Essential Insights into Text Complexity and the New Assessments: Three Articles from *Reading Today*

Elfrieda H. Hiebert

TextProject &

University of California, Santa Cruz

Table of Contents

Permission was obtained from the International Reading Association to reprint the following articles previously published in *Reading Today.*

The Common Core's Staircase of Text Complexity:
Getting the Size of the First Step Right

by Elfrieda H. Hiebert

"K–12 reading texts have actually trended downward in difficulty in the last half century" — Common Core State Standards, English/Language Arts (CCSS), Appendix A, page 2[1]

There can be no argument that texts used in high schools—and even those in middle schools—have been "dumbed" down over the past 50 years. If high school graduates are to be reading texts of college and careers, text difficulty levels for middle- and high-school students *do* need to be raised in accordance with the staircase of text difficulty identified by the CCSS.

However, it is not at all clear that the increase in text difficulty levels needs to start with primary-level texts, as is recommended in the CCSS staircase of text complexity. The claim that K-3 texts have been dumbed down over the past 50 years is simply not true. With respect to kindergarten, there were *no* kindergarten texts in core reading programs 50 or even 20 years ago.[2] The difficulty levels of kindergarten texts in current core reading programs are comparable to those of first-grade texts in the 1980s.[3]

For first-grade texts, the dumbing down claim was true when Jeanne Chall made it in 1967. Chall based her claim on a review of texts used between 1956-1962.[4] However, first-grade texts have changed a great deal since then. Shortly after Chall's critique, basal publishers retired characters such as Dick and Jane and along with them word repetitions (e.g., Run, run, run) and other "controlled" features that dominated beginning reading texts. By the late 1980s, controlled text had been entirely eliminated[5] from core reading programs.

In fact, an obstacle for many beginning readers today is that the texts are too hard. Several reviews describe the changes in first-grade texts that started in the late 1980s and remain prominent

Elfrieda H. Hiebert

regardless of whether the texts come from leveled books, decodables, or trade books.[6] In the texts of the 2010s, beginning readers must process large numbers of new words—typically 25 or more new words for every 100 words of text (regardless of the program's philosophy). The majority of words in today's beginning reading programs are included among the 300 most-frequent words in written English. However, many of the other words in the text—around 40%—appear a single time. Texts with many new words that are rarely if ever repeated make it hard for beginning readers to develop automaticity with core sound-letter patterns and critical words.

How much harder can the texts for primary-level students get? The CCSS suggests a whole lot. On the Lexile (L) scale that the CCSS uses for its staircase of text complexity, the step for grades two-three ends at 790L, approximately one grade level higher than previous recommendations. Two-thirds of the American fourth-grade cohort is failing to reach the current proficient reading standard on the National Assessment of Educational Progress.[7] Before raising the size of the first step on the staircase of text complexity and asking young students to climb bigger steps faster, two critical questions about the CCSS's staircase of text complexity need to be addressed.

First, what is the evidence that raising levels of text complexity, especially for primary-level texts, fosters the goal of college and career readiness? In particular, what is the evidence that attaining the 790L point at the end of third grade is necessary to be on track for college and career reading at high school graduation? Existing evidence suggests that exiting third-graders who read texts in the range of 540L to 585L proficiently

are likely to be successful in subsequent grades. This complexity level—540L to 585L—is approximately 200L to 250L *lower* than the recommendation of the CCSS. The evidence for 540L to 585L as a reading proficiency standard for exiting third graders comes from two venerable sources.

The first comes from the report *Double Jeopardy*[8] that used the National Longitudinal Survey of 4,000 children from birth through young adulthood. This study reported that students who left third grade without proficiency on a standardized reading test (the texts of tests average around 540L[9]) were unsuccessful in subsequent grades.

The second source of evidence is Jeanne Chall—the scholar who first identified both the dumbing down of text and the fourth-grade slump (a phenomenon among students who leave third grade without reading on-grade level). In 1996, Chall[10] and her colleagues identified benchmark texts from Grades 1-12. Panels of teachers and school administrators validated these benchmark texts. The third-grade texts, including science and social studies texts, have an average complexity level of 585L—200L lower than the third-grade exit level recommended in the CCSS.

The second question asks: Why aren't more third-graders able to read at proficient and advanced levels on the research-based standard identified by the NLS and Chall? At present, two-thirds of a third-grade cohort fails to attain the proficient standard with *current* levels of text complexity.[11] What if, rather than being harried by another new standard that has yet to be validated, we were to do some serious soul-searching? What if, rather than asking what is hot and not, we asked what is working and what is not? For example, how well do current texts and pacing guidelines support beginning and struggling readers?

To read proficiently at third grade means that students read many informational texts, use media, think critically about texts, and write responses to what they read—all prominent recommendations in the CCSS. But it is not at all clear that college and career readiness at high school graduation will be supported by raising the height of the staircase step at third-grade and asking young students to jump higher and faster. Before we increase the levels of text complexity in primary-level reading programs, we need to examine why it is that so many exiting third graders are not reading proficiently at current complexity levels—levels that

are linked to future school success. Until we do that, the pursuit of "harder, faster, earlier" will do little to support the many students who depend on schools to become literate at the levels required for the digital age.

Elfrieda H. Hiebert is President and CEO of TextProject, a California nonprofit whose mission is to bring beginning and struggling readers to higher levels of literacy through a variety of strategies and tools, particularly the texts used for reading instruction. She is also a member of the Division of Social Sciences at the University of California, Santa Cruz, hiebert@textproject.org.
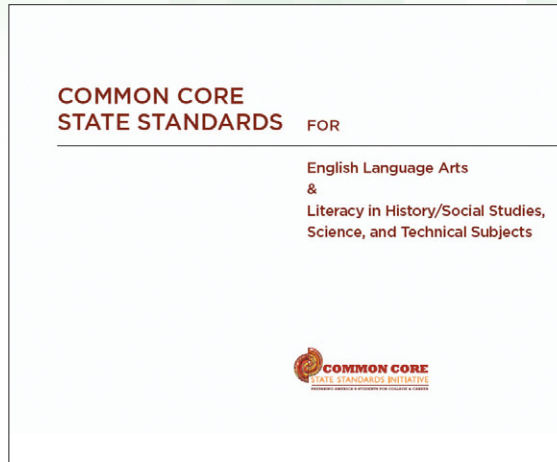
References

1. 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.

2. Morrow, L.M., & Parse, R. (1990). Early literacy strategies: Activities represented in current basal readers. In J. Zutell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms (39th Yearbook of the National Reading Conference).* Chicago, IL: National Reading Conference, Inc.

3. Hiebert, E.H. (2008). The (mis)match between texts and students who depend on schools to become literate. In E.H. Hiebert & M. Sailors (Eds.), *Finding the right texts for beginning and struggling readers: Research-based solutions* (pp. 1-21). NY: Guilford.

4. Chall, J. S. (1967/1983). *Learning to read: The great debate* (3rd ed.). Fort Worth, TX: Harcourt Brace.

5. Hiebert, E. H. (2005). State reform policies and the task textbooks pose for first-grade readers. *Elementary School Journal, 105,* 245-266.

6. 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. Hiebert, E. H. (2005). State reform policies and the task textbooks pose for first-grade readers. *Elementary School Journal,* 105, 245- 266.

7. 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, D.C.

8. Hernandez, D.J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation.* Baltimore, MD: The Annie E. Casey Foundation.

9. Hiebert, E.H. (2002). Standards, assessment, and text difficulty. In A.E. Farstrup & S.J. Samuels (Eds.). *What research has to say about reading instruction* (3rd Ed., pp. 337-369). Newark, DE: International Reading Association.

10. Chall, J. S., Bissex, G. L., Conard, S. S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty.* Cambridge, MA: Brookline Books.

11. Hernandez (2011)

COMMON CORE STATE STANDARDS FOR

English Language Arts &
Literacy in History/Social Studies,
Science, and Technical Subjects

COMMON CORE
STATE STANDARDS INITIATIVE

The CCSS Text Exemplars: Understanding Their Aims and Use in Text Selection

by Elfrieda H. Hiebert

The Common Core State Standards initiative (CCSS; CCSS Initiative, 2010) is the first standards document to have a separate standard (Standard 10) devoted to students' increasing capacity with complex text. Further evidence of text complexity's central role in the CCSS is its emphasis in the two appendices to the CCSS: Appendix A which addresses measurement of complex text and Appendix B which provides lists of texts which exemplify complexity.

Since text complexity has not been directly addressed in either state standards or reading programs over the past 25 years, considerable confusion exists about what text complexity means. The exemplar texts in Appendix B are part of the confusion. In some educational contexts, the exemplar texts from Appendix B have been embraced as the new curriculum. A quick search on the Internet shows that these texts can be purchased as programs and that blogs are devoted to lessons on them. Exemplar texts, however, are intended as models or examples, not as mandates or prescriptions. The aim of this article is to give educators grounding in exemplar texts by: (a) providing background, (b) illustrating how to select exemplar texts, and (c) describing their use in practice.

What Are Exemplar Texts?

Exemplar texts may be a new concept for many educators, but exemplars have a long history in reading and writing assessment. The National Assessment of Educational Progress uses such a system to evaluate students' written responses to comprehension questions. First, model responses are selected for advanced, proficient, basic, and below-basic performances. Then evaluators use these model responses to evaluate students' responses.

The CCSS's list of exemplars has a similar purpose: "to serve as useful guideposts in helping educators select texts of similar complexity, quality, and range for their own classrooms" (CCSS Initiative, Appendix B, 2010, p. 2). The CCSS writers claimed a prior project (Chall, Bissex, Conard, & Harris-Sharples, 1996) as the basis for using exemplar texts for text selection. But the CCSS departs from Chall et al.'s use of exemplars in several significant ways—departures that have contributed to educators' confusion about the role of the exemplar texts listed in Appendix B.

The biggest difference between the exemplars of Chall et al. (1996) and the CCSS is evident in Table 1: number of texts. Chall et al. provided a single text for a grade level and for each of six text types. All grade-level texts formed a scale for a text type (e.g., Literature). Each text for a grade level and text type had been selected after a lengthy process in which numerous texts were analyzed quantitatively and qualitatively.

By contrast, the CCSS provides a list of many texts for each grade band and text type (e.g., 13 stories and 12 informational texts for the grades 2-3 band). Standard 10 refers specifically to the need for students to have capacity with texts "at the high end" of a grade span's text complexity band" (CCSS, p. 14). But which of the texts in the list of exemplars represent the high end? On the grades 2-3 list, *Cowgirl Kate and Cocoa* and *The Fire Cat* are obviously at the low-end of the range but what about other texts? Are all complex for end-of-grade-three? Some for end-of-grade-two?

Identifying a list of potential texts to use as exemplars is only a first step. For exemplars to be useful guideposts, they need to be tied to specific points of reading development.

Table 1.

Comparison of Exemplar Texts in Chall et al. and Common Core State Standards

	Literature	Popular Fiction	Life Sciences	Physical Sciences	Narrative Social Studies	Expository Social Studies
Chall et al. (1996) Exemplars for Grade 3	*Magic Finger*	*Freckle Juice*	*Frogs & Toads*	*The Starry Sky*	*Phoebe and the General*	*The Fourth of July Story*
Exemplars for Grades 2-3	13 texts: •*Amos & Boris* •*Cowgirl Kate and Cocoa* •*Henry and Mudge: The First of Their Adventures* •*My Father's Dragon* •*Poppleton in Winter* •*Sarah, Plain and Tall* •*The Fire Cat* •*The Lighthouse Family: the Storm* •*The One-Eyed Giant* •*The Raft* •*The Stories Julian Tells* •*The Treasure* •*Tops and Bottoms*	12 texts				

## What Criteria should be used to Select Exemplar Texts?

Turning a list of texts into a true set of exemplars involves close study of texts and reading proficiencies associated with specific points in reading development. The basic steps in this process are sketched out with the 13 narrative texts for grades 2-3 from the CCSS's Appendix B.

To recognize the range of development during grades 2-3, six points have been identified across these two grades (see Table 2). The first step in matching texts with these six points of development was to conduct quantitative analyses of texts—vocabulary load, sentence length, cohesion, and text length. After these analyses, two texts were eliminated: one too simple (*Cowgirl Kate*) and the other too complex (*Amos and Boris*).

Three experienced primary-level teachers used the qualitative categories offered by the CCSS (levels of meaning, knowledge demands, language conventions and clarity, and structure) to analyze the remaining 11 texts. After independent analyses, the teachers discussed and finalized their choices. As well as agreeing on placement of exemplars at different points across grades 2-3 (see Table 2), the team decided to add a fifth criterion: length. A text of 50 or more pages (e.g., *Sarah, Plain & Tall*) requires students to retain information in a way that a picture book (*Tops and Bottoms*) does not. This grounding of texts to levels illustrates how exemplars are identified.

How Can Teachers Effectively Use Exemplar Texts?

Exemplars such as those in Table 2 are intended for all those involved in the instructional effort to use in selecting texts. A teacher might use the exemplars to decide which texts to use for lessons with particular students. Librarians might base recommendations of texts to students on
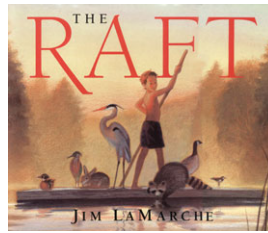
These exemplars represent terminal points on a range of reading development for grades 2-3.

the match of texts to exemplars. Publishers might use these models in ensuring appropriate selection of texts for programs.

Students should also be aware of the exemplars. These texts can be the focus of students' goal-setting and also of communication between teachers, students, and families. Students who are reading texts similar to *The Raft* at the end of third grade can see their progress from the beginning of the year when they were reading texts similar to *Henry and Mudge* but they are also aware that they have substantial ground to cover to be fully

prepared for the texts and tasks of grade four. A summer reading program might be put in place to support this growth.

Finally, the *process* of identifying exemplar texts is an important means for teachers to increase their knowledge of text complexity. Through discussions in schools, districts, and states, teachers deepen their facility in selecting texts and in conducting lessons which support students' increased capacity with complex texts.

Exemplar texts are a critical part of the text selection process, moving educators away from a reliance on quantitative indices. The CCSS Appendix B was a start but, without the connection to proficiencies, the lists of texts have left many believing that a curriculum has been sanctioned. Exemplars are not the texts of instruction but rather aid educators in selecting appropriate instructional texts. Exemplars are there to lead the way—not to be the curriculum.

References

1. Chall, J.S., Bissex, G.L, Conard, S.S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty.* Cambridge, MA: Brookline Books.
2. 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.

Elfrieda H. Hiebert is President and CEO of TextProject, a California nonprofit whose mission is to bring beginning and struggling readers to higher levels of literacy through a variety of strategies and tools, particularly the texts used for reading instruction. She is also a member of the Division of Social Sciences at the University of California, Santa Cruz, hiebert@textproject.org. ∎

Table 2.

Narrative Book Exemplars (Grade 2-3 Band)

Level	Title	Curriculum	
		Word Recognition[1]	Comprehension
2.1	•*The Fire Cat* •*Henry & Mudge*	600 most-frequent word families (MFWF) + less-frequent (LF) words with 1-3 letters	Follow story structure with common dilemma
2.2	•*Poppleton in Winter*	1,000 MFWF + LF 1-4 letters	Follow story structure with somewhat less familiar content and vocabulary
2.3	•*Tops and Bottoms* •*The Treasure*	1,000 MFWF + LF 1-5 letters	Understand new text structures
3.1	•*The Stories Julian Tells*	1,500 MFWF + LF 1-5 letters	Features of story structure become more multi-dimensional
3.2	•*The Lighthouse Family* •*The Raft*	2,000 MFWF + LF 1-5 letters	Maintain thread of story in longer texts and more complex plots and vocabulary
3.3	•*Sarah: Plain & Tall* •*My Father's Dragon* •*The one-eyed giant*	2,500 MFWF + LF 1-6 letters	Understand increasingly more complex character traits & literary language

[1] Approximately 96% or more of words in a text fall into the designated word recognition curriculum.

For the CCSS Assessments and Beyond:
Develop Your Students' Stamina
for Grappling with Complex Texts

by Elfrieda H. Hiebert

The inclusion of a standard on text complexity represents the most unique of several distinguishing features of the Common Core State Standards (CCSS) for English Language Arts. By reading texts at ever-accelerating complexity levels over a school career, students are expected to be proficient with the texts of college and careers upon high school graduation. Teachers are working vigorously to identify how to best apply this standard to their daily classroom instruction.

However, in the frenzy over ensuring that classroom texts are complex enough, little attention has been paid to the contexts in which students' reading of these complex texts will be assessed. A key distinction between assessment and instruction is what students are asked to *do* with texts.

In many elementary and even middle school classrooms, teachers read a new text aloud to students. In subsequent lessons, teachers often set the pace and content of students' reading of texts. That isn't going to happen during assessment. In the assessment context, students will need to read these texts on their own—an activity that existing evidence indicates is challenging for many students (Hiebert et al, 2010). Further, students must use evidence from what they have read to respond to questions and to write essays.

We can get a glimpse of the scope of this new challenge by examining the sample texts of the two new assessment consortia—Smarter Balanced Assessment Consortium (SBAC) and Partnership for Assessment of Readiness for College and Careers (PARCC). Table 1 provides a summary of time and task features of the assessments. Already at third grade, students are expected to spend considerable time reading on their own.

Assessment Texts: Much Harder?

The presence of a new standard on text complexity raises the question of how much harder the assessment texts will be from the typical texts of classrooms. Figure 1 shows that assessment texts follow the quantitative progression of the CCSS staircase of text complexity. SBAC texts have slightly lower Lexiles than PARCC texts, but remember that the sample is small. The texts of both assessments have slightly higher Lexiles than those of core-reading texts (Scott Foresman's *Reading Street* and Harcourt's *Storytown*) but the Lexiles fall within the appropriate grade-band ranges in Appendix A of the CCSS.

Because of the strong relationship between vocabulary and comprehension, profiles of vocabulary in the assessment and instructional texts are also provided (see Figure 2). These profiles convey an oft-overlooked feature of complex texts: core vocabulary accounts for most of the words in texts—typically 90% or more. The only case where the percentage of core vocabulary is less than 90%—SBAC Grades 6–8—involves informational text, but even here the percentage of core vocabulary is high (88%). For all other texts, whether assessment or instructional or grade 3 or 6, the percentage of core vocabulary is fairly consistent (92–94%).

The texts of Grades 3 and 6 of the core-reading program have similar levels of rare vocabulary as the assessment texts. Further, the vocabulary loads in the core-reading programs are similar in Grades 3 and 6. Most students can recognize the majority of words in texts, even by the end of second grade, but many are not automatic enough in recognizing words to sustain comprehension (Cummings et al, 2011). When third graders encounter rare vocabulary at the same rate as sixth graders, it's a challenge to develop automaticity with vocabulary.

Figure 1.
Texts of Sample Assessments and Core Reading Programs Relative to Staircase of Text Complexity
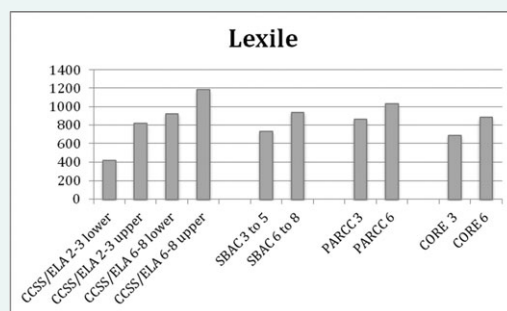
Figure 2.
Distribution of Core and Rare Vocabulary for Texts of Sample Assessments and Core Reading Programs
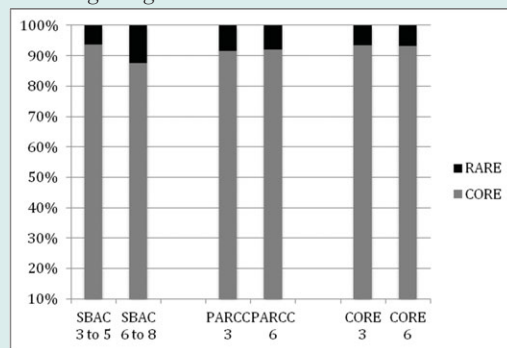
Table 1 gives information on the length of the assessment texts. Grade 3 assessment texts average 575 words, while Grade 6 texts average 675 words. Instructional texts average 950 (Grade 3) and 2,000 words (Grade 6). These differences in length can mean up to 27 more rare words for third graders and 93 more rare words for sixth graders in the average instructional text than in the average assessment text. With fewer total words and fewer rare words, it could even be argued that the assessment texts could present less of a challenge to students than instructional texts.

Tailoring Instruction for Student Success

In most cases core-reading texts are, in fact, complex enough for the majority of students—if not too complex. What are missing in many classrooms are not texts that contain an appropriate amount of complexity but rather the opportunities for students to be responsible for instructional texts and to read enough to be highly facile with these texts. What students aren't learning is how to interact with challenging texts by themselves.

Teachers should not rush to remove all instructional scaffolds but they need to remember that scaffolds are intended to be temporary and steadily withdrawn. Students need to be instructed in how to handle texts on new topics with challenging vocabulary. They need to be instructed in monitoring their comprehension. How can teachers make this happen?

- *Make students responsible for the first reads of texts.* This reading can occur in chunks, with teachers asking a purpose-setting question that requires students to find evidence in text. What is important is that the chunks get bigger over a school year.
- *Ensure that students reread critical parts of texts to demonstrate evidence for their interpretations.* Ask students to read their evidence for the purpose-setting question to a partner, with pairs sharing their evidence.
- *Conduct vocabulary lessons that uncover the critical vocabulary in texts prior to reading.* A short lesson on Ojibwa vocabulary prior to reading *The Birchbark House* or Yorkshire dialect prior to reading *The Secret Garden* illustrates how to support students' success when reading text on their own.
- *Hold explicit conversations with students about the role of challenge in learning.* Teachers need to draw students' attention to their proficiency with the majority of the words in texts and the pace at which new, potentially unknown words are included. The patterns of rare vocabulary in text need to be made visible to students.
- *Help students develop comprehension strategies to use when the task becomes difficult.* The vast majority of students know the majority of words in texts. What students often lack are strategies for proceeding when they encounter unknown words.

Preparing students for contexts where they are responsible for texts is not about test preparation. Ensuring that students are continually increasing their stamina in reading and responding to text is as essential to college and career readiness as ensuring that texts increase in their complexity over students' school careers.

Table 1.
Contexts and Text Complexity of Assessments

	PARCC	SBAC
Number and Length of Sessions	*End-of-Year*: 60 min. x 2 sessions (Gr. 3) to 70 min. x 2 sessions (Grs. 9–11) *Performance*: 40–60 min. per task (Gr. 3) to 50–85 min. per task (Grs. 9–11)	*Computer Adaptive Testing*: 1 hr. 45 min. (Grs. 3–8) to 2 hr. (Grs. 9–11) *Performance*: 105 min. (Grs. 3–8) to 120 min. (Grs. 9–11)
Text Length	200–800 words (Grs. 3–5); 400–1000 words (Grs. 6–8) 500–1500 words (Grs. 9–11)	650 words (Gr. 3) 750 words (Grs. 4–5) 950 words (Grs. 6–8) 1100 words (Grs. 9–11)

Summarized from K. Wixson (April 19, 2013). *Assessment and instruction in the era of the CCSS in English Language Arts.* Presentation given at the Preconvention Institute "Assessment in the Era of the Common Core." San Antonio, TX. Retrieved from: http://www.youtube.com/playlist?list=PLwIychIT3ICgqScYP0LWMspjk1B_VqcGC

References

Cummings, K.D., Kennedy, P.C., Otterstedt, J., Baker, S.K., & Kame'enui, E.J. (2011). *DIBELS Data System: 2010-2011 Percentile Ranks for DIBELS Next Benchmark Assessments* (Technical Report 1101). Eugene, OR: University of Oregon.

Hiebert, E.H., Wilson, K.M. & Trainin, G. (2010). Are students really reading in independent reading contexts? An examination of comprehension-based silent reading rate. In E.H. Hiebert & D. Ray Reutzel (Eds.), *Revisiting Silent Reading: New Directions for Teachers and Researchers.* Newark, DE. IRA.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: Authors.

Elfrieda H. Hiebert *is president and CEO of TextProject, a California nonprofit whose mission is to bring beginning and struggling readers to higher levels of literacy through a variety of strategies and tools, particularly for reading instruction. She is also a member of the Division of Social Sciences at the University of California, Santa Cruz, hiebert@textproject.org.*