

Examining Three Assumptions About Text Complexity: Standard 10 of the Common Core State Standards

Elfrieda H. Hiebert

TextProject &
University of California, Santa Cruz

Katie Van Sluys

DePaul University

TextProject Article Series

March 2014

TextProject, Inc.

SANTA CRUZ, CALIFORNIA

Hiebert, E.H., & Van Sluys, K. (2014). Examining three assumptions about text complexity: Standard 10 of the Common Core State Standards. In K.S. Goodman, R.C. Calfee, & Y.M. Goodman (Eds.), *Whose knowledge counts in government literacy policies? Why expertise matters* (pp. 144-160). New York, NY: Routledge

Running Head: THREE ASSUMPTIONS ABOUT TEXT COMPLEXITY

Examining Three Assumptions About Text Complexity:

Standard 10 of the Common Core State Standards

Elfrieda H. Hiebert, TextProject & University of California, Santa Cruz

Katie Van Sluys, DePaul University

To appear in K. Goodman, R.C. Calfee, and Y. Goodman (Ed.), *Using Knowledge from the Past to Create the Future: Perspectives from the Reading Hall of Fame*.

Please excuse any grammatical errors. This article is the original unedited version the authors submitted to the publisher.

Examining Three Assumptions About Text Complexity:

Standard 10 of the Common Core State Standards

At its core, reading involves a text and texts vary greatly in complexity—their structures, vocabularies, styles, and topics. Standard 10 of the Common Core State Standards (CCSS; CCSS Initiative, 2010) calls for students to grow capacity in reading texts of ever increasing complexity over the school years, culminating in high school graduates' ability to read the complex texts of college and careers. The CCSS is the first standards document to recognize this fundamental feature of literacy. The writers of the CCSS are to be applauded for their recognition of a central feature of reading instruction that has often been ignored.

As is always the case in any human endeavor, translating vision to practice means that tough choices need to be made. This was evident in the CCSS writers' description of a staircase of text complexity. Decisions were made about such thorny issues as ways to measure text complexity and which texts exemplify complexity at different points along the staircase.

In this response, we consider three assumptions about the view of text complexity as operationalized by the CCSS. We are concerned that these assumptions, if left unexamined, could increase the achievement gap as they become part of state and national policies. At the outset, we emphasize that we support strongly the goal of increased reading of complex texts and accompanying reading practices. A complex view of text complexity, however, is needed to ensure that appropriate texts and instruction are provided to students in order to increase their capacity to engage with complex texts. Before addressing the three assumptions and their potential consequences, we describe why text complexity is included as a distinct standard within the CCSS.

Overview: Why Text Complexity As A Standard?

Previous standards documents of state and national organizations describe students' comprehension, or their recognition of features such as figurative language, with respect to grade-level texts, but grade-level is never identified. An unarticulated assumption underlay these documents: that the texts seventh graders would be applying a compare-contrast strategy on would be more challenging than those used to measure third graders use of a compare-contrast strategy on state assessments, and so forth. Without an index of text complexity for establishing grade-appropriate texts, text level was determined by publishers and test-makers.

The results of leaving text levels ill defined are evident from three groups of analyses. First, comparisons of students' performances on assessments administered by their states and those of the National Assessment of Educational Progress (NAEP) showed that grade-level proficiency varied wildly from state to state (e.g., Bandeira de Mello, 2011). Second, American students failed to perform at appropriate levels on international comparisons of literacy proficiency (Mullis, Martin, Gonzalez, & Kennedy, 2003). Finally, the ACT's (2006) study of the preparedness of high school students for college board examinations and the relationship of these performances to their grades in college showed that students needed to achieve benchmark scores on tasks with *complex* texts, not complex tasks with simple or uncomplicated texts, to pass college courses. Unfortunately, only 51% of high school students achieved this benchmark level.

The response of the CCSS writers was to make the ability to read increasingly more complex text a centerpiece of the standards with an entire standard devoted to increased capacity with complex texts over the grades. The CCSS includes an appendix (B) where exemplars of complex texts are provided for different grade bands, beginning with the grade 2-3 band. However, the core of the definition of text complexity is in Appendix A where a tripartite model

is provided as the basis for establishing text complexity. This model includes qualitative dimensions (i.e., levels of meaning or purpose, structure, language conventionality and clarity, and knowledge demands), reader and task dimensions (i.e., elements of instruction that teachers address in assignments and lesson planning), and quantitative dimensions. Within Appendix A, the CCSS suggest that further guidance on qualitative dimensions would be forthcoming but, within the document that was distributed only one quantitative system was well described and operationalized. The readability system that is the focus of Appendix A is the Lexile Framework (Metametrics, 2000) which uses, as is typical of readability formulas, a semantic component (frequency of vocabulary) and a syntactic component (number of words in a sentence). Specific levels were identified within the Lexile Framework for grade bands, starting with grade 2-3. The point of initiation for the grade 2-3 band (450 on the Lexile scale) implicitly establishes an expectation for grade one.

In the almost two years since the release of the CCSS, the promise of additional guidance on qualitative systems has not materialized. Further, the organization responsible for the writing of the standards—Student Achievement Partners (SAP)—has supervised a grant from the Gates Foundation to establish the comparative predictive validity of different readability formulas. The report from this project includes co-authors from SAP and suggests an even stronger rationale for the use of readability formulas (Nelson, Perfetti, Liben, & Liben, 2012). Actions to date suggest a need for further examination of the assumptions about text complexity within the CCSS. Three such assumption are to be discussed here and include: (a) Text levels need to be accelerated at every level of students’ school careers, including the primary grades; (b) Students at all proficiency levels can be rapidly “stretched” to read substantially harder texts; and (c) Readability formulas provide sufficiently valid assessments of text complexity that can be used

as a guide for selections in instruction and assessment. We examine each assumption, related research, and potential consequences for young readers in the United States.

Assumption 1: Text Levels Need to be Accelerated At Every Level of Students' School Career, Including the Primary Grades

K–12 reading texts have actually trended downward in difficulty in the last half century.... quantitative measures should identify the college- and career-ready reading level as one endpoint of the scale.
(CCSS, p. 8).

The CCSS is recommending a reconfiguration of grade-level readability standards in order to prepare high school graduates for the texts of college and careers. Table 1 shows how the CCSS proposes to achieve this goal—by raising text levels in all grades, starting with second grade. There is another way to achieve this goal: create a strong foundation in the early grades and then raise the difficulty exponentially in the upper grades so that the increase, if it were graphed, would look like an up-sweeping curve. We believe that the research supports this second approach. But first, we'll examine what the CCSS proposes and the claims and sources used to support their recommendations: (1) that texts in all grades, including primary-level, have trended downward in difficulty over the years so that all students are reading easier material than they used to, and (2) that by accelerating difficulty in primary level texts, students will be better positioned to meet college and career readability expectations by high school graduation.

First, CCSS's recommendations assume that readability levels of texts at every grade level have trended downward over the past 50 years. Several papers refute the idea that primary grade texts have been simplified in kindergarten (Hiebert, 2011a), in first grade (Hiebert, 2010), or in third-grade (Hiebert, 2012a), and those arguments are summarized here. First, with regard

to kindergarten, it is impossible that expectations could have declined over the past 50 years since kindergarten texts were not part of core reading programs until after Reading First mandates in the first decade of the 21st century.

With regard to first grade, the first study cited in the CCSS is Chall's 1967/1983 reports on first-grade texts that summarized the features of texts in core reading programs with copyrights from 1956 to 1962. Massive changes occurred in first-grade texts in the subsequent decades (Foorman, Francis, Davidson, Harm, & Griffin, 2004), including the elimination of controlled vocabulary in first-grade texts that resulted in substantial increases in the number of unique and rare words.

The second study cited by the CCSS with implications for the primary grades is an analysis conducted by Hayes, Wolfer, and Wolfe (1996) of texts from grades one through eight and across three time periods (1919-1945, 1946-1962, 1963-1991). Hayes et al. showed numerous changes in texts over these three periods, including evidence of schoolbook simplification. But this pattern of simplification did not hold for the primary grades. In grade three, vocabulary difficulty was *highest* for the final period in that study.

CCSS also cites Williamson's (2008) analysis of a cohort of North Carolina students in grades 3-8 to show texts had become easier between the years 1999 and 2004. The end-of-the-year level for third grade on this curve was 700L. Table 1 shows that the exit level of the grade two-three band of the CCSS is 790, 1.6 standard deviations above the 700L level at which proficient third graders in North Carolina were performing.

The most recent evidence that the CCSS brings to bear on this claim of decreasing text difficulty is the ACT (2006) study that has already been described. The ACT research does not speak to primary level readabilities; the study begins with eighth-grade students.

To review, the data point only to a text-complexity gap at the middle and high school levels. None of these studies provides any evidence that primary-grade text difficulty has declined or that increasing grade-level reading expectations for the second and third graders will boost their trajectories as successful readers in the secondary-grades. This brings us to the next claim associated with CCSS's assumption: that "the earlier children learn to read, the better they will do." Despite substantial investments in early reading through No Child Left Behind/Reading First, gains have not been evident in higher grades (Gamse, Jacob, Horst, Boulay, & Unlu, 2008; Jackson et al., 2007). These findings are echoed by international data that tracks the effects of early reading instruction on later reading achievement (Suggate, 2009). Differences are not evident in the reading achievement of students at the end of elementary school in different countries as a result of different school entry ages. However, in countries with earlier starting ages, the achievement gap between those who were proficient and those who were not grew larger among 15-year-olds.

Potential Consequences

The stance of the CCSS to have students reading earlier and at higher levels will probably not help, nor hurt, those who come to school ready to read. However, it could make literacy a greater challenge for the students who primarily depend on America's public schools for their literacy instruction. At the present time, two-thirds of an American fourth-grade cohort do not meet current reading goals on the NAEP—one third of the cohort falls below the basic standard, another third below the proficiency standard (National Center for Educational Statistics, 2009). According to this framework for determining reading proficiency, students in the middle of fourth grade who can recognize the majority of what is considered fourth-grade vocabulary (Daane, Campbell, Grigg, Goodman, & Oranje, 2005) are considered proficient readers. Students

scoring in the below-basic and basic groups fail to read a portion of these words or read them too slowly to be proficient readers. Therefore, when increasingly difficult text is presented, students on the low end of either the proficient or basic groups drop down into basic and below basic, as *A First Look* (ACT, 2010) study verifies with eleventh graders.

CCSS recommendations could also increase the downward movement of formal reading instruction into kindergarten. As Pearson and Hiebert (2010) have illustrated, the proficiency distribution in kindergarten mirrors the pattern of fourth grade cohorts: one third enters kindergarten having mastered literacy content, another third has incomplete knowledge of this content, and the final third is entirely dependent on kindergarten to deliver this content. Knowing student experiences and current performance is essential when making informed curricular decisions such that learning can occur for all learners. If we fail to do this and instead craft ill-informed and/or arbitrary expectations the result will likely be separating, sorting and labeling students who “can/can’t” meet such expectations instead of the kinds of teaching and learning opportunities students really need.

Another consequence could be that the age at which children begin school will play a greater part in determining early academic success, as children born in the winter or spring will have a six to nine month advantage over their summer or fall born peers. Holding children back for an extra year will be a luxury that only some parents who can afford the childcare can indulge—another way in which this trend hurts students who depend upon public services. Finally, apart from any generalized group tendencies, there are individual differences in cognitive development rates that are not indicative of intelligence or future achievement. Any children given some breathing room in their early years of school might be able to catch up to

grade level standards. The pressure of higher expectations in the early grades may mean no grace period for these students.

Students will quickly come to know who is succeeding and who isn't. Such branding, whether it comes from the system or the students own observations about who is successful in class, can have long-lasting and often intractable effects on students' confidence as readers. This can, in turn, diminish their interest and willingness to engage in reading activities (Guthrie, Hoa, Wigfield, Tonks, Humenick, & Littles, 2007).

While a deeper reading of primary-grade focused research cited in the CCSS shows that texts have not grown easier over the years, there is research that points to another way to ensure that primary-level students are meeting current expectations. Research shows that students in third grade who achieve proficient literacy levels (under existing, not new and increased standards) are less likely to drop out of high school (Hernandez, 2011). Our energies would be better focused on attaining this useful goal that reflects the standards that are already in place, rather than grasping after an unsupported, aspirational standard that is even further out of reach.

**Assumption 2: Students at all levels can be rapidly stretched
to read substantially harder texts**

Students in the first year(s) of a given band are expected by the end of the year to read and comprehend proficiently within the band, with scaffolding as needed at the high end of the range. Students in the last year of a band are expected by the end of the year to read and comprehend independently and proficiently within the band (CCSS, p. 10).

This quote from the CCSS suggests students are to be stretched to read texts within an identified grade level band—first with scaffolding and then on their own with the expectation

that they can do so proficiently. Taking a closer look at this suggestion, and the texts recommended within Appendix B as exemplars, raises questions and concerns including issues of how texts are defined, how text complexity/difficulty is measured, and how the outcomes of interactions between readers and texts are defined and measured.

When it comes to texts and text selection, reading instruction has long been influenced by frameworks that focus on texts, ways of leveling texts, and then pairing particular texts with particular readers. For example, Betts' framework (1946), which assumes links between oral reading accuracy, reading comprehension, and possible emotional response (e.g. frustration), continues to influence instructional decisions within classrooms, schools, and policies (Halladay, 2008). A recent instantiation of this line of thinking is reflected in the CCSS's reference to text complexity and subsequent recommendations of grade-level anchor texts that further define particular perspectives toward texts, readers, learning, and instruction. One could challenge the footings on which this document attempts to stand in a number of ways, but we think it is more productive to think about the *readers* that walk into our schools and the kinds of *readers* that need to participate in present and future society.

The CCSS works with the notion of a generic x-grade reader. Readers are far from clones of other age peers. Readers of any age come to texts with distinct histories and diverse experiences with the purposes of reading—types and nature of different reading tasks, cultural and background knowledge, experience with words, print, and language(s), and academic positioning (including their own and others definitions of successful reading). It is imperative that any conversation or decision aiming to match readers and texts in ways that support or lead readers' development consider reader context interactions with texts, as well as text-based

factors, to create ideal instructional scenarios that offer the potential for engaging and teaching readers.

Readers come to texts with particular ideas and expectations often grounded in their cultural lives (Murata, 2007). Links between readers' prior and/or background knowledge and comprehension in first and/or additional languages has long been on the mind of educators and is well studied by researchers (Gee, 2000). In a recent study of fifth graders' social awareness, Dray and Selman (2011) showed that readers' social knowledge and experience, while moderated by reading skill, influenced comprehension. For good readers, there was almost no effect. But for those classified as poor readers, there was a fairly strong positive effect suggesting that readers' experiences influenced their comprehension of the focal text. While the study does not claim to illustrate causality, inferences suggest that the relationship between the text content and the lived experiences of the reader matter, especially for less experienced readers.

Readers not only vary in the experiences they bring with them into classrooms, they vary in terms how they approach reading—which may be based on their definitions of success, repertoire and use of reading strategies, social position, and/or engagement. Halladay's (2008) research into the lives of early readers finds that often young people's visions of successful reading varies from expectations for success held by other stakeholders—namely family, educators, policy makers. Findings from a study conducted by Blaxall and Willows (1984) highlight the ways readers vary in terms of known strategies as well as how/when to orchestrate known strategies. Their findings suggest that more skilled readers have greater flexibility when it comes to using reading strategies as evidenced by more syntactically or semantically acceptable miscues. Less experienced readers lean on graphophonemic cues that result in the dominance of graphically similar substitutions likely to disrupt meaning making. While focusing their research

primarily on young learners writing lives, Bomer & Laman's (2004), findings challenge linear and purely cognitive development trajectories. Their research emphasizes the complex social work in becoming literate, which includes the ways in which learners are perceived and positioned by others and their goals for any given literacy event. Guthrie, McGough, Bennett, & Rice (1996) and Cambourne's (1995) work focus on the social conditions that foster engagement. Their findings suggest that engaged readers are knowledgeable, strategic, motivated, and socially interactive. These findings point to the complexity of readers—their histories, expectations, skills, motivations, and surrounding learning conditions all influence what they do with texts, the needs and experiences they bring to a reading event, and the diverse range of what they need for the best learning to transpire.

Another part of puzzle concerned with making informed decisions about matching readers with texts, which was alluded to in the discussion of the Dray and Selman study but merits direct consideration, has to do with definitions, knowledge, actions, and contexts of reading. In other words, some research examines particular dimensions of reading practices and may or may not take into consideration all or enough factors to use such research to support claims about what makes a text “right” or appropriately “complex” for particular readers.

Consider research on the role of vocabulary. In one study, findings suggest that the percentage of high frequency words in a given text was a predictor of performance for low and average ability readers in terms of decoding, reading accuracy, and fluency (Compton, Appleton, & Hosp, 2004). Unfortunately, comprehension was not examined and we therefore do not know what sense readers made from text, even when the percentage of high frequency words impacted performance in measured areas.

Or consider what has been found on the topic of matching readers and reading material in independent reading contexts. Carver & Leibert's (1995) research suggests that mere reading of easy or instructional matched books alone is insufficient for increasing readers' skills. This may mean that such texts when placed in the hands of novice readers are insufficient for increasing readers' skills. This could be taken to mean learning to read requires more "complex" texts but, questions can be raised with regard to ways in which independent reading may have impacted engagement, and/or how level "determinations" were made, or how reading skills were designed and measured. A sweeping decision about the role of "easy" or "instructional" texts and the merits of independent reading seems unwise as other related research suggests that there may be some gain for older readers who engage in independent reading (Krashan & McQuillian, 2007) and for less fluent, marginalized readers who are less likely to have access to print in their homes (Kim, 2006).

Findings like these may push some to shift from "determining the right text" to inviting readers to choose. More studies reveal partial insights into the merits and concerns of such a shift. Donovan, Smolkin, and Lomax's (2000) findings suggest that the most emergent readers, when presented with choice of text, more often make text selections that are difficult. However, this research, which tracked 1st grade student selections in relation to determined reading levels, did not gather data with regard to the nature of reading experiences. Hence, we do not know what readers took away from and/or learned during/from their interaction with texts that were classified as difficult. Even still, this research does raise questions about the roles of interest and motivation, as well as conceptual and genre knowledge, in drawing readers to texts and engaging readers with texts for extended periods of interaction/reading. When Kim & Guryan (2010) examined choice in their study of 4th grade Latino students' summer reading, they found that

self-selected reading alone or with minimal support in the form of family literacy workshops impacts growth in reading ability. Again, this research points to the complexity of matching texts, readers, and instruction.

Potential Consequences

Focusing exclusively on texts, identified level of complexity, and accompanying expectations for achievement ignores the nuances of being and becoming a reader. It is probable that raising the bar to a particular height and then asking kids to jump harder and higher could result in kids who can say the words in a text or get through identified texts becoming readers who can decode, but choose not to actively read. Furthering the distance between where learners are and where they are expected to be will likely further marginalize or deny access to the literacy club for students whose experiences, interests, and cultures are not part of a leveled canon. As reading begets better readers, turning students off to reading affects our national aspirations as well. Already U.S. students rank among the lowest of any country in their interest in reading (Mullis, Martin, Gonzalez, & Kennedy, 2003). We also predict that the CCSS staircase will result in less proficient students spending less time reading. The CCSS advises that students who are reading at the low end of their grade band be exposed to the texts at the high-end of their grade band. Often, exposure to difficult texts takes the form of read-alouds or read-alongs (either facilitated by teachers or digital devices) so that students who cannot read difficult texts independently can experience them with scaffolding. Research shows that read-alouds can support students' listening comprehension, but to date, we do not know if such experiences help students develop independent reading proficiency. Depending on the distribution of instructional time, such expectations may disadvantage the most emergent readers.

The early emphasis on increasingly difficult texts, as outlined in the CCSS, could also create and reify simple definitions of what it means to read, to the point where students are miles away from critically engaging with the wide ranges of texts today's young readers are bound to encounter in their reading lives. Given this modest examination of research, it can be argued that a focus on the text alone in the hands of an emergent or novice readers will not produce the gains envisioned/called for by CCSS. It follows that while we need to think more about what is meant by "complex" when it comes to text, it's likely we ought to think more about the meaning of "complex" when it comes to readers, acts of reading, and reading instruction.

Assumption 3: Readability formulas provide sufficiently valid assessments of text complexity that they can be used as a guide for selections in instruction and assessment

...quantitative measures should identify the college- and career-ready reading level as one endpoint of the scale. MetaMetrics, for example, has realigned its Lexile ranges to match the Standards' text complexity grade bands and has adjusted upward its trajectory of reading comprehension development through the grades to indicate that all students should be reading at the college and career readiness level by no later than the end of high school. (CCSS, Appendix A, page 8)

We can all agree that text complexity is critical and that there needs to be progression in text complexity across the school years (although we can disagree where the points of greatest change needs to be and also the ways to support struggling and beginning readers in increasing their capacity). As with so many things in education, it is the measurement of text complexity that can greatly influence what happens. As the adage goes, the medium can be the message.

The tripartite system of the CCSS is described at the beginning of this chapter where qualitative and reader-task features are part of a system for establishing text complexity with quantitative measures. The assessments of readers and tasks are viewed to be within the purview of classroom teachers working with students, although particular features of readers and tasks can be presented within a rubric to support teachers in making these choices (see, e.g., Hiebert, 2011b).

With respect to support for qualitative measurement of texts, the CCSS writers (CCSS Initiative, 2010) indicated that examples of such systems would be forthcoming. At this point, fulfillment of this promise has been indirect through endorsement by CCSS writers of an effort developed by the Kansas state department of education (Pimentel, 2012). Within the Kansas rubric (Copeland, Lakin, & Shaw, 2012), four traits (levels of purpose, structure, language conventionality, and clarity/ knowledge demands) are described with generic descriptions at each of four levels (low, middle-low, middle-high, and high) for the two text types (narrative, informational). The Kansas system is at a nascent stage of development with vague descriptions, no examples of features in real texts, or evidence on reliability of the scheme. The offering of this system, after almost two years since the launch of the CCSS and the involvement of key CCSS writers in the already described predictive validity study of readability formulas (Nelson et al., 2012), suggest that readability formulas are likely to loom large within the evaluation of text complexity in assessments and text selection.

Nearly all readability formulas, regardless of small differences, analyze two main features of texts: (a) syntax and (b) vocabulary. Measuring syntax is fairly straightforward and is typically based on the number of words per sentence, although occasionally number of syllables is used. For vocabulary, some formulas (e.g., Spache, 1953) compare the words in a text to an

index of words that have been keyed to different grade levels, while others (e.g., Fry, 1968) use the number of syllables in words as an indicator of difficulty.

Until recently, readability formulas had to be applied manually by counting words or syllables and consulting word indices. Because of this, what was being measured by the readability formulas was abundantly clear to all who used them. The Lexile and several other digitized readability formulas take a different tack, which is to establish vocabulary complexity by taking the average frequency of all words in a sample of text (Smith, Stenner, Horabin, & Smith, 1989). The frequency of a word is established relative to its rank relative to all of the words in a digital database.

A readability formula is a good first resource in sorting large groups of texts. Hiebert and Pearson (2010) obtained Lexiles (L) of a large group of beginning reading texts that were sorted by text levels as assigned by publishers. The progression across seven levels showed a steady increase in complexity: 87L for the first level, 238L for the middle level, and 489L for the final level. However, within a level the variation was substantial. For the texts classified as the mid-point of beginning reading programs, Lexiles had a range of 760L (610L to -160L). This range is almost equivalent to growth expected through the end of grade three, according to the accelerated levels in Table 1.

Explanations for at least some of the discrepancies in readability levels are evident in a comparison of the Lexiles assigned to two texts within the exemplar texts in Appendix B of the CCSS (CCSS Initiative, 2010): *Boy, were we wrong about dinosaurs* (*Boy*; Kudlinski, 2005) and *In the time of the butterflies* (*Butterflies*; Alvarez, 1994). The Lexiles of the two texts fall closely to one another (within 60 Lexile points, the standard deviation in the Lexile Framework): 960L for *Boy* and 1000L for *Butterflies*). The CCSS writers offered these texts as exemplars for quite

different levels, however: *Boy* for the grade 2-3 band and *Butterflies* for the grade 9-10 band.

One feature of readability formulas that can explain such disparate outcomes is text genre. *Boy* is an informational text, while *Butterflies* is a narrative text. Narrative texts often contain dialogue and people typically speak in relatively short sentences. Further, a writer of a narrative often uses fairly straightforward vocabulary, sprinkled with rare vocabulary that provides nuance to the text. *Butterflies* contains words such as *anachuita* (name of a tree) and *guanabana* (name of a juice), but most of the words in the text are quite common. The mean of the frequency of words in a sample of *Butterflies* is 3.59 on a scale of 1 (hardest) to 5 (easiest) according to the Lexile Analyzer.

The mean for word frequency in *Boy*—3.49—indicates that the vocabulary is somewhat more difficult than that in *Butterflies*, even though *Boy* is intended for primary-level students. Rare words such as *iguanodon* are repeated often in this informational text, reflecting the need to use precise vocabulary to convey particular concepts (Cohen & Steinberg, 1983). Most, if not all, readability formulas do not compensate for this repetition of rare vocabulary within informational texts, even when repetition is known to support word learning (Finn, 1978). Readability formulas, then, can overestimate the difficulty of even a fairly straightforward informational text such as *Boy*, just as they underestimate the difficulty of a narrative with complex themes and content such as *Butterflies*.

Another feature of measuring vocabulary within digital readability formulas is also reflected in similar levels assigned to *Boy* and *Butterflies*. This feature is the use of relative rankings of words in large databases as the source for the vocabulary component within the formula. Using average word frequency is a tricky business because a few hundred words in English, about 2%, account for approximately 80% of the words that appear in texts (Zeno, Ivens,

Millard, & Duvvuri, 1995). The remaining 98% of the corpus appear much less frequently. In the Zeno et al. corpus, approximately 86% of the almost 150,000 unique words appeared less than once per million words of text.

Even when algorithms are used to normalize a distribution, the distribution is so skewed that many texts receive a similar index for vocabulary. For example, the mean of the vocabulary index within the Lexile equation for the exemplar CCSS texts is 3.5 and the standard deviation is .21. The measure of sentence length shows considerably more variation: $X=14.95$, $SD=4.8$. As a result of so many words with similar ratings, more pressure within the readability formulas is on sentence length (Deane, Sheehan, Sabatini, Futagi, & Kostin, 2006). Within the sample of CCSS exemplar texts, the correlation between the Lexile and mean sentence length is .92, while that for the vocabulary measure is -.46 (a negative correlation is expected since hard texts have lower vocabulary indices).

This heavy reliance on sentence length to define text difficulty creates its own set of problems, among them is the lack of knowledge as to how syntactic patterns can be taught, especially to English learners. Further, what we do know about syntax runs counter to assumptions made by readability formulas. According to the research, shorter sentences do not always make text easier. Short sentences tend to have fewer context clues and fewer links between ideas, requiring the reader to make more inferences. Even more troubling is the fact that sentence length is very easy to manipulate quickly. The consequences of such manipulation have been documented, as is discussed in the following section.

Potential consequences

The temptation to rely on quantitative indices can be all too easy. When a policy document provides specific ranges, the temptation can move to an imperative. Quickly, we could

return to the instructional and assessment contexts prior to *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985) when writers were often asked to write texts to satisfy readability formulas by using highly frequent words or short sentences (Davison & Kantor, 1982). The provision of specific quantitative ranges within the CCSS (see Table 1) could mean that states and large-city districts mandate particular readability levels in new assessments and textbook programs. Students could be reading inappropriate texts. For example, sixth graders (rather than high schoolers) could be asked to read texts such as *In the Time of Butterflies* (Alvarez, 1994) and fifth graders (rather than second graders) given texts such as *Boy, were we wrong about dinosaurs* (Kudlinski, 2005).

A second potentially unintended consequence is that publishers could return to rewriting texts to comply with readability formulas, rather than using the readability formulas as a general guideline (Davison & Kantor, 1982). A quick review of websites on the internet shows that some publishers are claiming to offer sets of CCSS-appropriate texts that have the same content but are at all of the grade-bands in Table 1. How would a publisher achieve such a feat, when the vocabulary of texts at different levels varies considerably across grades? Hiebert (2012b) has illustrated how a classic text—*The Wind in the Willows*—can be transformed from college and career readiness to first grade Lexile levels by simply manipulating syntax.

The third potential and unintended consequence of specific quantitative indices is perhaps the most egregious. An easy to obtain measurement of a text's complexity can give educators the sense that they have identified appropriate texts and matched particular students with the texts. To provide texts that support students on the staircase of text complexity, educators need to understand the features of text that contribute to students' challenges with comprehension, vocabulary, and word recognition. Quantitative data may be quickly obtained but it fails to

support teachers in understanding the features of complexity that challenge students at particular points on the staircase of reading.

Summary

By including a stand-alone text complexity standard, the CCSS shines a light on a issue worthy of research, discussion, professional learning, and informed classroom decisions with regard to texts and accompanying teaching practices. The effort was put together quickly—as is almost always the case with large-scale projects. We are hopeful, in writing this chapter, that there is an opportunity to consider potential consequences before their effects change the opportunities for thousands of students, especially those who live in high-poverty communities, to achieve high levels of literacy. The collective wisdom of scholars within the field of literacy needs to be brought to bear on the issues raised in this chapter—as well as to other issues (e.g., definitions of close reading). National conversations are needed to ensure that teachers can make informed decisions that support, as well as challenge, readers so that they not only learn to navigate increasingly more complex texts over time, but also become readers who engage with sophisticated texts as active citizens.

References

- ACT (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: Author.
- ACT (2010). *A first look at the Common Core and college and career readiness*. Iowa City, IA: Author.
- Alvarez, J. (1994). *In the time of the butterflies*. Chapel Hill, NC: Algonquin.
- Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers: The report of the Commission on Reading*. Champaign, IL: The Center for the Study of Reading, National Institute of Education, National Academy of Education.
- Bandeira de Mello, V. (2011), *Mapping State Proficiency Standards Onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005–2009* (NCES 2011-458). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC: Government Printing Office.
- Betts, E.A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York: American Book Company.
- Blaxall, J. & Willows, D. (1984). Reading ability and text difficulty as influences on second graders' oral reading errors. *Journal of Educational Psychology*, 76(1), 330-340.
- Bomer, R. & Laman, T. (2004). Positioning in a primary writing workshop: Joint action in the discursive production of writing subjects. *Research in the Teaching of English*, 38(4), 420-466.
- Cambourne, B. (1995). Toward an educationally relevant theory of literacy learning: Twenty years of inquiry. *The Reading Teacher*, 49, 182-192.

- Carver, R. & Leibert, R. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly*, 30(1), 26-48.
- Chall, J. S. (1967/1983). *Learning to read: The great debate* (3rd ed.). Fort Worth, TX: Harcourt Brace.
- Cohen, S.A., & Steinberg, J.E. (1983). Effects of three types of vocabulary on readability of intermediate grade science textbooks: An application of Finn's transfer feature theory. *Reading Research Quarterly*, 19(1), 86-101.
- Common Core State Standards Initiative. (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*. Washington, DC: CCSSO & National Governors Association.
- Compton, D., Appleton, A., & Hosp, M. (2004). Exploring the relationship between Considering the reader-text match in first graders' self-selections during recreational reading. *Reading Psychology*, 21, 309-333.
- Copeland, M., Lakin, J., & Shaw, K. (January 26, 2012). *Text complexity and the Kansas Common Core Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*.
http://www.ccsso.org/Resources/Digital_Resources/The_Common_Core_State_Standards_Supporting_Districts_and_Teachers_with_Text_Complexity.html
- Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M.J., & Oranje, A. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading* (NCES 2006-469). Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- Davison, A., & Kantor, R. N. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17(2), 187-209.

- Deane, P., Sheehan, K. M., Sabatini, J., Futagi, Y., & Kostin, I. (2006) Differences in text structure and its implications for assessment of struggling readers, *Scientific Studies of Reading*, 10(3), 257-275.
- Donovan, C. A., Smolkin, L. B., & Lomax, R. G. (2000). Beyond the independent-level text: Considering the reader-text match in first-graders' self-selections during recreational reading. *Reading Psychology: An International Quarterly*, 21(4), 309-333.
- Dray, A. & Selman, R. (2011). Culture and comprehension: A mixed methods study of children's responses to a fictional story about interracial conflict. *Reading and Writing Quarterly*, 27, 48-74.
- Finn, P.J. (1978). Word frequency, information theory, and cloze performance: A transfer theory of processing in reading. *Reading Research Quarterly*, 13(4), 508-537.
- Foorman, B. R., Francis, D. J., Davidson, K. C., Harm, M. W., & Griffin, J. (2004). Variability in text features in six grade 1 basal reading programs. *Scientific Studies of Reading*, 8 (2), 167—197.
- Fry, E. B. (1968). A readability formula that saves time. *Journal of Reading*, 11, 513-516, 575-578.
- Gamse, B.C., Jacob, R.T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First Impact Study Final Report*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Gee, J. (2000). Discourse and Sociocultural Studies in Reading. In M. L. Kamil, P. B. Mosenthal, D. Pearson, & R. Barr (Eds.). *Handbook of Reading Research v. 3* (pp. 195-207). Mahway, NJ: Lawrence Erlbaum.

- Guthrie, J. T., McGough, K., Bennett, L., & Rice, M. E. (1996). Concept-oriented reading instruction: An integrated curriculum to develop motivations and strategies for reading. In L. Baker, P. Afflerbach, & D. Reinking (Eds.). *Developing engaged readers in school and home communities* (pp. 165-190). Mahwah, NJ: Lawrence Erlbaum Associates.
- Guthrie, J.T., Hoa, A.L.W., Wigfield, A., Tonks, S. M., Humenick, N. M., & Littles, E. (2007). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology, 32*, 282-313.
- Halladay, J. (2008). *Difficult texts and the students who choose them: The role of text difficulty in second graders' text choices and independent reading experiences*. (Doctoral dissertation). Retrieved from ProQuest. (1685138491)
- Hayes, D.P., Wolfer, L.T., & Wolfe, M. F. (1996). Schoolbook simplification and its relation to the decline in SAT-Verbal Scores. *American Educational Research Journal, 33*(2), 489-508.
- Hernandez, D.J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Baltimore, MD: The Annie E. Casey Foundation.
- Hiebert, E.H. (June 11, 2010). *Have the texts of beginning reading been dumbed down over the past 50 years?* Retrieved from <http://www.textproject.org/frankly-freddy/have-the-texts-of-beginning-reading-been-dumbed-down-over-the-past-50-years/>
- Hiebert, E.H. (August 3, 2011a). *Is Reading in Kindergarten the Means for Ensuring College and Career Readiness?* Retrieved from <http://textproject.org/frankly-freddy/is-reading-in-kindergarten-the-means-for-ensuring-college-and-career-readiness/>
- Hiebert, E. H. (2011b). *Benchmark texts: Stepping up complexity*. Santa Cruz, CA: TextProject. Retrieved from <http://textproject.org/teachers/benchmark-texts-stepping-up-complexity/>

Hiebert, E.H. (2012a). The common core's staircase of text complexity: Getting the size of the first step right. *Reading Today*, 29(3), p26-27.

Hiebert, E. H. (February 27, 2012b). *Syntax and Text Complexity: A Classic Text Goes from College-Career Level to First Grade*. Retrieved from <http://www.textproject.org/frankly-freddy/syntax-and-text-complexity-a-classic-text-goes-from-college-career-level-to-first-grade/>

Hiebert, E.H., & Pearson, P.D. (2010). *An examination of current text difficulty indices with early reading texts* (Reading Research Report 10.01). Santa Cruz, CA: TextProject, Inc. Retrieved April 1, 2011 from: <http://www.textproject.org/research/an-examination-of-current-text-difficulty-indices-with-early-reading-texts/>

Jackson, R., McCoy, A., Pistorino, C., Wilkinson, A., Burghardt, J., Clark, M., Ross, C., Schochet, P., & Swank, P. (2007). *National Evaluation of Early Reading First: Final Report*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Kim, J. S. (2006). The effects of a voluntary summer reading intervention on reading achievement: Results from a randomized field trial. *Educational Evaluation and Policy Analysis*. 28(4), 335-355.

Kim, J. S. & Guryan, J. (2010). Efficacy of a voluntary summer book reading intervention for low income Latino children from language minority families. *Journal of Educational Psychology*, 102(1), 20-31.

Krashan, S. & McQuillian, J. (2007). The case for late intervention. *Educational Leadership*, 65(2), 68-73.

Kudlinski, K.V. (2005). *Boy, were we wrong about dinosaurs*. New York, NY: Dutton.

Lobel, A. (1970). *Frog and toad are friends*. New York, NY: HarperCollins.

- MacLachlan, P. (1985). *Sarah, Plain and Tall*. New York, NY: HarperCollins.
- MetaMetrics. (2000). *The Lexile framework for reading*. Durham, NC: Author. [Online].
Available: http://cdn.lexile.com/m/cms_page_media/135/The-Lexile-Framework-for-Reading.pdf
- Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary school in 35 countries*. Chestnut Hill, MA: International Study Center, Boston College.
- Murata, K. (2007). Unanswered questions: cultural assumptions in text interpretation. *International Journal of Applied Linguistics*, 17(1), 38-59.
- National Center for Education Statistics (2009). *The Nation's Report Card: Reading 2009* (NCES 2010-458). Institute of Education Sciences, U.S. Department of Education, Washington, DC.
- Nelson, J., Perfetti, C., Liben, D., & Liben, M. (2012). *Measures of text difficulty: Testing their predictive value for grade levels and student performance*. New York, NY: Student Achievement Partners.
- Pearson, P.D., & Hiebert, E.H. (2010). National reports in literacy: Building a scientific base for practice and policy. *Educational Researcher*, 39(4), 286-294.
- Pimentel, S. (January 26, 2012). *The Common Core State Standards: Supporting districts and teachers with text complexity*. Webinar presented through the Council of Chief State School Officers. Retrieved at:
http://www.ccsso.org/Resources/Digital_Resources/The_Common_Core_State_Standards_Supporting_Districts_and_Teachers_with_Text_Complexity.html
- Smith, D., Stenner, A.J., Horabin, I., & Smith, M. (1989). *The Lexile scale in theory and practice*

(Final report). Washington, DC: MetaMetrics. (ERIC Document Reproduction Service No. ED 307 577)

Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal*, 53, 410–413

Suggate, S.P. (2009). School entry age and reading achievement in the 2006 Programme for International Student Achievement (PISA). *International Journal of Educational Research*, 48, 151-161.

Williamson, G.L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, 19 (4), 602-632.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York, NY: Touchstone Applied Science Associates.

Table 1¹

Original and Recalibrated Lexile Ranges for CCSS Grade Bands

Text Complexity Grade Band	Original Lexile Ranges	Recalibrated Lexile Ranges
K-1	N/A	N/A
2-3	450-725	450-790
4-5	645-845	770-980
6-8	860-1010	955-1155
9-10	960-1115	1080-1305
11-CCR	1070-1220	1215-1355

¹Summarized from CCSS, Appendix A, p. 8.