new words they should encounter through this reading would equal 12,900 per year. Yet, if the likelihood that the students will understand and retain each of these words is only 0.05, then their vocabulary can only be expected to grow by 645 per year, giving them but 5,160 new words by the time they graduate from high school.

Recalling that even texts that are for students in grades 1 through 8 presume knowledge of at least 100,000 different words, it is clear that both estimates for learning vocabulary fall way short of the need. At the same time, however, both estimates also seem at odds with the intuitive sense that a high school student need be neither a genius nor a tireless scholar to read and understand most materials written for grade-school children.

Insights from a Computer Model of Vocabulary Acquisition

For another way to think about vocabulary acquisition, let's consider an intriguing computer model called Latent Semantic Analysis (LSA) that was developed by Tom Landauer and his colleagues.[31] The core mechanism underlying the LSA model is "associative learning." When a text is input into the LSA model, the computer builds an association between each individual word of the text and the total set of words—that is, the context—in which the word has appeared. Where a word shows up in multiple contexts, the strength of the association between the word and each of the separate contexts is weakened through competition. Where a word arises repeatedly in one particular context, the association between the two is strengthened.

Importantly, the associations between words and contexts in the LSA model are bidirectional. That is, there are links from each word to each of its contexts and also from each context to all of its words. As a result, the full complex of knowledge that is called forth as each word is "read" extends through its contexts to other words, and through those words to other contexts and words. Thus, as the model "reads" the next word of the text and the next and the next, activation spreads to other, related complexes of knowledge, which may well include clusters that have never before been directly represented by any combination of words and contexts the model has ever "read" before.

Moreover, because the model's knowledge is represented relationally, the addition or modification of any one connection impacts many others, pulling some closer together, pushing some further apart, and otherwise altering the strengths and patterns of connections among words and contexts. Through this dynamic, reading causes the connections that collectively capture LSA's knowledge of words to grow, shrink, and shift continuously, continually, and always in relation to one another.

In short, the model's response to any text it "reads" extends well beyond what is denoted by the specific words of the text. Further, the richness of the model's representation of any text that it "reads" depends on how much it already knows. Just as with people,[32] the larger its starting vocabulary and the more it has read before, the more it will learn and understand from the next text.

In comparing LSA's word-learning to that of schoolchildren, the researchers began by "training" it with a set of texts judged comparable to the lifelong learning of a typical seventh-grader. The researchers then gave the model new texts to "read" and measured its vocabulary growth. The results showed that the likelihood that the computer gained adequate understanding of new words it encountered in these new texts was 0.05—just exactly the same as researchers have found for schoolchildren.[33]

But the results showed something else, too. It turned out that, with each new reading, the model effectively increased its understanding not just of words that were in the text but also of words

that were *not* in the text. Indeed, measured in terms of total vocabulary gain, the amount the model learned about words that did *not* appear in a given reading was three times as much as what it learned about words that *were* in the reading.

"What?" we cry, "How can that be? How can reading a text produce increases in knowledge of words that it does not even contain? That is not credible! It makes no sense!" But wait. If we were talking about knowledge rather than words, then it would make lots of sense. Every concept—simple or complex, concrete or abstract—is learned in terms of its similarities, differences, and relationships with other concepts with which we are familiar. As a simplistic example, when we read about tigers, then, by dint of both similarities and contrasts, we learn more about all sorts of cats and, further, about every subtopic mentioned along the way. The more deeply we read about tigers, the more nuanced and complex these concepts and their interrelations become.

As explained earlier, it was to be expected that LSA's full response to any new text would spread beyond the content of the text itself. The unexpected discovery was that this dynamic would impact the model's understanding of individual words. Given that words are really nothing more than labels for interrelated bundles of knowledge, perhaps this should not have been surprising.

In the study that modeled a seventh-grader, the researchers were able to gauge LSA's overall vocabulary growth by computationally examining changes in the representation of every word to which it had ever been exposed. Yet here is a mull-worthy correlate: unavoidably, the bundles of concepts and relations that emerged or were strengthened through LSA's reading experience included many that pertained to words that the model had never seen before. An analogous effect might explain why researchers have found time and again that the strength of students' vocabulary predicts the likelihood that they will learn new words from context,[34] the probability that they will correctly infer a new word's meaning from context,[35] and both the amount and nature of their reasoning when they are asked to explain how they do so.[36] Even when students are *told* the meaning of a new word, their prior vocabulary strength predicts the likelihood that they will retain it.[37] (These are known as "Matthew effects," referring to the notion that the rich get richer and the poor get poorer.) As the reader's linguistic and conceptual knowledge grows in richness and complexity, it will increasingly support the meanings of many new words and the representation of many new spheres of knowledge.

Cognitive psychologists broadly agree that the meaning of any word consists of bundles of features and associations that are the cumulative product of the reader's experience with both the word in context and the concepts to which it refers. What is unique about the LSA model is its demonstration that this structure and dynamic can so richly and powerfully evolve through accrued experience with the various contexts in which words do and do not occur—that is, sheerly through reading.

Another way to state the larger point here is that words are not just words. They are the nexus—the interface—between communication and thought. When we read, it is through words that we build, refine, and modify our knowledge. What makes vocabulary valuable and important is not the words themselves so much as the understandings they afford. The reason we need to know the meanings of words is that they point to the knowledge from which

we are to construct, interpret, and reflect on the meaning of the text. A core implication of the LSA model is that students' knowledge of words grows less through any process of inferring their meanings, one by one, based on the sentences in which they arise, than as a product of learning more generally about the contexts in which they arise and of understanding the concepts and relationships to which they refer.

Knowledge, Cognitive Strategies, and Inferences

If reading results in so rich a network of knowledge through nothing more than overlaps and contrasts in associations, then shouldn't students learn far more efficiently, given active, incisive inference and comprehension strategies? Research indicates that such strategies can be taught and suggests that doing so may improve comprehension.[38] However, inference and comprehension strategies seem to do little to compensate for weak domain knowledge.[39] Instead, research repeatedly shows prior domain knowledge to be a far stronger predictor of students' ability to comprehend or to learn from advanced texts.[40] Of course, students' comprehension and learning is also influenced by their reading skills (such as decoding and fluency). But even the advantage of strong reading skills turns out to be greatest for students with strong domain knowledge.[41]

Again, such findings should not be surprising. Cognitive research affirms that there are two modes of reasoning.[42] The first, most common mode is knowledge-based. This sort of reasoning is rapid, extensive, and automatic. This is the sort of reasoning that ensures, for example, that we properly understand the meaning of *fan* depending on whether the text is about a *soccer fan*, a *ceiling fan*, or a *peacock's fan*. This is the sort of reasoning that computer models such as LSA statistically emulate.

The second mode of reasoning is conscious and rule-based. Such logical, analytic thought also warrants instructional attention in our schools, as it is our means of deliberately evaluating

and vetting our thoughts for bias, happenstance, and inconsistencies. However, no reasoning strategy, however well-structured, can rival the speed, power, or clarity of knowledge-driven understanding;[43] nor can it compensate for an absence of sufficient information.

There may one day be modes and methods of information delivery that are as efficient and powerful as text, but for now there is no contest. To grow, our students must read lots. More specifically, they must read lots of "complex" texts—texts that offer them new language, new knowledge, and new modes of thought. Beyond the basics, as E. D. Hirsch, Jr., the founder of Core Knowledge, has so forcefully argued, the reading deficit is integrally tied to a knowledge deficit.[44]

Back to the Classroom: A Strategy for Developing Advanced Reading

The capacity to understand and learn from any text depends on approaching it with the language, knowledge, and modes of thought, as well as the reading skill, that it presumes. It would seem, then, that when assigning materials from which students are to learn, there are basically but two choices. Either the materials must be sufficiently accessible in language and concept for the students to read and understand on their own, or the students must be given help as they read. Some students receive such help in their homes, but many do not and, as I have argued elsewhere, this is likely the major factor underlying the achievement gap.[45] In any case, because opportunities for one-on-one reading assistance are limited in the typical school setting, educators often feel that their only alternative is to restrict assignments to materials that are within their students' independent reach. There follows the popularity of so-called high-low texts, intended to offer high interest or information alongside low demands on vocabulary and reading skill.

It was in this spirit, through earnest efforts to ensure full curricular access to all, that the complexity of schoolbooks came to be relaxed. Sadly, as this strategy pulled vortically upon itself, it did not solve the access problem but, instead, made it worse. In terms of literacy growth, making the textbooks easier is an approach that ultimately denies students the very language, information, and modes of thought they need most in order to move up and on. Is there any escape from this dilemma?

The answer is yes, there is, and it follows directly from Zipf's law. Again, according to Zipf's law, every coherent text is made up of a few words that recur again and again, and many words that occur just once or only a few times. And, again, Zipf's law is shown to hold for virtually every natural language domain, regardless of its size, topic, modality, or sophistication.

Let us first consider the implications of Zipf's law with respect to word-frequency counts such as the one undertaken for *The American Heritage School Dictionary*.[46] Recall that the goal of such large frequency counts is to capture as broad and representative a picture of the language as possible. For this reason, the collective texts from which they are constructed are chosen to represent as broad and representative a range of topics and genres as possible while avoiding repetition of any particular topic or text. A consequence of this text-sampling strategy is that the low-frequency words within these word counts fall into two different categories. In the first category are words that are rare because they are com-

plex, technical, obsolete, or esoteric (e.g., *caprifoliaceous*, *omphaloskepsis*, and *mumpsimus*). In the second category, however, are words that are rare because their meanings are relatively specific and are often tied to specific contexts, topics, and genres.[47] For example, a high-frequency word such as *home* may be expected in texts of many different types and topics of which only a small subset would accept such low-frequency synonyms as *condominium*, *wigwam*, *hospice*, *habitat*, *birthplace*, *burrow*, or *warren*. The same holds for the high-frequency word *strong* versus the more specific alternatives *valid*, *virile*, *tensile*, *pungent*, *dominant*, *vibrant*, *durable*, *lethal*, *tyrannical*, and *undiluted*. More generally, the greater the information that a word carries, the fewer the topics and contexts in which it will arise.

Because words in both of these two categories are low frequency, both tend to be excluded by readability formulas that are based on large word-frequency counts. Yet, the "information" in

> We must organize our readings in every subject so each text bootstraps the language and knowledge needed for the next. Gradually, students will be ready for texts of greater complexity.

a text is shown to depend disproportionately on words in this second category.[48] Because of this, when words in this second category are removed or substituted so as to "simplify" the text, much of the information in the text is removed along with them.

A more specific statement of Zipf's law is this: which words appear frequently and infrequently in any given text depends on what the text is about. So, in a text about cooking, the word *habitat* would be infrequent, but in a text about ecology, it would not. The problem with large word-frequency counts—and, by extension, with the readability formulas that are based on them—is that, by design, the texts from which they are generated are collectively topic-neutral. Similarly, if your students were to read a little of this and a little of that, without rereading anything or dwelling on any topic, then the likelihood of their encountering any given information-bearing word would be quite small.

In contrast, if your students read several texts on a single topic, they would encounter a number of domain-specific, information-bearing words. In such texts, the words that rise to the top are those most useful for describing the concepts and relationships that are central to that topic. For example, a quick sampling of informational texts about Mars that I picked off the Internet affirms that, without exception, and whether the intended audience was young children or scientists, the nouns *Mars* and *planet* are among the five most frequent in each. The balance of the dominant nouns in each text depends on the subtopic in focus—variously, its moons, its geography, our efforts at its exploration, etc.

With this in mind, and combined with what else we know

about literacy growth, Zipf's law prescribes a self-supporting strategy for developing the sorts of knowledge structures that complex texts require. That is, we know that even for young[49] and delayed[50] readers, any new word encountered (and identified correctly) in print becomes a sight word with little more than a single encounter, provided its meaning is known. We know that the more that students already know about the topic of a text, the greater their understanding and learning will be as they read.[51] We know that vocabulary strength predicts the speed and security with which students learn the meanings of unfamiliar words, whether through reading[52] or direct instruction.[53]

The challenge, then, lies in organizing our reading regimens in every subject and every class such that each text bootstraps the language and knowledge that will be needed for the next. Zipf's law tells us that this can be done by carefully sequencing
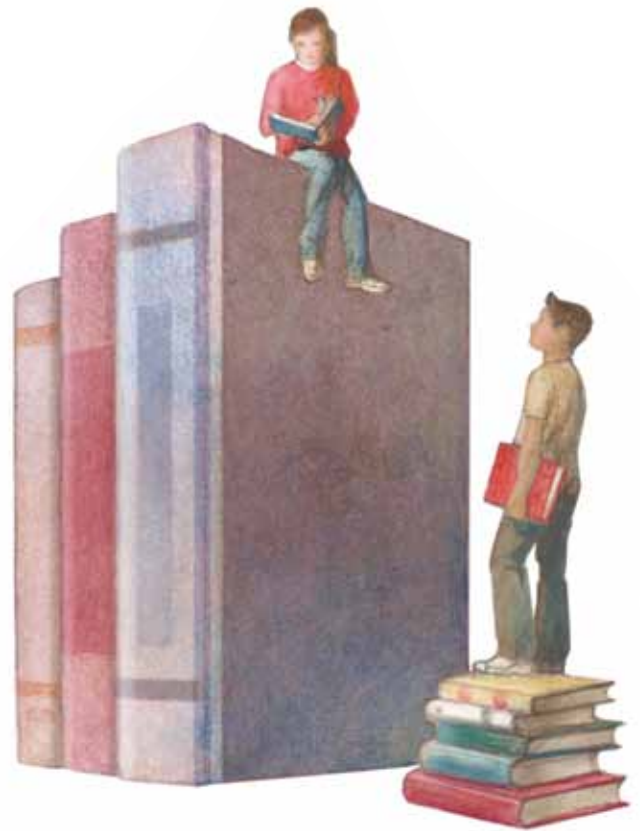
A great benefit of a common core curriculum is that it would drive an overhaul of the texts we give students to read, and the kinds of learning and thought we expect their reading to support.

and scaffolding students' reading materials within any given topic. *Ideally, such scaffolding would begin on the very first day of school, with prekindergarten and kindergarten teachers reading aloud stories and nonfiction texts that build on each others' vocabulary and ideas.*

Teachers in any grade (and parents) would do well to follow this relatively straightforward strategy:

1. Select a topic about which your students need to learn. (There will be plenty of time for other topics once you've started this process.) If the students are below grade level, begin with shorter, simpler texts.
2. Teach the key words and concepts directly, engaging students in using and discussing them to be sure they are well anchored.
3. As the students learn the core vocabulary, basic concepts, and overarching schemata of the domain, they will become ready to explore its subtopics, reading (or having read aloud to them) as many texts as needed or appropriate on each subtopic in turn.

Gradually and seamlessly, students will find themselves ready for texts of increasingly greater depth and complexity. Better yet, as their expertise on, say, Mars, expands, they will find themselves in a far better position to read about Venus, Jupiter, earth sciences, space exploration, and on and on.

Can advanced texts really be made accessible to less proficient readers in this way? Yes. As a concrete example, no text on dinosaurs would get through a readability formula for second-graders.

However, having built up their vocabulary and domain knowledge, many second-graders are able to read and understand remarkably sophisticated texts about dinosaurs with great satisfaction. Similarly, I have rarely met a Boston cabby—no matter how much he decried reading—who wasn't quick to pick up and read a news article about the Red Sox. *Knowledge truly is the most powerful determinant of reading comprehension.* The greatest benefits of literacy grow through reading deeply in multiple domains and about multiple topics. We can and must do a better job of leading—and enabling—our students to do so. If education is the key to opportunity, then their options, in school and beyond, depend on it.

The Role of a Common Core Curriculum

There are some who object reflexively to the notion of a common core curriculum. Yet, if you think about it, the very concept of publicly supported schooling is predicated on the belief that there is a certain body of knowledge and abilities that is needed by every citizen for a safe, responsible, and productive life.

Under the Massachusetts School Law of 1642, every town was made responsible for teaching every child "to read perfectly the English tongue," and to understand the capital laws of the commonwealth and the principles of religion, as well as for ensuring every child was provided an apprenticeship in "some lawful calling, labour, or employment." In effect, these requirements constituted the colony's *common core curriculum.*

In the centuries since then, responsibility for our children's religious education has been reassigned from the school to families and churches. However, the educational and literacy levels required by the other dimensions of life, liberty, and the pursuit of happiness have exploded. In our times, written language has

become the major medium not just for education but for information in every aspect of life. Further, as priest, professor, and historian Walter Ong has pointed out, the ubiquity of audio support hardly matters: written language is the underlying medium for educated communication regardless of modality.[54]

The arguments for a *common* core curriculum are partly that it would be readily accessible to every teacher and school, partly that it would provide continuity and coherence for the millions of students who frequently change schools (an issue E. D. Hirsch, Jr., explores beginning on page 30), and partly that a vocabulary-building curriculum is too big and too hard a job for any teacher or school to put together alone. Creating each unit, for each grade K–12, will depend on judicious selection not just of topics but of the reading materials comprising each unit. From the billions of pages of print that are available, finding those that are both well written and appropriate will take work. The task of building a good core curriculum will require intense effort by teams of educators and scholars, including the best minds and sensibilities available.

In creating a common core curriculum, the goal is neither to dictate nor to limit what all students should be able to know and do. As detailed within this issue of *American Educator*, the core curriculum might fill only two-thirds of students' instructional time. Perhaps, too, the units would be populated with alternate sets of readings. After all, as reviewed in this article, the greatest benefit of a well-structured program of reading and learning is that it prepares the student to read other materials with competence and thoughtful comprehension. If education is to nurture interest and support relevance, it must also leave room for some choice. The purpose of a core curriculum is to build the foundations that will put students in good stead to choose and pursue what they wish to learn and do—which, of course, depends integrally on their being able to learn and do it.

From my perspective, a great benefit of a common core curriculum is that it would drive a thorough overhaul of the texts we give students to read, and the kinds of learning and thought we expect their reading to support. Amid the relatively few SAT headlines this fall, the one written by the College Board, which administers the SAT, stood out: "2010 College-Bound Seniors Results Underscore Importance of Academic Rigor."[55] As the College Board went on to explain, "students in the class of 2010 who reported completing a core curriculum—defined as four or more years of English, three or more years of mathematics, three or more years of natural science, and three or more years of social science and history—scored, on average, 151 points higher on the SAT than those who did not complete a core curriculum." We've known at least since Socrates that challenging, well-sequenced coursework leads to more learning. It is time for us, as a nation, to act on that knowledge and give all students the common core curriculum they need to be prepared for advanced reading and learning. □

Endnotes

1. Catherine Gewertz, "Few Changes on SAT Posted by Class of 2010," *Education Week*, September 22, 2010.

2. Eric Gorski, "Scores on SAT College Entrance Test Hold Steady," Associated Press, September 13, 2010.

3. Sara Rimer, "Class of 2008 Matches '07 on the SAT," *New York Times*, August 26, 2008.

4. See, for example, Willard Wirtz et al., *On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline* (New York: College Entrance Examination Board, 1977); Tamar Lewin, "Students' Paths to Small Colleges Can Bypass SAT," *New York Times*, August 31, 2006; Jesse M. Rothstein, "College Performance Predictions and the SAT," *Journal of Econometrics* 121, no. 1–2 (2004): 297–317; and Lawrence C. Stedman, "Respecting the Evidence: The Achievement Crisis Remains Real," *Education Policy Analysis Archives* 4, no. 7 (1996): 1–31.

5. Andrew Sum, Irwin Kirsch, and Robert Taggart, *The Twin Challenges of Mediocrity and Inequality: Literacy in the U.S. from an International Perspective* (Princeton, NJ: Educational Testing Service, 2002).

6. Stephen Provasnik, Patrick Gonzales, and David Miller, *U.S. Performance across International Assessments of Student Achievement: Special Supplement to the Condition of Education 2009* (Washington, DC: National Center for Education Statistics, 2009).

7. Bobby D. Rampey, Gloria S. Dion, and Patricia L. Donahue, *NAEP 2008 Trends in Academic Progress* (Washington, DC: National Center for Education Statistics, 2009).

8. Marilyn Seastrom, Lee Hoffman, Chris Chapman, and Robert Stillwell, *The Averaged Freshman Graduation Rate for Public High Schools from the Common Core of Data: School Years 2001–02 and 2002–03* (Washington, DC: National Center for Education Statistics, 2005).

9. ACT, *Reading between the Lines: What the ACT Reveals about College Readiness in Reading* (Iowa City, IA: ACT, 2006).

10. Wirtz et al., *On Further Examination*; and William W. Turnbull, *Student Change, Program Change: Why the SAT Scores Kept Falling* (New York: College Entrance Examination Board, 1985).

11. Wirtz et al., *On Further Examination*.

12. Albert E. Beaton, Thomas L. Hilton, and William B. Schrader, *Changes in the Verbal Abilities of High School Seniors, College Entrants, and SAT Candidates between 1960 and 1972* (New York: College Entrance Examination Board, 1977).

13. Lawrence C. Stedman, "The Condition of Education: Why School Reformers Are on the Right Track," *Phi Delta Kappan* 75, no. 3 (1993): 215–225.

14. Turnbull, *Student Change, Program Change*.

15. Jeanne S. Chall, Sue S. Conard, and Susan H. Harris, *An Analysis of Textbooks in Relation to Declining SAT Scores* (New York: College Entrance Examination Board, 1977).

16. Donald P. Hayes, Loreen T. Wolfer, and Michael F. Wolfe, "Schoolbook Simplification and Its Relation to the Decline in SAT-Verbal Scores," *American Educational Research Journal* 33, no. 2 (1996): 489–508.

17. Hayes, Wolfer, and Wolfe, "Schoolbook Simplification," 497.

18. Hayes, Wolfer, and Wolfe, "Schoolbook Simplification," 489–508.

19. Donald P. Hayes, "The Growing Inaccessibility of Science," *Nature* 356, no. 6372 (1992): 739–740.

20. Douglas Biber, "Are There Linguistic Consequences of Literacy? Comparing the Potentials of Language Use in Speech and Writing," in *The Cambridge Handbook of Literacy*, ed. David R. Olson and Nancy Torrance (Cambridge: Cambridge University Press, 2009), 75–91.

21. Donald P. Hayes and Margaret G. Ahrens, "Vocabulary Simplification for Children: A Special Case of 'Motherese'?" *Journal of Child Language* 15, no. 2 (1988): 395–410.

22. Emmett A. Betts, *Foundations of Reading Instruction* (New York: American Book Company, 1946); Ronald P. Carver, "Percentage of Unknown Vocabulary Words in Text as a Function of the Relative Difficulty of the Text: Implications for Instruction," *Journal of Reading Behavior* 26, no. 4 (1994): 413–437; Marcella Hu Hsueh-chao and Paul Nation, "Unknown Vocabulary Density and Reading Comprehension," *Reading in a Foreign Language* 13, no. 1 (2000): 403–430; and Batia Laufer, "What Percentage of Text-Lexis Is Essential for Comprehension?" in *Special Language: From Humans Thinking to Thinking Machines*, ed. Christer Lauren and Marianne Nordman (Clevedon, UK: Multilingual Matters, 1988), 316–323.

23. George K. Zipf, *The Psycho-Biology of Language* (Boston: Houghton Mifflin, 1935).

24. *The American Heritage School Dictionary* (Boston: Houghton Mifflin, 1977).

25. John B. Carroll, Peter Davies, and Barry Richman, *The American Heritage Word Frequency Book* (Boston: Houghton Mifflin, 1971).

26. William E. Nagy and Richard C. Anderson, "How Many Words Are There in Printed School English?" *Reading Research Quarterly* 19, no. 3 (1984): 304–330.

27. Steven A. Stahl and Marilyn M. Fairbanks, "The Effects of Vocabulary Instruction: A Model-Based Meta-Analysis," *Review of Educational Research* 56, no. 1 (1986): 72–110.

28. William E. Nagy, Patricia A. Herman, and Richard C. Anderson, "Learning Words from Context," *Reading Research Quarterly* 20, no. 2 (1985): 233–253.

29. William E. Nagy, Richard C. Anderson, and Patricia A. Herman, "Learning Word Meanings from Context during Normal Reading," *American Educational Research Journal* 24, no. 2 (1987): 237–270.

30. Richard C. Anderson, Paul T. Wilson, and Linda G. Fielding, "Growth in Reading and How Children Spend Their Time Outside of School," *Reading Research Quarterly* 23, no. 3 (1988): 285–303; and Nagy, Anderson, and Herman, "Learning Word Meanings."

31. Thomas K. Landauer, "Learning and Representing Verbal Meaning: The Latent Semantic Analysis Theory," *Current Directions in Psychological Science* 7, no. 5 (1998): 161–164; and Thomas K. Landauer and Susan T. Dumais, "A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge," *Psychological Review* 104, no. 2 (1997): 211–240.

32. Anne E. Cunningham and Keith E. Stanovich, "Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later," *Developmental Psychology* 33, no. 6 (1997): 934–945; and Harry L. Gradman and Edith Hanania, "Language Learning Background Factors and ESL Proficiency," *Modern Language Journal* 75, no. 1 (1991): 39–51.

Advanced Texts

(Continued from page 11)

33. Nagy, Anderson, and Herman, "Learning Word Meanings."

34. Meredyth Daneman and Ian Green, "Individual Differences in Comprehending and Producing Words in Context," *Journal of Memory and Language* 25, no. 1 (1986): 1–18; Patricia A. Herman, Richard C. Anderson, P. David Pearson, and William E. Nagy, "Incidental Acquisition of Word Meaning from Expositions with Varied Text Features," *Reading Research Quarterly* 22, no. 3 (1987): 263–284; John L. Shefelbine, "Student Factors Related to Variability in Learning Word Meanings from Context," *Journal of Reading Behavior* 22, no. 1 (1990): 71–97; and Robert J. Sternberg and Janet S. Powell, "Comprehending Verbal Comprehension," *American Psychologist* 38, no. 8 (1983): 878–893.

35. Louise Morrison, "Talking about Words: A Study of French as a Second Language Learners' Lexical Inferencing Procedures," *Canadian Modern Language Review* 53, no. 1 (1996): 41–75; and Hossein Nassaji, "The Relationship between Depth of Vocabulary Knowledge and L2 Learners' Lexical Inferencing Strategy Use and Success," *Canadian Modern Language Review* 61, no. 1 (2004): 107–134.

36. Manuel G. Calvo, Adelina Estevez, and Margaret G. Dowens, "Time Course of Elaborative Inferences in Reading as a Function of Prior Vocabulary Knowledge," *Learning and Instruction* 13, no. 6 (2003): 611–631; and Nassaji, "Relationship."

37. Andrew Biemiller and Catherine Boote, "An Effective Method for Building Meaning Vocabulary in Primary Grades," *Journal of Educational Psychology* 98, no. 1 (2006): 44–62; Joseph R. Jenkins, Marcy L. Stein, and Katherine Wysocki, "Learning Vocabulary through Reading," *American Educational Research Journal* 21, no. 4 (1984): 767–787; and Claudia Robbins and Linnea C. Ehri, "Reading Storybooks to Kindergartners Helps Them Learn New Vocabulary Words," *Journal of Educational Psychology* 86, no. 1 (1994): 54–64.

38. National Reading Panel, *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction; Reports of the Subgroups* (Washington, DC: National Institute of Child Health and Human Development, 2000).

39. Tenaha O'Reilly and Danielle S. McNamara, "The Impact of Science Knowledge, Reading Skill, and Reading Strategy Knowledge on More Traditional 'High-Stakes' Measures of High School Students' Science Achievement," *American Educational Research Journal* 44, no. 1 (2007): 161–196.

40. Filip Dochy, Mien Segers, and Michelle M. Buehl, "The Relation between Assessment Practices and Outcomes of Studies: The Case of Research on Prior Knowledge," *Review of Educational Research* 69, no. 2 (1999): 145–186; and Amy M. Shapiro, "How Including Prior Knowledge as a Subject Variable May Change Outcomes of Learning Research," *American Educational Research Journal* 41, no. 1 (2004): 159–189.

41. O'Reilly and McNamara, "The Impact of Science Knowledge."

42. See Steven A. Sloman, "The Empirical Case for Two Systems of Reasoning," *Psychological Bulletin* 119, no. 1 (1996): 3–22; and Steven A. Sloman, *Causal Models: How People Think about the World and Its Alternatives* (Oxford: Oxford University Press, 2005).

43. K. Anders Ericsson, Neil Charness, Robert R. Hoffman, and Paul J. Feltovich, eds., *The Cambridge Handbook of Expertise and Expert Performance* (Cambridge: Cambridge University Press, 2006).

44. E. D. Hirsch, Jr., *The Knowledge Deficit: Closing the Shocking Education Gap for American Children* (Boston: Houghton Mifflin, 2006).

45. Marilyn Jager Adams, "The Promise of Automatic Speech Recognition for Fostering Literacy Growth in Children and Adults," in *International Handbook of Literacy and Technology*, vol. 2, ed. Michael C. McKenna, Linda D. Labbo, Ronald D. Kieffer, and David Reinking (Hillsdale, NJ: Lawrence Erlbaum Associates, 2006), 109–128; and Marilyn Jager Adams, "The Limits of the Self-Teaching Hypothesis," in *Educating the Other America: Top Experts Tackle Poverty, Literacy, and Achievement in Our Schools*, ed. Susan B. Neuman (Baltimore: Paul H. Brookes, 2008), 277–300.

46. Carroll, Davies, and Richman, *The American Heritage Word Frequency Book*.

47. Michael Stubbs, "Language Development, Lexical Competence and Nuclear Vocabu-lary," in *Educational Linguistics*, ed. Michael Stubbs (Oxford: Blackwell, 1986), 98–115.

48. Patrick J. Finn, "Word Frequency, Information Theory, and Cloze Performance: A Transfer Feature Theory of Processing in Reading," *Reading Research Quarterly* 13, no. 4 (1977-1978): 508–537.

49. Anne E. Cunningham, "Accounting for Children's Orthographic Learning while Reading Text: Do Children Self-Teach?" *Journal of Experimental Child Psychology* 95, no. 1 (2006): 56–77.

50. David L. Share and Carmit Shalev, "Self-Teaching in Normal and Disabled Readers," *Reading and Writing: An Interdisciplinary Journal* 17 (2004): 769–800.

51. O'Reilly and McNamara, "The Impact of Science Knowledge"; and Shapiro, "Including Prior Knowledge."

52. Daneman and Green, "Individual Differences"; Herman et al., "Incidental Acquisition of Word Meaning"; Shefelbine, "Student Factors Related to Variability"; and Sternberg and Powell, "Comprehending Verbal Comprehension."

53. Biemiller and Boote, "An Effective Method for Building"; Jenkins, Stein, and Wysocki, "Learning Vocabulary through Reading"; Charles A. Perfetti, Edward W. Wlotko, and Lesley A. Hart, "Word Learning and Individual Differences in Word Learning Reflected in Event-Related Potentials," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, no. 6 (2005): 1281–1292; and Robbins and Ehri, "Reading Storybooks to Kindergartners."

54. Walter J. Ong, *Orality and Literacy: The Technologizing of the Word*, 2nd ed. (New York: Routledge, 2002).

55. College Board, "2010 College-Bound Seniors Results Underscore Importance of Academic Rigor," press release, September 13, 2010, www.collegeboard.com/press/releases/213182.html.