
TEXT COMPLEXITY

State of the Art and the Conundrums It Raises

ABSTRACT

Standards for literacy in the twenty-first century raise the bar on the complexity of texts and the tasks for which they are used. The strengths and limitations of contemporary approaches to text complexity are discussed with respect to major points raised in the six articles in this special issue. In addition, four features of text that are of central importance for reading to learn disciplinary content are discussed: topic complexity, genre and function, word-level indicators, and task complexity. We argue that if we are to ascertain sources of complexity and challenge when readers engage with text for purposes of accomplishing interpretive and explanatory levels of understanding, we will need to take into account text, task, and reader situated in sociocultural contexts of schools and communities.

Susan R. Goldman

UNIVERSITY OF ILLINOIS
AT CHICAGO

Carol D. Lee

NORTHWESTERN
UNIVERSITY

As is evident in the articles in this special issue, the Common Core State Standards (CCSS) in English Language Arts (ELA), History, and Science and Technical Subjects (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) call for an increase in the complexity of texts. But it is not only the texts of instruction that are to be more complex. The tasks that students are called on to perform with texts are more complex as well, requiring comparison and contrast across multiple texts and multiple media. Students are expected to engage in comprehension, critique, and production of arguments based on information in texts using discipline-appropriate criteria and practices using the discourse genres and practices typical of each discipline (Gold-

man, 2012). In effect, the task standards attempt to capture developmentally appropriate forms of the literacy practices engaged in by specialists in the various disciplines. These disciplinary practices involve thinking and reasoning processes that have typically been beyond the purview of traditional definitions of comprehension as assessed by multiple choice, cloze techniques, or even extended-response essays. In effect, then, the CCSS raises the bar on two central aspects of the instructional system—texts and tasks—to better prepare students for success in college and career.

Despite the fact that the CCSS raises the bar both for texts and tasks, the lion's share of attention to challenges posed by the CCSS has been devoted to texts and methods for determining the complexity of texts. The emphasis on text complexity blurs some important distinctions, such as comprehension complexity as compared to text complexity (Valencia, Wixson, & Pearson, 2014, in this issue) and text difficulty as compared to text complexity (Cunningham & Mesmer, 2014, in this issue). These are important distinctions to maintain because they imply very different sources of complexity for learners and, concomitantly, different emphases in the design of instruction by teachers. Several of the articles in this special issue raise the task \times text issue explicitly (Cunningham & Mesmer, 2014, in this issue; Pearson & Hiebert, 2014, in this issue; Valencia et al., 2014, in this issue), but, as is the case with much of the interpretation related to the CCSS, these critical distinctions are obscured when the perspective on assessments of complexity is text-centric, without consideration of reader and task.

This commentary considers the strengths and limitations of the extant approaches by highlighting major points raised in the six articles in this special issue. We then discuss particular limitations of text-centric perspectives on complexity that have become salient in our work in Project READI, an effort funded by the Institute of Education Sciences based on reading and argumentation using multiple texts in the disciplines in middle and high school. In that work, we have been wrestling with how to take into account text, task, and reader situated in sociocultural contexts of schools and communities to ascertain sources of complexity and challenge when readers engage with text for purposes of accomplishing interpretive and explanatory levels of understanding.

Assessing Text Complexity: A Long-Term Conundrum

It seems fairly obvious that we want to provide readers with texts appropriate to their knowledge and skill levels. By analogy to Vygotsky's notions of the zone of proximal development, some have referred to this as the zone of proximal text difficulty (Wolfe et al., 1998). That is, texts should challenge readers sufficiently that they are pushed to improve their existing knowledge and skills for reading, comprehending, and learning from texts “just beyond” their current levels—but not too far beyond. Texts that lie outside the zone will lead to little enhancement of comprehension or learning. And texts that are too difficult may have negative affective consequences, along with the obvious cognitive consequences.

However, it has turned out to be difficult to effect good matches between readers and texts. Pearson and Hiebert (2014, in this issue) point out that efforts to achieve the “match” began over 100 years ago with qualitative analyses in the form of “rich descriptions” of features of texts that contributed to comprehension difficulty. These approaches have been almost exclusively text-centric, a tradition that continues to

this day: “difficulty” or “readability” are properties of texts. Interestingly, the features that were richly described were sentence length, obscure vocabulary, and rare syntax (Pearson & Hiebert, 2014, in this issue). The practice of focusing on sentence- and word-level features continued as readability indices developed through the 1900s, but the rich description gave way to “counts” of words per sentence, syllables per sentence, and paragraph as well as other linguistic indices that could be quantified. These indices have been applied to texts across the board.

With the increases in computational power over the past 15 years, the text features that can be used to calculate text difficulty have expanded enormously, as discussed in the Graesser et al. (2014) and Sheehan, Kostin, Napolitano, and Flor (2014) articles in this issue and reviewed recently by Nelson, Perfetti, Liben, and Liben (2012). Both the TextEvaluator (Sheehan et al., 2014, in this issue) and the Coh-Metrix-derived TextEasibility systems (Graesser et al., 2014, in this issue) reduce a multitude of features of texts to a smaller number of dimensions of text complexity. These dimensions roughly index complexity sources that impact surface, basic meaning, and some inferential aspects of comprehension and thus align with contemporary models of text comprehension (Goldman, 2004; Graesser & McNamara, 2011; Kintsch, 1998). Each of these quantitative approaches produces ratings that can assist in creating relative text-complexity predictions (Cunningham & Mesmer, 2014, in this issue). They are highly correlated with Lexiles and other more holistic quantitative indices (Nelson et al., 2012), but they enable a more nuanced examination of features of texts than other quantitative text-complexity measures.

Even with statistical and face validity, uncertainties about interpretation remain. For example, relationships among dimensions can be complex and may differ depending on the specific content and topic. Connecting dimensions, or combinations of dimensions, to comprehension is much more complicated to specify than in the case of a single number. In response, Graesser et al. (2014, in this issue) report that they have been exploring ways to combine different dimensions and expand the analyses to include psychological dimensions. These developments show promise for producing more complete characterizations of the properties of texts. Even so, we see two important limitations of even these highly sophisticated automated analyses of text complexity that are particularly relevant to reading to learn disciplinary content: conceptual complexity of text topics and the pragmatic/functional intent of texts and authors.

Turning to qualitative systems for capturing text complexity, Pearson and Hiebert (2014, in this issue) indicate the variety of motivations that led to renewed efforts to develop alternatives to quantitative indices. They do an excellent job of summarizing three classes of approaches and their strengths and weaknesses: text leveling, rubrics plus exemplars, and text maps. The text-complexity rubrics reflect human judgment about where to place particular texts, stated in terms of descriptions and features of the texts. Once more, however, text-complexity judgments are based on the texts, absent considerations of what students are asked to do with them. Thus, rubric-based approaches also locate difficulty in properties of texts.

Valencia et al. (2014, in this issue) are closest to our view of what needs to be considered when examining complexity. They first reinstate the learning goals of the CCSS: the development of deep learning as compared to shallow, fact-based, reproductive learning or facile procedural execution of skills without regard to understanding what or why these skills are useful. In this regard, Valencia et al. make the

important and practically useful distinction between comprehension complexity, text complexity, and text difficulty. They attend to two-thirds of the reader-text-task triangle by proposing that the appropriate unit of analysis for comprehension complexity is a text-task pairing. They provide compelling illustrations of how the complexity of comprehension moves up or down depending on the task (type of question) paired with a particular text.

Conundrums and Challenges in Moving Forward

Among the important takeaways from these articles are (1) the limitations of a text-centric complexity focus for understanding the challenges texts pose and (2) the importance of conceptualizing difficulties in reading comprehension in the intersections among features of texts, the demands of tasks, the resources that readers bring, and the broader sociocultural contexts under which reading comprehension takes place. This kind of broad conceptualization of the spaces in which reading for understanding occurs will be central to efforts to enact the demands for increased text and task complexity called for in the CCSS.

Another critical issue raised in several articles is whether aspects of the CCSS may increase the difficulty of successful implementation, particularly the apparent arbitrariness of the specifics of standards by grade-level bands (Pearson & Hiebert, 2014, in this issue; Valencia et al. 2014, in this issue; Williamson, Fitzgerald, & Stenner, 2014, in this issue). Valencia et al. offer a useful insight in questioning whether the “big nine” ideas as end goals in English language arts ought not be the goals at each grade band; and, if not, there ought to at least be recognition that the bands should be broader (e.g., K–5, 6–9, 10–12).

Williamson et al. (2014, in this issue) call for examining growth over longer stretches of time through growth-curve modeling. In conjunction with findings reported in several articles in this issue and in Nelson et al.’s (2012) study of accelerations in growth in the primary grades and deceleration of growth in middle and high school, Williamson et al. suggest that the time frames for significant shifts in text and task complexity may take longer than the current skill ladders in the CCSS. Importantly, Williamson et al. point out that, even with time-frame adjustments, text-complexity exposure in absence of changes in instructional practices “is likely to be insufficient to the task at hand” (Williamson et al., 2014, in this issue).

In the remaining sections of this commentary, we raise salient issues related to four features of text that are briefly addressed or alluded to in the articles that our work indicates are of central importance for reading to learn disciplinary content during the middle and high school years: topic complexity, genre and function, word-level indicators, and task complexity. We draw on our experiences in Project READI to research and develop instructional interventions that address many of the teaching and learning challenges brought to the fore by the CCSS and the college and career demands of the twenty-first century. Specifically, in the Project READI work, the ways that texts and tasks become increasingly discipline-centered as students move into middle and high school are a central concern. Further, as we collaborate with teachers, we are continuously confronted with the need for analytical tools that address the problems of practice that teachers face on a day-to-day basis as they attempt to meet the instructional needs of diverse students.

Topic Complexity

Graesser et al. (2014, in this issue) and Valencia et al. (2014, in this issue) raise the issue of topic complexity. Graesser et al. argue that various linguistic features are already surrogates for complexity of the content and that there are ways to computationally derive topic-complexity indices that will improve upon these surrogates. For example, considerate texts may be more likely to use simpler syntax when communicating complex topic information. However, especially with technical reports in science and primary-source documents in history (e.g., exemplary texts for middle and high school in Appendix B of the CCSS), this may not be the case and may even increase the comprehension challenges by not making relationships between concepts explicit.

In addition, any given topic can be understood at different levels of complexity, as is evident in an example of the treatment of the topic of structure and function of seeds in textbooks at primary, middle, and high school levels (Advisory Council on Advancing Adolescent Literacy, 2008). These levels of complexity could be measured by some of the features of Coh-Metrix (e.g., syntactic complexity, markers of cohesion, etc.) or by NAEP's use of text maps (National Assessment Governing Board, 2010). However, teachers hoping to have measures that would shed light on sources of difficulty for their students would need to go to multiple tools to figure out how topic complexity and syntactic complexity, cohesive markers, and vocabulary were represented in a given text. Coh-Metrix can provide such insights, but its outputs are not yet available in a teacher-friendly format. In addition, besides conceptual and syntactic complexity, complex texts can include graphic displays and data tables that are essential sources of information, the complexity of which is not addressed by any of the existing readability measures, either as stand-alone texts or embedded within traditional verbal texts.

A related issue about topic complexity and its influence on readers' comprehension is the analytic grain size that makes information useful, especially for teachers on the ground. We wonder about how detailed the categories for topics need to be in such tools to be practically useful. The grain-size issue is probably most clear and easy to address for topics treated in informational texts. In disciplines like history, mathematics, and the sciences, we are more optimistic that it will not be that difficult to assess differences in topic complexity with automated tools (as discussed by Pearson & Hiebert, 2014, in this issue).

However, when it comes to topic complexity in literature, another set of ontological considerations arise. The counterpart to topic in literature would likely be theme. There are archetypal themes that recur over time, across different national literary traditions, and even across developmental trajectories for children, adolescents, and adults (e.g., love, justice, courageous action, coming of age, etc.). What criteria should be used to determine levels of complexity with such themes? One swath could be through the lens of extant literature on life-course development and moral development (e.g., Kohlberg, 1981). Certainly, there are topics and levels of complexity in understanding topics that will distinguish what a 5-year-old can understand from what a 15-year-old can understand. At the same time, it is important to view this issue through a cultural lens in the sense that, for example, children and adolescents in different cultural communities (e.g., ethnic, national, religious, etc.) can well bring different repertoires for wrestling with varying kinds of thematic complexities (Spem-

cer et al., 2006). For example, Steinbeck's *The Grapes of Wrath* has a reading level of about fifth grade. Many would argue that the themes of the novel are ones not likely appropriate for fifth graders. However, it is possible that the fifth-grade children of migrant worker families might well bring insights into the novel that their middle-class urban or suburban counterparts would not (Gutierrez & Rogoff, 2003; Lee, 2007).

Finally, related to topic complexity is the genre of text used to convey the information and the attendant features of these texts, including the pragmatic functions of texts and the communicative intent of authors (e.g., to communicate given information, to communicate partial information about a model, to critique propositions about a concept, etc.). As we discuss below, pragmatic functions of genres differ among disciplines and affect both topic and the interpretive challenges of texts within the disciplines.


Genre and Function

Although genre is a text feature addressed across many of the articles in this special issue, every tool referenced in the articles takes a big-grain size view of genre, differentiating among them at what boils down to two categories: narrative versus informational, with some distinguishing science from other disciplinary texts (e.g., Graesser & McNamara, 2011). These broad conceptions of genre are reflected in some multidimensional quantitative and qualitative tools as described in Valencia et al. (2014, in this issue). However, the utility of these tools for teachers is questionable, given the grain size.

The NAEP text maps (National Assessment Governing Board, 2010), although not without their problems, are one promising approach to useful tools for teachers. NAEP's use of text maps for literary or narrative texts is based on basic story grammar elements (theme, plot structure, setting, character, and a broad bucket called "author's craft"). On the positive side, the literary text matrices for fiction and literary nonfiction are nicely detailed and serve as guides for experts to select texts at the NAEP grade levels (grades 4, 8, and 12) along with at least traditional, empirically supported quantitative readability formulas. As well, the NAEP matrices indicate that particular features and genres identified for one grade level will be tested at more complex levels in later grades. On the negative side, there are two fundamental problems with these narrative text maps serving as effective tools for teachers. First, there is no discussion of what aspect(s) of any of these features become more complex and in what ways for literary texts. Second, there are no explanations offered as to why one set of text features would be tested at one grade level versus another. Note that there is some attention to such issues in documents teachers use to prepare students for advanced-placement exams in literature. However, even these documents do not address any developmental issues with regard to this knowledge. This is a similar challenge with the CCSS.

For informational text, the NAEP 2011 Reading Framework is better able to capture how text complexity for argument or persuasive texts increases across grades in sources and genres. For example, informational trade books that argue a position are pegged at grade 4, editorials and letters to editors at grade 8, and political and social commentary essays and historical accounts at grade 12. To be useful for teachers,

quantitative and qualitative tools for measuring text complexity will need to include at least this level of specificity about genre distinctions.

One aspect of genre that is almost completely ignored in taxonomies of genres is **are**  tions that they play within a disciplinary context (but see Goldman & Bisanz, 2002). That is, in disciplinary communities of practice of, for example, scientists, there are developed norms regarding the purposes or functions of different disciplinary genres (e.g., review paper vs. empirical report in science). The implications for comprehension of the functions of genre within a discipline, and associated authorial intent in writing particular genres, are not discussed within the NAEP framework nor in the CCSS. This is especially so for literary texts. In literature, knowledge of literary genres—at a grain size much finer than broad distinctions between stories, plays, and poetry—serves as a useful heuristic for readers to make predictions about the kinds of people they will meet, the kinds of actions that are likely to take place, and the kinds of pragmatic functions that the text is likely to engage (Lee, 2011; Smith & Hillocks, 1988). In many respects, specific features of specialized genres in literature serve functions comparable to indicators of text structures in expository texts (e.g., comparison-contrast, cause-effect, etc.).

In history, genre is important for comprehension particularly as it relates to sourcing, the process by which historians examine information about the text such as who wrote it, for what purpose, and when in relation to the event that is the topic of the text (Wineburg, 2001). Author intent is critical to interpreting historical documents and introduces sources of complexity not derivable from features of the text per se. The general structure of genres in history—for example, autobiographies, genres of legal documents such as constitutions, editorials, political speeches, political memoranda, and personal letters—can aid the reader in anticipating the functions and intents of the author and the text. The same is true in science, where mapping across alternative forms of representing the “same” information is an important sense-making practice (Shanahan & Shanahan, 2008).

The general point is that genre plays a critical role in disciplinary content learning. However, for genre to serve useful functions, especially for teachers making decisions about the selection and sequencing of texts, existing tools, including the range of tools described in the articles in this volume, need improvement. In their current state, they do not specify genre at a sufficient level of detail, describe the specialized ways that genres function within disciplines, or discuss how these understandings of genres can serve as useful heuristics that aid readers’ comprehension.

Word- and Sentence-Level Indicators

A third area of discussion across the articles focuses on word- and sentence-level features of texts. Graesser et al.’s (2014, in this issue) discussion of uses of Pennebaker’s Linguistic Inquiry and Word Count (LIWC: Pennebaker, Booth, & Francis, 2007) in assessing word concreteness as an index of text complexity raised some interesting questions for us. Again, we use reasoning about literary texts as the odd-man-out discipline in these discussions. Word concreteness in literary texts will often serve very different functions than it serves in expository or informational texts (Lee, 2011). To illustrate, we entered Alice Walker’s short story “The Flowers” into the LIWC site. The results present dimensions of words on social worlds, positive emotions, negative emotions, overall cognitive words, and big words (bigger than six

letters) in a chart that compares the data from this story with averages for personal texts and formal texts. Overall, it is not clear to us what these distinctions tell us. However, of interest is the data on positive and negative emotions. First, it shows more positive emotion words than negative: positive = 2.47, negative = 0.53. However, it is precisely the shift in the last two paragraphs of the story from positive to negative emotion words that signals the theme. Using Coh-Metrix, the words are identified as easy (learned at younger ages) and, in terms of concreteness and imageability, are not considered difficult. However, the significance of Walker's word choices (e.g., concrete and simple) is not so much in their individual meanings, but rather in the cumulative patterns across words. The reader is expected to infer ~~from~~ across patterns of similarity and dissonance of word choices a thematic abstraction regarding what the story may be trying to convey, both about the protagonist, intentionally named Myop, and for others in the real world who did and do have similar experiences. These patterns are conveyed through short and simple sentences that are not the source of complexity in the story. Our point is that what makes "The Flowers" a complex literary text is not captured by existing quantitative measures of text complexity. Also, available criteria for making qualitative judgments typically do not sufficiently draw attention to the sources of complexity in texts like this one. There are many such texts in ~~current~~ literature—with simple words that belie complex themes.

Task Complexity

Finally, the issue of task complexity arose across several of the articles as part of a broader triad of text, reader, and task complexity. It is a crucial element of the interactive triad because, as Valencia et al. (2014, in this issue) note, the challenges of demonstrating comprehension can shift with the same text depending on the demands of the task (Goldman, 1997; Goldman & Wiley, 2011). Our concern is the need to elaborate further the features of tasks that make them complex so that we can understand how those demands will shift depending on the features of texts and the resources of readers (Lee, 2007, 2011). Additionally neither the articles (nor have we in this commentary) addressed sufficiently the role of the context in which acts of comprehension take place. This issue of social context is essential because we know that students can show one level of text comprehension in the context of schooling and quite different levels of comprehension as they engage in reading activities outside of school (Fisher, 2004).

Clearly, attending to multiple texts raises the complexity of tasks. In Project READI, where, as in the CCSS, the emphasis is on argumentation by drawing critically from across multiple texts, we are grappling with how such multiple-text comprehension tasks differ by discipline. For example, prototypical tasks in science may involve inferring a model of a scientific construct or process from several texts, each of which offers partial information of the whole. In history, a multiple-text analytic task often requires inferring contributing factors to explain a historical event. When these multiple texts convey conflicting explanations, readers often invoke criteria for making judgments, including credibility and author bias/vested interests.

Literature again offers interesting conundrums that can differ substantively from the comprehension demands in other disciplines. In many ways, the substance of task complexity in literary analysis may not be affected by the number of texts being

analyzed. Rather, it is the structure of the reasoning about problems such as using criteria to make judgments about characters and/or themes and analyzing how rhetorical devices employed by authors help to shape interpretations that are the substance of the intellectual work (Hillocks & Ludlow, 1984). Then, comparing across texts poses rhetorical demands about how to communicate the reader's perceptions of points of similarity or difference. The Hillocks taxonomy of interpretive tasks in literature provides an informative example of an approach to task complexity that may be useful in other disciplinary areas like science and history. It begins with key details that are stated literally to inferences that require pulling together relevant details from across the text to generalizations about themes and, arguably the most challenging, explanations of how choices made by authors influence what readers take from the text. It would be very useful if comparable progressions could be developed for history and science genres.

Conclusions

In closing, despite or maybe because of the challenges of the CCSS, there is renewed attention to questions around what makes a text complex. Such a question might actually be just the Trojan horse needed to unravel a complicated set of tacit assumptions about what we want students to know and be able to do as they attempt to gain knowledge through reading and why. We base this claim on two generalizations we draw from the articles in this special issue. First, despite advances in automated quantitative text metrics, these indices are only a starting point for practitioners. Text selection must also take into account the match or mismatch between what students bring to particular texts and what comprehension of those texts requires in the way of knowledge of the conventions of text structure, disciplinary content, and disciplinary-inquiry practices. Second, the tasks that are paired with particular texts jointly define the comprehension challenge, so the work of text selection must be situated in the demands of the task(s) for which texts are being selected and vice versa. The outcomes of such text \times task analyses include a far better understanding of what we are asking students to know and be able to do, thereby setting the stage for asking why we want them to know and do these things. That is, are these instructional outcomes what we intend? Do they prepare our students for the life challenges they face now and into the future (the CCSS are, after all, college and career readiness standards)? Such questions enable an ongoing process of refining and reforming our instructional designs all the while keeping student learning outcomes front and center.

We suspect that in and of themselves these suggestions seem noncontroversial. Perhaps more controversial is how to make the time for practitioners to develop the knowledge and skills needed to engage in such text \times task analyses and develop instructional sequences that will enable students to develop the knowledge, skills, and stamina to persist with the challenging texts and tasks called for by the CCSS. Our collaborations with teachers on the design of Project READI interventions suggest that teacher learning trajectories for implementing instructional practices that realize the CCSS, including text \times task selection and sequencing, occur over cycles of classroom implementation and opportunities to deepen their own understanding of core disciplinary content and practices. This learning does not take place overnight. From a policy perspective, we close by stressing the need to provide adequate oppor-

tunities and time for teacher learning in the evaluation of student learning under the CCSS.

Note

We would like to acknowledge our collaborators on the Project READI Literature Team for their contributions to our thinking about the comprehension demands of literature: Stephen Briner, Candice Burkett, MariAnne George, Sarah Levine, Joseph Magliano, Kathryn McCarthy, and Teresa Sosa. Project READI is funded by the Institute of Education Sciences, U.S. Department of Education, through grant R305F100007 to the University of Illinois at Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.



References

- Advisory Council on Advancing Adolescent Literacy. (2008). *Adolescent literacy: Time to act*. New York: Carnegie Corporation of New York.
- Cunningham, J. W., & Mesmer, H. A. (2014). Quantitative measurement of text difficulty: What's the use? *Elementary School Journal*, *115*(2).
- Fisher, M. T. (2004). "The song is unfinished": The new literate and the literary and their institutions. *Written Communication*, *21*, 290–312.
- Goldman, S. R. (1997). Learning from text: Reflections on the past and suggestions for the future. *Discourse Process*, *23*, 357–398.
- Goldman, S. R. (2004). Cognitive aspects of constructing meaning through and across multiple texts. In N. Shuart-Ferris & D. M. Bloome (Eds.), *Uses of intertextuality in classroom and educational research* (pp. 313–347). Greenwich, CT: Information Age.
- Goldman, S. R. (2012). Adolescent literacy: Learning and understanding content. *Future of Children*, *22*(2), 89–116.
- Goldman, S. R., & Bisanz, G. (2002). Toward a functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 19–50). Mahwah, NJ: Erlbaum.
- Goldman, S. R., & Wiley, J. (2011). Discourse analysis: Written text. In N. K. Duke & M. H. Mallette (Eds.), *Literary research methodologies* (2nd ed., pp. 104–134). New York: Guilford.
- Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science*, *3*, 371–398.
- Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *Elementary School Journal*, *115*(2).
- Gutierrez, K., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice? *Educational Researcher*, *32*, 19–25.
- Hillocks, G., & Ludlow, L. (1984). A taxonomy of skills in reading and interpreting fiction. *American Educational Research Journal*, *21*, 7–24.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kohlberg, L. (1981). *Essays on moral development: Vol. 1. The philosophy of moral development*. San Francisco: Harper & Row.
- Lee, C. D. (2007). *Culture, literacy and learning: Taking bloom in the midst of the whirlwind*. New York: Teachers College Press.
- Lee, C. (2011). Education and the study of literature. *Scientific Study of Literature*, *1*, 49–58.
- National Assessment Governing Board. (2010). *Reading framework for the 2011 National Assessment of Educational Progress*. Washington, DC: U.S. Department of Education.
- 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 with Appendices A–C*. Washington, DC: Authors.

- 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: Student Achievement Partners.
- Pearson, P. D., & Hiebert, E. H. (2014). The state of the field: Qualitative analyses of text complexity. *Elementary School Journal*, **115**(2).
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *LIWC2007: Linguistic inquiry and word count*. Austin, TX: LIWC.net.
- Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, **78**, 40–59.
- Sheehan, K. M., Kostin I., Napolitano, D., & Flor, M. (2014). The *TextEvaluator* tool: Helping teachers and test developers select texts for use in instruction and assessment. *Elementary School Journal*, **115**(2).
- Smith, M., & Hillocks, G. (1988, October). Sensible sequencing: Developing knowledge about literature text by text. *English Journal*, 44–49.
- Spencer, M. B., Harpalani, V., Cassidy, E., Jacobs, C., Donde, S., & Goss, T. N. (2006). Understanding vulnerability and resilience from a normative development perspective: Implications for racially and ethnically diverse youth. In D. Chicchetti & E. Cohen (Eds.), *Handbook of developmental psychopathology* (pp. 627–672). Hoboken, NJ: Wiley.
- Valencia, S., Wixson, K. K., & Pearson, P. D. (2014). Putting text complexity in context: Refocusing on comprehension of complex text. *Elementary School Journal*, **115**(2).
- Williamson, G. L., Fitzgerald, J., & Stenner, J. (2014). Student reading growth illuminates the Common Core text-complexity standard: Raising both bars. *Elementary School Journal*, **115**(2).
- Wineburg, S. (2001). *Historical thinking and other unusual acts: Charting the future of teaching the past*. Philadelphia: Temple University Press.
- Wolfe, M. B. W., Schreiner, M. E., Rehder, B., Laham, D., Kintsch, W., & Landauer, T. K. (1998). Learning from text: Matching readers and text by Latent Semantic Analysis. *Discourse Processes*, **25**, 309–336.

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

1