

Text Complexity: Primary Teachers' Views

Jill Fitzgerald

MetaMetrics® and The University of North Carolina at Chapel Hill Emerita
and Research Professor

Elfrieda H. Hiebert

TextProject and The University of California at Santa Cruz

Kimberly Bowen

MetaMetrics®

E. Jackie Relyea-Kim and Melody Kung

The University of North Carolina at Chapel Hill

Jeff Elmore

MetaMetrics®

May 12, 2014

Corresponding author: Jill Fitzgerald, jfitzger@email.unc.edu

RUNNING HEAD: Primary Teachers' Views of Text Complexity

**This is an Accepted Manuscript of an article published by Taylor and Francis in
Literacy Research and Instruction, 2015, volume 54, pp. 19-44, available through library
subscription or online at <http://www.tandfonline.com/10.1080/19388071.2014.954086>.
(The latter number in the address is the DOI.) For citation/quotation, please use the
published version.**

Text Complexity: Primary Teachers' Views

The Common Core State Standards (CCSS) have put text complexity in the spotlight to a degree not encountered in prior standards. Although linguists have been interested in text complexity for some time (e.g., Merlini Barberesi, 2003), complexity of text was not highlighted in U.S. schools until recently. While many previous standards documents emphasized students' ability to answer increasingly more challenging *questions* about the texts they read, the CCSS also include a separate Standard—Anchor Standard for Reading 10—that focuses on all students' ability to read increasingly more challenging texts over their school careers (National Governors Association [NGA] Center for Best Practices & Council of Chief State School Officers [CCSSO], 2010a). Controversy about the text-complexity standard is heating up, and teachers are investing time and energy into CCSS implementation (Shanahan, 2011; Williamson, Fitzgerald, & Stenner, 2012). Even primary grade teachers (kindergarten through second grade teachers) are considering how to support their children's advancement through progressively challenging texts. Although the Standards require *all* children in kindergarten through twelfth grade to read more complex texts (NGA & CCSSO, 2010b, Appendix A, p. 2), the Standards do not specifically address text-complexity *factors* for early grades texts. However, primary teachers (kindergarten through second grade teachers) will need to consider text characteristics in relation to text complexity when selecting or recommending texts for young students (in kindergarten through second grade). The extent to which teacher knowledge about early grades text complexity will fill the void left by the CCSS is unknown. Better understanding of teachers' views could be an important first step to enhancing everyday implementation of the standard in primary grade classrooms.

Consequently, the purpose of the present study was to explore primary teachers'

outlooks on text complexity. The research question was “What text characteristics do primary teachers think are important for early grades text complexity?” Teachers accomplished a two-part task. First, to stimulate teachers’ thinking about important text characteristics, we asked primary teachers from across the United States to accomplish an online paired text comparison task and to think about which text characteristics mattered most for their decisions. Next, teachers completed an online questionnaire focused on revealing the text characteristics teachers thought mattered most for early grades text complexity.

The CCSS Text Complexity Standard and Implications for Primary Grade Children

First, the CCSS authors rationalize the significance of requiring students to read increasingly complex texts *throughout schooling* through six claims: (a) An ACT (2006) report claimed that higher-performing college students were differentiated from lower-performing students in their ability to answer questions associated with *complex* texts. The ACT (2006) report was perhaps among the first in the U.S. to highlight the complexity of texts. (b) College texts’ difficulty levels increased in the past five decades, science journals’ word difficulty increased between 1930 and 1990, and modern day workplace reading material far exceeds twelfth grade level (e.g. Adams, 2009). (c) Kindergarten through twelfth grade texts’ challenge levels decreased in the past six decades (e.g., Hayes, Wolfer, & Wolfe, 1996). (d) High school students are generally not held accountable for reading complex text independently, but college students are (e.g., Heller & Greenleaf, 2007). (e) Expository texts are not read much in kindergarten through twelfth grade, but such texts account for the majority of college and workplace reading (e.g., Moss & Newton, 2002). (f) College students must be able to read well in order to make sufficient academic progress (Wirt, Choy, Rooney, Provasnick, Sen, & Tobin, 2004).

Second, the CCSS authors reference both quantitative and qualitative characteristics

of text complexity as well as the importance of considering text complexity in relation to the particular students who are reading a text and the task required of the students. Quantitative text-complexity ranges are delineated in two staircases for each of six grade bands—K-1, 2-3, 4-5, 6-8, 9-10, 11-College/Career-Ready). However, no text complexity ranges are provided by the CCSS for the K-1 band. One staircase shows “old” ranges of text complexity, and the other shows aspirational bands, with the latter reaching higher than “old” peaks. Significantly, by the end of grades 3, 5, 8, 10, and 12, students must read at the high end of the “new” aspirational complexity band, “independently and proficiently” (NGA & CCSSO, 2010b, Appendix, p. 10).

Although no CCSS aspirational text-complexity levels are provided for kindergarten and first grade, the end of third grade shift has serious implications for earlier grades. The “old” upper boundary for third grade was 725 Lexiles, the 75th percentile of current-day third-grade texts found in one study (Koons, 2011). (A Lexile is a unit of measure, with one Lexile equal to 1/1000th of the difference between the mean difficulty of mid-first-grade texts [200L] and *USA Today* [1200L].) The re-established aspirational boundary is 820 Lexiles (L) (NGA & CCSSO, 2012). Notably, the 95L increase is not insignificant as it rises to a similar level as the upper boundary (845L) of the former band for fourth/fifth grade. Moreover, when considered in relation to student reading abilities, the end of third grade bottom two quartiles of students’ historical estimated average reading levels in one study were 415L and 631L, respectively (Williamson, Fitzgerald, & Stenner, in press). Clearly, many average and lowest-performing end-of-third-grade students will likely have considerable difficulty reading texts that have complexity levels of 820L. Most importantly in relation to the present study, to achieve the heightened level of text complexity at the end of third grade, it seems probable that kindergarten through second grade students would need to read more complex texts as well.

The CCSS authors also propose four qualitative indicators of text complexity: levels of meaning or purpose, structure, language conventionality and clarity, and knowledge demands. However, the extent to which the four indicators are considered applicable to the earliest grade levels remains unclear. Exemplar texts are given for the earliest grade levels, but in an accompanying footnote, the CCSS authors state:

Children at the kindergarten and grade 1 levels should be expected to read texts independently that have been specifically written to correlate to their reading level and their word knowledge. Many of the titles listed . . . are meant to supplement carefully structured independent reading with books to read along with a teacher or that are read aloud to students to build knowledge and cultivate a joy in reading (NGA & CCSSO, 2010a, p. 32).

Because kindergarten and first grade students' reading levels vary considerably, the statement offers substantial latitude when considering the complexity of text appropriate for beginning readers. The latitude places weighty demands on the expertise of primary teachers.

Theoretically, What Text Characteristics Might Matter Most for Text Complexity?

Just as the Common Core authors considered text complexity as situated in the relation between particular readers, texts, and tasks, a founding principle for the present study was that early grades text complexity is relational to the students who are reading them (cf. the RAND Reading Study Group relational model of reading, Snow, 2002). Early grades texts are generally written to heighten certain factors related to young children's processing ease as they begin to learn to read (cf. "optimal" texts and "optimality theory" for the fundamental assumption that an "optimal" text is one in which text characteristics are configured such that readers can create meaning while reading with the most ease and the greatest depth of processing, Merlini Bararesi, 2003; also see Juola, 2003 who discusses the necessity of complex systems to reflect processes, including cognitive processes of "users").

Importantly, to our knowledge, early grades text complexity per se has not been a topic of research. That is, while many text characteristics have been investigated individually or in clusters in relation to many young students' performance factors, very little is known about the collective characteristics that matter most for determining text complexity levels. Only recently has an attempt been made to even suggest a theoretical set of characteristics that ought to be considered collectively (Mesmer, Cunningham, & Hiebert, 2012). Text characteristics that are prominent in the Mesmer and colleagues' exposition include multiple linguistic levels: subword, word, sentence, and discourse level characteristics. The following summary draws on, and adds to, Mesmer and colleagues' (2012) review.

A deep research base suggests that, although meaning creation is at the heart of learning to read, "cracking the code" requires focal effort for beginning readers. Critical cognitive factors inherent in the early learning to read phase of development are phonological awareness and word recognition (e.g., Adams, 1990; Clay, 2001; Fitzgerald & Shanahan, 2000; Storch & Whitehurst, 2002). Consequently, hypothetical critical text characteristics that would support development of phonological awareness and word reading are texts that include: repetition of simple words (to facilitate sight word development and orthographic pattern knowledge (e.g., Metsala, 1999; Vadasy, Sanders, & Peyton, 2005; and cf. Howes & Solomon, 1951, where children's accuracy and speed of recognition were influenced by word frequency); words with relatively simple orthographic configurations to encourage orthographic pattern knowledge (e.g., Bowers & Wolf, 1993); rhyming words to support aspects of phonological awareness (e.g., Adams, 1990); words with meanings that are familiar to young children in oral language—such words likely reduce challenges to meaning creation while reading, permitting more attention to word recognition (e.g., Muter, Hulme, Snowling, & Stevenson, 2004); word imageability, concreteness, and age-of-word acquisition

assist comprehension and/or word recognition (e.g., Woolams, 2005); repeated phrases which may reinforce phonological awareness and sight word development as well as varied word recognition strategies such as guessing from context (e.g., Ehri & McCormick, 1998); and inclusion of cohesive ties (e.g., referential cohesion when a noun, pronoun, or noun phrase references another text element) is related to reading time and comprehension (e.g., McNamara & Kintsch, 1996).

Among the many and varied text characteristics that are potentially important for early grades text complexity, in general, only a modest number of them have been included in text complexity measurement systems. Nearly all of the most commonly used measurement systems focus on word frequency, word length, and/or sentence length (generally considered a proxy for syntactical complexity) (e.g., Klare, 1974-1975; the Lexile Framework for Reading [Stenner, Burdick, Sanford, & Burdick, 2006]; Degrees of Reading Power [Koslin, Zeno, & Koslin, 1987]; SourceRater [Sheehan, Kostin, Futagi, & Flor, 2010]; and Coh-Metrix [Graesser, McNamara, & Kulikowich, 2011]). The systems depend on word frequency, word length, and/or sentence length because those characteristics historically have been shown to be strong predictors of text complexity. It is worth noting that there is a difference between predictive capacity and explanatory capacity however. Some systems include additional text-characteristics, but notably, none of the commonly occurring measurement systems specifically provides explanation of what constitute *early grades* text complexity (cf. Graesser, et al., 2011 and van der Sluis & van den Broek, 2010).

Understanding Primary Teachers' Views of Text Complexity:

Why It's Important

Common Core State Standards guidance on text complexity for beginning readers is elusive, and criteria for evaluating text complexity for more mature readers may not always be applicable for beginning readers. However, it is possible that many primary teachers

have developed their own criteria for evaluating the texts of beginning reading.

Understanding teachers' perspectives about which text characteristics are important for early grades text complexity could have wider implications for how primary practitioners teach to the text-complexity Standard. In the midst of changes in practices and policies, teachers' voices are rarely formally heard and shared. Yet teachers' opinions matter significantly, in part because it has been well established that their perspectives weigh heavily during classroom implementation of reforms in general (Greenfield, Rinaldi, Proctor, & Cardarelli, 2010). The impact of their perspectives is no less likely as they guide their students to meet the goals of the text complexity Standard.

Further, reading is an interactional event, especially during its instruction, and one of the primary agents in that instruction is the teacher. For beginning readers in classroom settings, teachers are often a direct part of the reading event, as they typically choose texts, and they also focus on particular aspects of texts during instruction with their students. We hypothesized that, in an area where guidance about text complexity has been lacking, teachers have acquired substantial amounts of knowledge. Within our profession and also as we interact with more extended communities, it is important to recognize teachers' expertise.

Methods

Design

Participants were 90 primary teachers. Teachers accomplished a two part activity, making judgments about text complexity in pairs of texts and responding to a questionnaire about which text characteristics are important for text complexity. The first task was designed solely to stimulate the teachers' thinking about text characteristics that matter most for text complexity. The teachers completed an online 30-minute paired text comparison task during which they saw two texts side by side on the computer screen. The teachers

were asked to decide which text in the pair was more complex than the other. While doing the paired text comparison task, they were asked to think about which text characteristics they were using to make their decisions. After the paired text comparison activity, they completed (online) a questionnaire to capture the teachers' thoughts about important text characteristics (vis-à-vis the main purpose of the study), reporting which text characteristics they thought they used during the paired-text-comparison activity, and they rated their confidence in determining text complexity. They also provided additional contextual information about selected classroom instruction habits as well as demographic information. The main analysis was examination of the extent to which teachers indicated they used various text characteristics to determine text complexity. The study was conducted soon after the Standards were released.

Participants

The participants constituted a voluntary convenience sample (cf. Fink, 2003 for convenience sampling). A message about the project was sent on a pre-existing U.S. email list that reached educators, including primary teachers, across the country. The email list was administered by MetaMetrics in Durham, NC. Of all the teachers on the email list, two hundred fifty primary teachers expressed an initial interest in participating in the study. They were then given detailed information about the project purpose and task as well as benefits of participating (a set of classroom books). Representing 33 states and 75 school districts, 90 teachers chose to participate. Information on participant characteristics appears in Table 1. On the whole, they taught in urban or suburban schools, represented a wide range of years of teaching experience, had taught students from a broad socioeconomic range (slightly more than half of the teachers came from schools where 50% or more of the students participated in a subsidized lunch program), and had considerable exposure to information about basic beginning reading processes and instructional practices.

Data Sources: The Two Part Task

The paired comparison task was designed solely to provide an immediate context—looking at authentic texts—to stimulate teachers' thinking about which text characteristics they considered when making judgments about text complexity. Different texts could be more or less complex for different reasons. The paired comparison task should have encouraged teachers to consider the fullest range of text characteristics possible.

Text selection and the paired text comparison activity. Using maximum variation purposive selection (Patton, 1990), a set of 350 texts intended for kindergarten through second grade was chosen primarily from an existing larger corpus of early grades texts (MetaMetrics, n.d.). As well, to ensure that some CCSS exemplar texts were well represented, 18 CCSS exemplar books designated for kindergarten, first, and second-grade that were not available in the larger corpus were purchased. The goal of maximum purposive selection was to ensure comprehensive representation of a wide variety of early grades text types, text levels, and publishers. We chose 350 texts, a large enough number to ensure wide representation of types of texts, but also a small enough corpus to be manageable for development and implementation of the paired-text-comparison task. Text examples are provided in Table 2.

To create a text corpus for the paired-comparison task, six categories of commonly occurring early grades texts were determined: code based (books that are intended to be highly decodable and/or that have phonics regularities, or emphasize certain phonics patterns), whole word (books that focus on high frequency words), trade books (texts that appear in school or public libraries), leveled books (texts that have a "level" of difficulty assigned to them, generally other than a grade level), texts of assessments, and other (e.g., label books, which are texts that show many pictures with a word or a few words beside them naming the picture). The first four types were previously defined and identified by Aukerman (1984)

and Hiebert (2011) and considered to be common labels for educators and publishers (cf. Mesmer, 2006). The last two types were included because texts appearing in assessments also commonly occur in early grades, and texts of assessments may be even more prominent during implementation of the CCSS. It should be noted that the six categories are not mutually exclusive. For instance, leveled books may also be trade books. Mutually exclusivity was not required for our study though. The purpose for determining common categories was to ensure representation of a wide variety of commonly used early grade texts. Publisher's designation of text type was used. When publishers used labels not listed above, or where there were no labels, we used a researcher-devised set of rules for classification. For example, if a publisher characterized a text as primarily attending to high-frequency or sight words, it was labeled "whole word." Forty one percent of the texts were categorized as leveled, 17% were code-based, 15% trade books, 10% whole-word, 9% were texts of assessment, and 8% were other. Using a set of definitions modified from Duke (2000), 66% of the texts were labeled narrative, 24% informational, and 10% were labeled hybrid or other. (Inter-rater reliability for genre classification was .96).

Text levels were determined by the researchers using publisher stated grade, level, or age ranges. Approximately 36% were labeled by the researchers as easiest (kindergarten or kindergarten ages), 37% moderately hard (first grade or first-grade ages), and 27% hardest (second grade or second-grade ages). Among the 350 texts, thirty two publishers were represented, ranging from 3 to 15 different publishers for each of five of the six text types, with one publisher for the text of assessment type.

Each entire text, including images, was scanned by computer. A computer scientist created a program for the paired text comparison task and means for the teachers to access it through a web address. When teachers logged on, teachers saw the directions. Teachers were told that they would see two excerpts of several pages of text, side by side, on the screen,

they could scroll through each excerpt, and they should click on a button at the bottom of a text to tell which of the two texts was more complex (see a screenshot in Table 3). Teachers were also directed to think about which text characteristics they were using to make their text complexity decisions, and they were told they would complete a questionnaire after doing the paired-text-comparison task. To avoid biasing the teachers, the directions provided no information about possible text characteristics that the teachers might consider. Pairs were randomly computer generated for teachers “in the moment” from the original 350 texts while they were on the computer, so that each teacher received different sets of pairs. Each teacher saw a total of 125 comparisons involving excerpts from 35 books.—not all 350 texts (though all 350 texts were used for the task). The reduction in number of texts that teachers read was done to limit the amount of time teachers would need to spend on the task. After reviewing a pair of texts, teachers clicked a button at the bottom of the screen to indicate which text they thought was more complex. The task took approximately 30 to 45 minutes. All teachers completed the paired text comparison task.

Online questionnaire. After completing the paired text comparison task, teachers then responded to the online questionnaire shown in Table 4 (Survey Monkey [n.d.] was used for the questionnaire delivery and response.) We designed the questionnaire primarily to explore the text characteristics that teachers thought were most influential for early grade text complexity. That is, the questionnaire responses were our main research interest. We also collected additional contextual information about demographics and selected teachers’ reading instruction habits along with their confidence in determining text complexity for their students and matching text complexity levels to students’ reading levels.

On the questionnaire, to prompt teachers’ consideration of text characteristics, we offered a selection of text characteristics along with an opportunity to add others (see item 5 on Table 4). We did not want to unduly influence the teachers’ thinking, but at the same

time, we thought it important to provide examples while encouraging teachers to offer commentary on additional characteristics. To decide which characteristics to include, we made a list of the characteristics discussed above in the section, “Theoretically, What Text Characteristics Might Matter Most for Text Complexity,” and randomly selected six while insuring that the characteristics that most commonly occur in text complexity measurement systems (high frequency words and sentence length) were included. As former first grade teachers, two of the researchers also thought some teachers might consider pictures as an aid to students’ word recognition, and so we added pictures/illustrations to the list. The “other” category was also added.

Three types of items were offered: (a) a prompt for open response (questions 1 through 4 in Table 4); (b) Likert-type items (questions 5, 7, and 8 in Table 4); and (c) request to rank order entries (question 6 in Table 4).

Results

The participants’ responses to the questionnaire were divided into three clusters, and in the following sections, we provide results by each of three clusters of items: (a) items related to the teachers’ context of reading instruction; (b) items addressing the main research question, “What text characteristics do primary teachers think are important for early grades text complexity”; and (c) results from items indicating teachers’ confidence in determining text complexity and student-text matches. For each cluster, we also first state the statistics and/or measurement methods used to examine the item responses. The statistics were ones typically used to summarize questionnaire results.

The First Item Cluster: The Context of Reading Instruction

The first part of the questionnaire addressed the amount of time teachers devoted to particular activities (e.g., guided reading), the materials used for these instructional activities, and topics of professional development. For each item in the cluster, the percent of teachers

who gave selected responses was calculated. The teachers' responses are depicted in Figure 1. Of three major reading instruction activities—teacher reading aloud, guided reading, independent or individual reading—nearly half of the teachers (47.7%) said that guided reading consumed from 41 to 100% of their reading instruction time (see the last three grids, moving left to right in the top row of Figure 1). Approximately two-thirds (65.6%) of the teachers reported that they devoted up to 40% of reading time to student individual reading. Almost 60% reported that read-alouds consumed up to 20% of instructional time devoted to reading.

Teachers' answers to the next question about the types of texts that they used for the three reading instructional activities are summarized in Figure 2. For guided reading and student independent reading, 77% of teachers reported using leveled readers and 83% reported using decodable books. When reading aloud to children, more than 86% of the teachers said they used non-leveled trade books. Only 6%, or five teachers, of the entire sample identified additional materials for activities. The additional materials included technology and multimedia reading materials, reading articles on the computer, and/or doing PowerPoint presentations with the teacher's own created text.

Next, we were interested in the extent to which teachers reported having learned about various reading instruction topics and whether they used the topics in their own classroom situations. A large number of teachers identified professional learning experiences (see the darker lines in Figure 3) aligned with several of the processes named in the influential *Report of the National Reading Panel* (National Institute of Child Health and Human Development [NICHD], 2000)—from about 78% to 81% for topics related to decoding (78.9%), phonological awareness (77.8%), phonics (80.0%), and fluency (81.1%). The topics encountered the most were guided reading (86.7% of teachers), fluency (81.1%),

and differentiation (81.1%). Least encountered were repeated readings (43.3%) and whole language (40.0%).

For several entries, considerable correspondence was noted between professional development topics and actual classroom implementation. Actual implementation is represented in Figure 3 in gray lines. Specifically, for a primary instructional activity of guided reading [77.8%], for the content of instruction (decoding [74.4%], phonological awareness [73.3%], phonics [73.3%], fluency [78.9%], and comprehension 80.0%), and for differentiation, the percent of teachers who implemented the topic (76.7%) was similar to the percent of teachers exposed to the topic during professional development (81.1%). For the remaining instructional practices and/or philosophies (e.g., whole language), less correspondence between topic exposure and classroom implementation was apparent.

The Second Item Cluster, The Focus of the Study: Teachers' Views of Text

Characteristics That Determine Complexity

The second item cluster addressed the main research question of the study. For the first item in the second cluster, summary statistics were examined for the Likert item response levels. The statistics provided the means (the average extent of text characteristic use) and standard deviations for each response level (extent to which the characteristic was used during the paired text comparison activity). Additionally, the total group of participants was broken into reading/literacy specialists and other educators, and then the summary statistics were compared across groups. The subgroup split was done because, as noted earlier, there was some possibility that reading/literacy specialists, who typically have more knowledge about reading instruction, might respond differently. For the second item in the second cluster, teachers rank ordered the text characteristics they reported using in the previous item. The mean rank for each characteristic was calculated and used to rank the characteristics from most to least used.

In the following, we describe the results about teachers' views of text characteristics, frequently referring to parts of Table 5. The second column in Table 5 shows the average extent of text-characteristic usage reported by the entire group of teachers. The average response for all but one characteristic and the "other" category was around 5, which corresponds to *usually*. The characteristics with the highest rating were word decodability (5.3) and high-frequency words (5.2), between *usually* or *always*. Teachers identified font/text size as the variable that least influenced their ratings (mean of 4.3).

Just over half of the teachers reported using additional text characteristics (55%), but only approximately a third of the teachers (29%) actually wrote added text characteristics. The additional characteristics named by the teachers fell into five categories. Some teachers named two or more categories, totaling 30 comments. The additional characteristic categories mentioned by teachers were: concept or content difficulty relative to children's background knowledge (named by 12 teachers); genre (6 teachers); repeated patterns in text (4 teachers); topic and whether it was of interest to students (3 teachers); and "miscellaneous"—sentence structure/syntax (2 teachers), author popularity (1 teacher), use of quotation marks (1), and format factors (1).

Even for text characteristics with the highest average ratings, some teachers had alternative perspectives. For every text characteristic, at least one teacher responded "Never," and at least one responded "Always." The variability is evident in the standard deviations in Table 5. As might be expected, the most variability as indicated by the largest standard deviation was for "other", and the second most was for font/text size, while word decodability was the characteristic that showed the least amount of variability.

When teachers were asked to rank the text characteristics relative to one another, the characteristics were ranked (on average) as follows (with the most important listed first) (see Table 5 for the mean rank and SD statistics): word decodability, high frequency words,

pictures/illustrations, words meaningful to young children, amount of text per page, sentence length, font/text size, and other. As would be expected, the mean rank order is consistent with the mean extent of usage (from the Likert items) for the different characteristics shown in Table 5.

Reading/literacy teachers or specialists have somewhat different teaching responsibilities than other educators (Bean, Cassidy, Grumet, Shelton, & Wallis, 2002). For example, because they spend the school day teaching reading, reading teachers could have more background knowledge about reading instruction and materials than classroom teachers. For this reason, the mean ratings of reading/literacy teachers and other teachers are separated in the last two columns of Table 5. However, a comparison shows few differences in the text characteristics used for the paired-text-comparison task between the reading/literacy teachers and other teachers. Classroom teachers reported using the same characteristics about as frequently as reading/literacy specialists. The one exception to this pattern was the amount of text per page, which classroom teachers reported using slightly more frequently than reading/literacy teachers. The rank ordering of text characteristics by each of these subgroups was identical.

The relationship between teachers' text complexity level designations and researcher/publisher text complexity level designations. To garner a sense of the match between teachers' text complexity level designations and the easy/moderately hard/hardest researcher text complexity level designations (with the latter based on publisher stated grade, level, or age designation), it was necessary to first aggregate teachers' responses so that each text could be assigned a teacher designated text complexity measure. To accomplish that, through Rasch modeling (Bond & Fox, 2007) using the 90 teachers' paired-comparison decisions, a text complexity logit scale was created. Through the scale development, each of the 350 texts was assigned a teacher text complexity level designation (a logit measure) (a

level that came from the aggregation of the 90 teachers' decisions). The logit measures for text complexity ranged from approximately -13 to approximately +13. Measurement reliability was .99 (using the Separation Index Method [Wright & Stone, 1999]).

The correlation between teachers' text complexity level designations for the 350 texts and the researcher/publisher designations was .79 (Spearman ρ , $p < .001$), indicating that on the whole, the teachers' determinations of more versus less complex texts agreed well with the researcher's easy/moderately hard/harder designations garnered from publishers' grade, level, or age assignments. To provide an additional sense of the correspondence between teacher designations and the researcher/publisher designations, see the last two columns in Table 2.

The relationship is shown in Figure 4. For the texts designated as "easiest" by researcher/publisher assignment, teachers, on the whole, thought they were less complex than the other texts, and likewise, researcher/publisher designations of "harder" texts were judged to be more complex than other texts. That is, as researcher/publisher text complexity designation increased, so did teacher determination of text complexity increase.

The Third Item Cluster: Teachers' Confidence in Determining Text Complexity and Student-Text Matches

For each item in the third cluster, summary statistics were examined for the Likert responses providing the extent of teachers' confidence, on the whole. As well, percent of teachers who reported each response level was examined. When teachers were asked to express confidence in determining text complexity, on the whole, the average of 5.2 ($SD = .9$) suggests a predominance of "somewhat" to "completely agreed" responses. Only 3% of 90 teachers gave ratings of "sometimes" or "never" to the statement of knowing how to establish text complexity.

When teachers were asked to indicate the extent to which they knew their students' reading ability levels and understood their students' reading interests, they were again in

strong agreement. Their average responses ($M = 5.4$, $SD = .9$ and $M = 5.0$, $SD = .8$) for the two questions, respectively, were “somewhat” to “completely agree.” For both questions, over 97% of the teachers indicated they agreed, somewhat agreed, or completely agreed that they knew their students’ reading ability levels and reading interests.

Teachers also agreed that they do a good job of matching students to books that are appropriately challenging, responding on average between “agree” and “somewhat agree” ($M = 4.7$, $SD = .8$). “Matching students to books” was intended to reference helping students to read books that are written at the students’ reading levels and/or that are in students’ interest areas. Over 94% of teachers responded they agreed, somewhat agreed, or completely agreed that they do a good job of matching students to texts.

Conclusion, Discussion, and Implications

The main conclusion was that the primary teachers reported using the following text characteristics when determining text complexity. Word decodability and high-frequency words were used the most. The next most influential characteristics, from most used to least used, were pictures/illustration, extent to which the text included words meaning to young children, sentence length, and the amount of text per page. There were only slight differences in extent of usage for those top six characteristics, all of which hovered around “usually” used when determining text complexity in the paired text comparison task. Finally, font/text size was used “often,” and over half of the teachers also used additional text characteristics “often.” However, considerable variability existed in the additional characteristics that teachers named. The additional characteristics fell into five categories: concept or content difficulty relative to children’s background knowledge, genre, repeated patterns in text, topic and whether it was of interest to students, and miscellaneous characteristics.

Understanding primary teachers' views about text complexity can help bridge the gap of guidance missing from the CCSS regarding how to address the text complexity Standard in the primary grades. In the following sections, we discuss the study results, especially with regard to how the teachers' views can matter for teaching to the text complexity standard, but also in relation to our earlier theorization about text characteristics that matter for text complexity.

As context to interpretation of the findings and their implications, it is important to consider limitations of the study. Despite the diversity in teacher backgrounds, the views reported in the present article come from a specific group of teachers reporting on the practices within their own school settings. Because we did not have a random sample, the degree to which participant views represent those of a national population of primary teachers cannot be established. Possibly, for instance, teachers who volunteer for studies are different from other teachers in important ways, such as in the extent of their confidence in ability to teach reading, or perhaps participants were more interested in obtaining the set of books for their classrooms than other teachers. Additionally, we did not contact non-responders to determine the extent to which non-responders and responders were similar or dissimilar. As well, it is possible that the questionnaire format limited teachers' revelation of text characteristics they thought they used. That is, selected text characteristics were listed, but some were not (e.g., text structure, genre, content), and some teachers may not have taken sufficient time to add to the list when prompted.

Important Contextual Information: Teachers' Recognition of Unique Texts for Beginning Readers

Before we discuss our main findings about teachers' views of key text complexity characteristics, it is also important to discuss and understand implications of contextual information the teachers reported about materials they use for reading instruction. Teachers

reported using a mix of text types, which shows their recognition of the unique nature of the beginning reading period. By using a mix of texts students would be exposed to materials in which varied text-complexity characteristics are emphasized. Approximately three-quarters of the teachers reported using leveled texts for guided reading (77.8%) or for student individual reading (83.3%) and decodable texts for student individual reading (74.4%). By using both, teachers are exposing students to the different word level characteristics that they identified as important. Leveled books tend to provide more opportunities for students to read texts with higher level word meanings, while decodable texts provide more experiences with easier word decodability and high frequency words (Murray, Munger, & Hiebert, in press).

The percentage of teachers reporting the core or basal reading programs was low: approximately a third of the sample. The percentage varies considerably from reports prior to NCLB (Baumann, Hoffman, Duffy-Hester, and Ro, 2000) when the majority of teachers, including primary teachers, reported using textbooks for at least a portion of their reading instruction. While the low percentage in this study may be an artifact of the sample (e.g., teachers who volunteered to participate, or of primary teachers rather than elementary teachers in general), or even an artifact of the way we asked the question, it is also possible that a diminished reliance on core programs or basals translates into more variety in text-complexity exposure for students.

Teachers' reported use of content area textbooks was at the same level as core reading texts. The choice is understandable at the primary levels in that content area textbooks often provide generic descriptions of phenomena in science or events in social studies (Beck & McKeown, 1991). But the pattern does raise the question of amount of attention given to informational texts in primary classrooms, texts that may employ complexity characteristics that are different from decodable or leveled readers or that may have exaggerated presence of

selected characteristics. For example, informational text structures can vary from narrative text structures, and they can have more unfamiliar words (Duke, 2000). Few decodable texts can be described as truly expository or informational in genre. There are leveled texts that attend to content but not necessarily in a systematic manner as would be expected in content area instruction.

**Teachers' Voices Can Help to Define Primary Grades Text Complexity Characteristics:
How Their Views Were Similar to, and Different from, Prior Positions on Text
Complexity**

Primary teachers bring a particular kind of wisdom to understanding text complexity, a wisdom that arises from the real experience of watching young children interact with texts and of guiding and instructing children through beginning phases of learning to read (Fitzgerald, 2000). They could contribute valuable understandings to affect a robust conception of primary-grade text complexity.

The teachers in our study were opinionated and also confident about their ability to determine text complexity for young students. Aspects of their responses resonated well with, and are supported by, our earlier theorization about potential key text characteristics for text complexity as well as the Mesmer and colleagues' (2012) theorization. The teachers in our study emphasized word decodability and word frequency, and to a slightly lesser degree, word meaning—three factors that are prominent in prior literature as related to aspects of young children's reading performance (cf. Adams, 1990; Fitzgerald & Shanahan, 2000). That is, we might infer that teachers thought word structure influences students' ability to pronounce a word and higher word frequency levels could mean students would have greater familiarity with the words. It is also plausible that teachers considered sentence length to be a proxy for syntactical complexity, another previously hypothesized key characteristic related to complexity. It is worth noting that the teachers' use of decodability, frequency, and

sentence length is consonant with the main variables used in many of the current text-complexity measurement systems. Could it be possible that many teachers know which variables are used in measurement systems and that has influenced them? Unfortunately, we have no data to clarify the possibility.

At the same time, the primary teachers' views diverged from our earlier theorization in at least one important way. No teachers mentioned cohesion, and only a few mentioned genre or text structure—two discourse level characteristics. Discourse-level text characteristics may play an important role in early grades text complexity in so far as patterning and repetition, along with greater presence of cohesive ties could interplay with other text characteristics, such as inclusion of somewhat more difficult words to decode or lower frequency words, to lower readers' perceptions of complexity (Merlini Barbaresi, 2003). As another example, some evidence suggests that concreteness/abstractness or imageability interacts with structural complexity and word familiarity to influence readers' word recognition (e.g., Schwanenflugel & Akin, 1994).

The teachers' emphasis on word and sentence level characteristics also diverged from the CCSS's perspective of qualitative indicators of text complexity: levels of meaning or purpose, structure, language conventionality and clarity, and knowledge demands (NGA & CCSSO, 2010a). That is, the CCSS's list of qualitative indicators does not include word and sentence level characteristics. Among the 90 teachers, only 12 teachers mentioned concept or content difficulty, which could signal the CCSS dimension labeled "levels of meaning." As already noted, few emphasized text structure or organization. None referenced language conventionality as defined by the CCSS (i.e., literal and conversational language versus figurative and unfamiliar language).

Perhaps had we specifically asked the participants to focus on the characteristics of texts for read-alouds, they would have raised some of the CCSS qualitative characteristics to

a greater extent. Although the writers of the CCSS did not distinguish between text characteristics for independent reading and read alouds, the exemplars identified in the CCSS Appendix B appear to differ in text characteristic emphases (NGA & CCSSO, 2010c). Many of the exemplars for independent reading for kindergarten and first grade in the CCSS Appendix B are from “I can read” series where words are high in decodability, frequency, and familiarity. The exemplar read-alouds follow the pattern of trade books that typically include words that are harder to decode, less frequent, and less familiar. Trade books dominate in the CCSS exemplars for read-on-your-owns in second grade and beyond.

One Size Does Not Fit All: Individualized Text Complexity

An important theme in the primary teachers' comments is the consideration of text complexity *in relation to their students*. They reported using different types of texts that reflect multiple levels of text complexity for various classroom activities and tasks. Shifting text types according to reading task suggests strong teacher awareness of a relationship between text characteristics, text complexity, and goals for students' reading development. When a few teachers wrote additional text characteristics that they used for the paired text comparison task, they noted the concept or content difficulty of the text *in relation to* children's background knowledge, and/or they considered *students'* interest in the topic, again signaling the importance of considering printed text characteristics relative to particular students. The teachers' focus on text-to-student complexity match-up is entirely consistent with the CCSS definition of text complexity: “The inherent difficulty of reading and comprehending a text combined with consideration of reader and task variables . . .” (NGA & CCSSO, 2010b, Appendix A, Glossary of Key Terms, p. 43). It also suggests that their internal theorization about text complexity may be based in a relational outlook that involves a particular text's complexity as it interfaces with a particular reader—an outlook that is consistent with an originating principle for the present study.

Implications for Professional Development, Teacher Education, and Policy

The primary teachers in our project share some understandings with the CCSS authors about what makes early grades texts complex, but they also bring divergent views. Similar to the clarion call of Mesmer and colleagues (2012) for text complexity frameworks that emphasize the unique aspects of texts for beginning readers, the participants placed heavy emphasis on text characteristics that are not prominent within either the quantitative or qualitative systems of text complexity in the CCSS. A major take-away from learning about the teachers' views is that professional development should be differentiated for teachers at different grade levels.

The findings of our project suggest that professional development and teacher education efforts to support the CCSS vision should build on what primary teachers know—in order to ensure that inappropriate practices are not mandated and to use professional wisdom as a bridge to new learning. Structural and other text level characteristics of text, such as repeated refrains and patterned syntax, are important in particular ways at the early grades level (e.g., Mesmer et al., 2012). Workshop and webinar providers who value teachers' outlooks on text complexity but also expand their outlooks, for instance, including information about the importance of text level characteristics, will likely have greater impact on teachers' decisions about text complexity and text choices for their students. Helping inservice and preservice teachers understand that structure and text characteristics that cut across sentences, such as repetitive refrains, matter at the early grade level would appear to be an important goal of professional development and teacher education. They matter because as was pointed out in the opening rationale, structure and repetition can facilitate children's word recognition strategies.

Additionally, the role of more complex characteristics of text could be considered in relation to a comprehensive literacy program. While our study participants reported using

different types of texts for different purposes, it would also be helpful for teachers to think about text characteristics of particular texts in relation to their total reading programs. For instance, teachers might use texts with deeper levels of meaning, more complicated structures, and unconventional language more often during literature circles as compared to small group reading instruction.

A strong policy implication is that standards makers such as CCSS might listen to teachers' opinions and experience about early grades text complexity and consider the relevance for the policies they create. Specifically, if the CCSS authors invited experienced primary teachers to offer views about which text characteristics ought to play into early grades text complexity, it is possible that the text-complexity standard might include a wider array of qualitative text characteristics for teachers to consider. Valuing educators' perspectives could enhance the possibilities for more thoughtful, careful, thorough, and robust implementation of classroom activities, which in turn could lead to an enhanced probability that students attain the standards that are set.

Future Research Directions

An important next step could be employing different research methods for giving voice to primary teachers' views of text characteristics that matter for text complexity. For instance, structured focus groups or structured individual interviews might reveal a wider array of text characteristics that teachers consider.

Another avenue for research is to develop new automated measurement systems designed specifically for early grades texts. Currently no text measurement system is designed specifically for the early grades, although a number of the existing systems provide text complexity measurement for at least a selected range of early grades texts. For instance, the numerical Lexile measure extends to just above 0L, below which the designation "Beginning Reader" is provided.

To develop such automated systems, researchers would first need to digitize a wide array of text characteristics—an array as wide or wider than that considered in the Coh-Matrix system research where 53 text variables were included in the analyses (Graesser et al., 2011). Then, the digitized measures could be used to explore which of the many characteristics best predicts text complexity levels. It is commonly understood that most current automated text complexity measurement systems rely principally on a small subset of text characteristics, namely, word frequency, word length, and/or sentence length (e.g., Degrees of Reading Power [DRP; Koslin, Zeno, & Koslin, 1987], Reading Maturity Metric [n.d.], and SourceRater [Sheehan, Kostin, Futagi, & Flor, 2010]). It is possible that a broader set of text characteristics may matter for measuring early grades text complexity. The possibility is supported by the teachers' voices in the present research in that they suggested additional characteristics such as decodability.

Since no current text complexity measurement systems are designed specifically for early grades texts, such research could benefit teachers by describing important text characteristics through statistical analyses. Moreover, comparing the resultant list of important text characteristics that contribute to text complexity to those suggested by teachers in the present (and future) studies could provide further insights into possible voids in teachers' outlooks.

The Importance of a Strong Foundation

As we work to implement the vision of the CCSS, students' ability to scale the staircase of text complexity successfully in second grade through college and career will depend on a strong foundation in students' beginning reading experiences. By listening to the voices of teachers who have expertise in providing such a base, educators and policy makers can learn much about how the tasks and texts of beginning reading can be best

designed and implemented to ensure that students are truly highly literate for the tasks and college and career texts.

References

- ACT. (2006). *Reading between the lines: What the ACT reveals about college readiness in reading*. Iowa City, IA: Author.
- Adams, M. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
- Adams, M. (2009). The challenge of advanced texts: The interdependence of reading and learning. In E. H. Hiebert (Ed.), *Reading more, reading better: Are American students reading enough of the right stuff?* (pp. 163-189). New York, NY: Guilford.
- Aukerman, R. C. (1984). *Approaches to beginning reading* (2nd ed.). New York: John Wiley & Sons.
- Baumann, J. F., Hoffman, J. V., Duffy-Hester, A. M., & Ro, J. M. (2000). The First R yesterday and today: U.S. elementary reading instruction practices reported by teachers and administrators. *Reading Research Quarterly*, 35(3), 338-377.
- Bean, R. M., Cassidy, J., Grumet, J. E., Shelton, D. S., & Wallis, S. R. (2002). What do reading specialists do? Results from a national survey. *The Reading Teacher*, 55(8), 736-744.
- Beck, I. L. & McKeown, M. G. (1991). Social studies texts are hard to understand: Mediating some of the difficulties. *Language Arts*, 68, 482-490.
- Bond, T. G., & Fox, C. M. (2007). *Fundamentals in the human sciences* (2nd Edition). Mahwah, NJ: Erlbaum.
- Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal*, 5, 69-85.
- Clay, M. M. (2001). *Change over time in children's literacy development*. Portsmouth, NH: Heinemann.

- Duke, N. K. (2000). 3.6 minutes a day: The scarcity of informational text in first grade. *Reading Research Quarterly, 35*, 202-224.
- Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 14*, 135-163.
- Fink, A. (2003). *How to sample in surveys* (2nd ed.). Thousand Oaks, CA: Sage.
- Fitzgerald, J. (2000). On distances: From campus teacher preparation to first grade literacy instruction and back again. *Phi Delta Kappan, 81*, 455-461.
- Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychology, 93*, 3-22.
- Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*, 223-234.
- Greenfield, R., Rinaldi, C., Proctor, P., & Cardarelli, A. (2010). Teachers' perceptions of RtI reform in an urban elementary school: A consensual qualitative analysis. *Journal of Disability Policy Studies, 21*(2), 47-63.
- Hayes, D. P., Wolfer, L. T., & Wolfe, M. F. (1996). Sourcebook simplification and its relation to the decline in SAT-Verbal scores. *American Educational Research Journal, 33*, 489-508.
- Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement*. Washington, DC: Alliance for Excellent Education.
- Hiebert, E. H. (2011). Texts for beginning readers: The search for optimal scaffolds. In C. Conrad & R. Serlin (Eds.), *The Sage handbook for research in education: Pursuing ideas as the keystone of exemplary inquiry* (pp. 413-428). Thousand Oaks, CA: Sage.

- Howes, D. H., & Solomon, R. L. (1951). Visual duration thresholds as a function of word probability. *Journal of Experimental Psychology*, 92, 248-255.
- Juola, P. (2003). Assessing linguistic complexity. In M. Miestamo, K. Sinne Sinnemäki, & F. Karlsson (Eds.), *Language complexity: Typology, contact, change* (pp. 89-108). Philadelphia, PA: John Benjamins Publishing.
- Klare, G. R. (1974-1975). Assessing readability. *Reading Research Quarterly*, 10, 62-102.
- Koons, H. (2011, February). *Examining the text difficulty continuum and its role in preparing students for success*. Paper presented at the annual meeting of the Pacific Coast Research Conference, Coronado, CA.
- Koslin, B. I., Zeno, S., & Koslin, S. (1987). *The DRP: An effective measure in reading*. New York: College Entrance Examination Board.
- Martin, S. D., & Shapiro, E. S. (2011). Examining the accuracy of teachers' judgments of DIBELS performance. *Psychology in the Schools*, 48(4), 343-356.
- McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-287.
- Meltzer Kleinhenz, S., & Kleinhenz, S. M. (2008). *Elephants*. North Mankato, MN: Capstone Press.
- Merlini Barbaresi, L. M. (2003). Towards a theory of text complexity. In L. Merlini Barbaresi (Ed.), *Complexity in language and text* (pp. 23-66). Pisa, Italy: Edizioni Plus.
- Merlini Barbaresi, L. M. (2003). Towards a theory of text complexity. In L. Merlini Barbaresi (Ed.), *Complexity in language and text* (pp. 23-66). Pisa, Italy: Edizioni Plus.
- Mesmer, H. A. (2006). Beginning reading materials: A national survey of primary teachers' reported uses and beliefs. *Journal of Literacy Research*, 38, 389-425.

- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly, 47*, 235-258.
- Metsala, J. L. (1999). Young children's phonological awareness and non-word repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*, 3-19.
- Moss, B., & Newton, E. (2002). An examination of the informational text genre in basal readers. *Reading Psychology, 23*, 1-13.
- Murray, M.S., Munger, K.A., & Hiebert, E.H. (in press). An analysis of two reading intervention programs: How do the words, texts, and programs compare? *Elementary School Journal*.
- Muter, V., Hulme, C., Snowling, M. J., Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*, 665-681.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010a). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors. Available at www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010b). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*, Appendix A. Washington, DC: Author. Available at http://www.corestandards.org/assets/Appendix_A.pdf
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010c). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects*, Appendix B. Washington, DC: Author. Available at http://www.corestandards.org/assets/Appendix_B.pdf

- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2012). *Supplemental information for Appendix A of the Common Core State Standards for English language arts and literacy: New research on text complexity*. Washington, DC: Author. Retrieved from Common Core State Standards Initiative website: <http://www.corestandards.org/resources>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction: Reports of the Subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Patton, M. (1990). *Qualitative evaluation research methods*. Beverly Hills, CA: Sage.
- Reading Maturity Metric. (n.d.). <http://www.readingmaturity.com/rmm-web/#/>
- Schwanenflugel, P. J., & Akin, C. E. (1994). Developmental trends in lexical decisions for abstract and concrete words. *Reading Research Quarterly, 29*, 250-264.
- Shanahan, T. (2011). Common Core Standards: Are we going to lower the fences or teach kids to climb? *Reading Today, 29*(1), 20-21.
- Sheehan, K. M., Kostin, I., Futagi, Y., & Flor, M. (2010, December). *Generating automated text complexity classifications that are aligned with targeted text complexity standards* (ETS RR-10-28). Princeton, NJ: Educational Testing Service.
- Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation.
- Stenner, A. J., Burdick, H., Sanford, E., & Burdick, D. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement, 7*, 307-322.

- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 41*, 225-234.
- Survey Monkey. (n.d.) www.surveymonkey.com
- Vadasy, P. F., Sanders, E. A., Peyton, J. A. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities, 38*, 364-382.
- van der Sluis, F., & van den Broek, E. L. (2010). Using complexity measures in information retrieval. In *Proceedings of the third symposium on information interaction in context* (pp. 18-22). New Brunswick, NJ: ACM.
- Williamson, G., Fitzgerald, J., & Stenner, A. J. (in press). Student Growth in Reading Illuminates the Common Core Text-Complexity-Exposure Standard: Raising Both Bars. *Elementary School Journal*.
- Williamson, G., Fitzgerald, J., A. J. Stenner. (2013). The Common Core State Standards Complexity Goal: Figuring out how much complexity is enough. *Educational Researcher, 42*(2), 59-69.
- Wirt, J., Choy, S., Rooney, P., Provasknik, S., Sen, A., & Tobin, R. (2004). *The condition of education 2004 (NCES 2004-077)*. U. S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Retrieved from <http://nces.ed.gov/pubs2004/2004077.pdf>.
- Woolams, A. M. (2005). Imageability and ambiguity effects in speeded naming: Convergence and divergence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 878-890.
- Wright, B., & Stone, M. (1999). *Measurement essentials* (2nd edition). Wilmington, DE: Wide Range Incorporated.

References: Texts in Table 2

- Adler, D. A., & Adler, M. S. (2010). *A picture book of John and Abigail Adams*. New York: Holiday House.
- Averill, E. (1960/1988). *The fire cat*. New York: HarperCollins.
- Bogacki, T. (2008). *The champion of children: The story of Janusz Korczak*. New York: Farrar Straus Giroux.
- Brown, C. L. (2006). *After the dinosaurs: Mammoths and fossil mammals*. New York: HarperCollins.
- Burns, D. (2000). *Clare's secret*. Columbus, OH: SRA/McGraw-Hill.
- Capucilli, A. S. (1996). *Biscuit*. New York: HarperCollins.
- Cole, J. (2010). *The magic school bus and the climate challenge*. New York: Scholastic.
- Cowley, J. (1998). *Microscope*. Columbus, OH: Wright Group/McGraw-Hill.
- Davidson, S. (retold from J. & W. Grimm) (2005). *The frog prince*. London: Usborne.
- Eastman, P. D. (1960/1988). *Are you my mother?* New York: Random House.
- Elliott, C. (2006). *Building a birdhouse*. Logan, IA: Perfection Learning.
- Floca, B. (2009). *Moonshot: The flight of Apollo 11*. New York: Simon & Schuster.
- Fredericks, A. D. (2001). *Under one rock: Bugs, slugs and other ughs*. Nevada City, CA: Dawn Publications.
- Hawes, A. (2000). *Quork attack*. Boston: Rigby/Houghton Mifflin Harcourt.
- Ipcizade, C. (2008). *Zebras*. Mankato, MN: Capstone.
- Jarman, J. (2000). *Bully bear*. Boston: Rigby/Houghton Mifflin Harcourt.
- Kortenkamp, S. (2009). *Spacecraft*. Mankato, MN: Capstone.
- Llewellyn, C., & Lovsin, P. (2010). *Boo-hoo, Baby!* London: Wayland.
- Maslen, B. L. (2006). *Bump!* New York: Scholastic.

- Maslen, B. L. (1976/2006). *Mat*. New York: Scholastic.
- Mackinnon, M. (2007). *The sun and the wind (retold)*. London: Usborne.
- McClintock, B. (2005). *Cinderella*. New York: Scholastic.
- Meister, C. (1999). *Catch that cat!* New York: Scholastic.
- Melser, J. (2000). *One, one, is the sun*. Columbus, OH: Wright Group/McGraw-Hill.
- O'Brien, A. (2002). *Seth's bath*. Columbus, OH: SRA/McGraw-Hill.
- O'Connor, J. (2008). *Fancy Nancy at the museum*. New York: HarperCollins.
- Osborne, M. P. (2009). *A good night for ghosts*. New York: Random House.
- Ransom, C. (2005). *I like shoes!* New York: Scholastic.
- Rueda, J. (2006). *Plant life*. Washington, DC: National Geographic Society.
- Rylant, C. (2001). *Poppleton in winter*. New York: Scholastic.
- Seuss, Dr. (2003). *Green eggs and ham*. London: HarperCollins.
- Silverman, E. (2005). *Cowgirl Kate and Cocoa*. Orlando, FL: Harcourt.
- Verne, J. (adapted by Jane Bingham). (2004). *Around the world in eighty days*. London: Usborne.
- Williams, L. D. (2005). *Cranes*. Mankato, MN: Capstone.
- Willems, M. (2008). *I love my new toy!* New York: Scholastic.

Table 1

Participant Demographics (N=90)

Category	Sub-categories	Percentages
Gender	Female	93.3
	Male	6.7
Role	Reading specialist	40.0
	Classroom teacher	43.3
	Media specialist	16.7
Level taught in past 3 years	K	14.4
	1 st grade	8.9
	2 nd grade	12.2
	2 or more primary grades	64.5
Ethnicity	Caucasian	86.7
	African American	6.7
	Asian or Pacific Islander	2.2
	Other	4.4
Teaching experience	5 or fewer years	28.9
	6-10 years	22.2
	11-15 years	20.0
	16 years or more	28.9
Type of school	Public school	87.7
	Private or charter schools	12.3
Geographic context	Urban & suburban	72.2
	Rural	27.8
Socioeconomic status of school (subsidized lunch)	Fewer than 25%	24.4
	26-55%	21.1
	56-85%	32.2
	86 or more	22.2

Table 2

Text Examples with Study “Level” Designation and Teacher Ranking Logit Measure

Title	Author	Designated “Level” for the Study	Teacher Ranking Logit Measure
<i>Catch that Cat!</i>	C. Meister	Easiest	-13.62
<i>Mat</i>	B. L. Maslen	Easiest	-9.46
<i>One, One, Is the Sun</i>	J. Melser	Easiest	-7.55
<i>I Like Shoes!</i>	C. Ransom	Easiest	-6.81
<i>Boo-hoo, Baby!</i>	C. Llewellyn & P. Lovsin	Easiest	-4.83
<i>Biscuit</i>	A. S. Capucilli	Easiest	-4.49
<i>Microscope</i>	J. Cowley	Easiest	-4.33
<i>The Sun and the Wind</i>	M. Mackinnon retold from Aesop	Easiest	-3.59
<i>I Love My New Toy!</i>	M. Willems	Easiest	-3.43
<i>Are You My Mother?</i>	P. D. Eastman	Easiest	-2.43
<i>Green Eggs and Ham</i>	Dr. Seuss	Easiest	-1.92
<i>Cranes</i>	L. D. Williams	Easiest	-0.76
<i>Bump!</i>	B. L. Maslen	Moderately	-2.64

PRIMARY TEACHERS' VIEWS OF TEXT COMPLEXITY

<i>Bully Bear</i>	J. Jarman	hard Moderately	-1.46
<i>Zebras</i>	C. Ipcizade	hard Moderately	-0.44
<i>Seth's Bath</i>	A. O'Brien	hard Moderately	0.18
<i>Building a Birdhouse</i>	C. Elliott	hard Moderately	0.90
<i>Plant Life</i>	J. Rueda	hard Moderately	2.26
<i>Fancy Nancy at the Museum</i>	J. O'Connor	hard Moderately	2.44
<i>Poppleton in Winter</i>	C. Rylant	hard Moderately	3.63
<i>The Fire Cat</i>	E. Averill	hard Moderately	3.70
<i>After the Dinosaurs: Mammoths and Fossil Mammals</i>	C. L. Brown	hard Moderately	4.96
<i>Spacecraft</i>	S. Kortenkamp	hard Moderately	5.14
<i>Cinderella</i>	B. McClintock retold from the C. Perrault version	hard Moderately	5.20
<i>Cowgirl Kate and Cocoa</i>	E. Silverman	Hardest	2.90
<i>The Frog Prince</i>	S. Davidson	Hardest	3.56

PRIMARY TEACHERS' VIEWS OF TEXT COMPLEXITY

<i>Clare's Secret</i>	retold from J. & W. Grimm D. Burns	Hardest	3.94
<i>Moonshot: The Flight of Apollo 11</i>	B. Floca	Hardest	4.74
<i>A Good Night for Ghosts</i>	M. P. Osborne	Hardest	5.87
<i>The Magic School Bus and the Climate Challenge</i>	J. Cole	Hardest	7.90
<i>Quork Attack</i>	A. Hawes	Hardest	7.98
<i>Under One Rock: Bugs, Slugs and Other Ughs</i>	A. D. Fredericks	Hardest	8.42
<i>Around the World in Eighty Days</i>	J. Verne (adapted by Jane Bingham)	Hardest	11.45
<i>A Picture Book of John and Abigail Adams</i>	D. A. Adler & M. S. Adler	Hardest	12.02
<i>The Champion of Children: The Story of Janusz Korczak</i>	T. Bogacki	Hardest	13.71

Note. Complete references are provided in "References: Texts in Table 2."


PRIMARY TEACHERS' VIEWS OF TEXT COMPLEXITY

Table 3

Screen Shot of the Paired-Text-Comparison Text-Complexity Task

Please select the book you believe presents more challenge for independent reading

Enduro motorcycles have headlights and brake lights. Riders use these for racing at night or on paved streets. These bikes also have tires with deep **tread** for off-road racing.



tread—a series of bumps and deep grooves on a tire

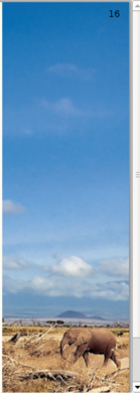
FACT: Enduro courses often wind through woods and over hills. Some courses even cross streams.

10

This word is the keyword found at

This book presents more challenge

Some months are dry on the savanna. Elephants must walk far to find food and water.



16

This book presents more challenge

Note. Texts are from Doeden (2010) (left) and Meltzer Kleinhenz and Kleinhenz (2008) (right). Permission granted Capstone Publishers, April 28, 2014.

PRIMARY TEACHERS' VIEWS OF TEXT COMPLEXITY

Table 4
Questionnaire Items

Category	Questions
Contextual information about teachers' reading instruction	<ol style="list-style-type: none"> 1. Approximately what proportion of your reading instruction time do you spend doing the following activities (activities that predominate United States classrooms today)? <ul style="list-style-type: none"> · Reading aloud to children · Guided reading · Time for student individual reading 2. For the instructional activities in number 1, if you do an activity, what type(s) of text do you use? <ul style="list-style-type: none"> · Basal reader · Content textbook · Leveled reader · Decodable book · Other 3. For which of the following topics have you received information/education through in-service workshop(s), conference session, university course, or similar activity? <ul style="list-style-type: none"> · Guided reading · Repeated reading · Phonological awareness · Decoding · Phonics · Comprehension · Reading workshop · Literature circles · Whole language · Literacy Centers · Differentiation · Fluency 4. Which of the entries in number 3 do you use in your instruction?
Text characteristics used to determine texts' complexity in the paired-text-comparison activity	<ol style="list-style-type: none"> 5. To what extent did you use each of the following text characteristics during the paired-text-comparison task to determine which text was more challenging? (Never = 1, Seldom, Sometimes, Often, Usually, Always = 6) <ul style="list-style-type: none"> · Font/text size · Amount of text per page · Pictures/illustrations · Word decodability · Proportion of high frequency words · Word meaning · Sentence length · Other (please comment) 6. Rank order each of the characteristics you used from number 5 above.
Teachers' confidence in determining text complexity	<ol style="list-style-type: none"> 7. Indicate your agreement (Completely Disagree = 1, Somewhat Disagree, Disagree, Agree, Somewhat Agree, Completely Agree = 6): <ul style="list-style-type: none"> · I know how to determine text complexity.
Teachers' confidence in matching text-complexity levels to students' reading levels	<ol style="list-style-type: none"> 8. Indicate your agreement for each of the following (Completely Disagree = 1, Somewhat Disagree, Disagree, Agree, Somewhat Agree, Completely Agree = 6): <ul style="list-style-type: none"> · I know each of my students' reading ability levels. · I understand each of my students' reading interests. · I am confident about matching students with books that are appropriately challenging.

Table 5

Mean Likert Rating for Importance (Standard Deviations)(n) and Mean Rank Ordering (Standard Deviations)(n) for the Extent to Which Teachers Reported Using Various Text Characteristics During the Paired-Text-Comparison Activity

Text Characteristic	All teachers				Reading/ Literacy Specialists				Other teachers			
	<i>M (SD)</i>	<i>N</i>	<i>M Rank (SD)</i>	<i>N</i>	<i>M (SD)</i>	<i>n</i>	<i>M Rank (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M Rank (SD)</i>	<i>n</i>
Word decodability	5.3 (0.8)	90	2.6 (1.9)	88	5.4 (0.8)	36	2.6(1.9)	34	5.3 (0.8)	54	2.6 (1.9)	54
High frequency words	5.2 (0.8)	90	3.0 (1.9)	90	5.4 (0.8)	36	3.0(1.8)	36	5.1 (0.8)	54	3.1 (2.0)	54
Pictures/illustrations	5.0 (1.0)	88	3.8 (1.9)	90	4.9 (1.1)	35	3.9 (2.1)	36	5.0 (0.9)	53	3.7 (1.7)	54
Meaningful words to young children	4.9 (1.0)	89	4.3 (1.9)	90	5.0 (1.1)	35	4.2 (1.9)	36	4.9 (0.9)	54	4.3 (2.0)	54
Sentence length	4.9 (1.0)	90	4.6 (1.9)	87	4.9 (1.0)	36	4.3 (2.0)	34	4.9 (1.0)	54	4.8 (1.8)	53
Amount text per page	4.9 (1.0)	90	4.9 (1.7)	89	4.7 (1.1)	36	5.0 (1.5)	36	5.0 (0.8)	54	5.0 (1.8)	53
Font/text size	4.3 (1.4)	89	5.2 (1.6)	90	4.3 (1.7)	35	5.1 (1.7)	36	4.2 (1.4)	54	5.2 (1.6)	54
Other	4.4 (1.6)	50	7.8 (1.1)	90	4.6 (1.6)	21	7.8 (0.7)	36	4.3 (1.7)	29	7.7 (1.3)	54

Note. For Means: 1 = Never, 2 = Seldom, 3 = Sometimes, 4 = Often, 5 = Usually, and 6 = Always. For Mean Ranks, 1 = Most used text characteristic, 2 = Next most used, etc. Where N's/n's for mean Likert rating and mean rank differ, some teachers either rated more characteristics than they ranked or ranked more characteristics than they rated.

PRIMARY TEACHERS' VIEWS OF TEXT COMPLEXITY

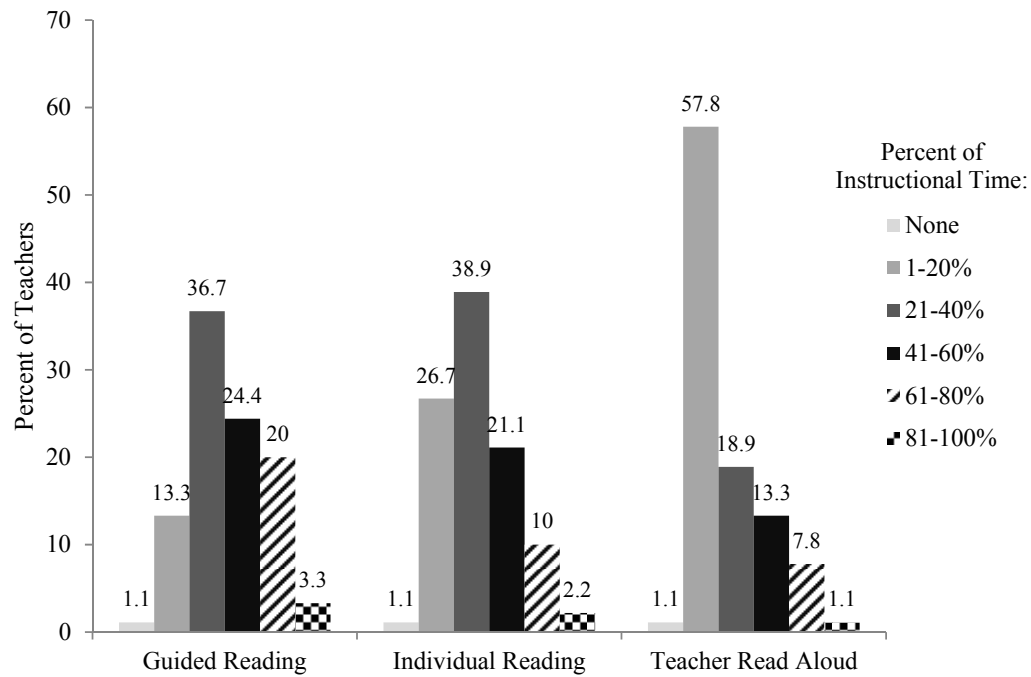


Figure 1. Percent of teachers reporting various amounts of instructional time for three types of reading instruction activities.

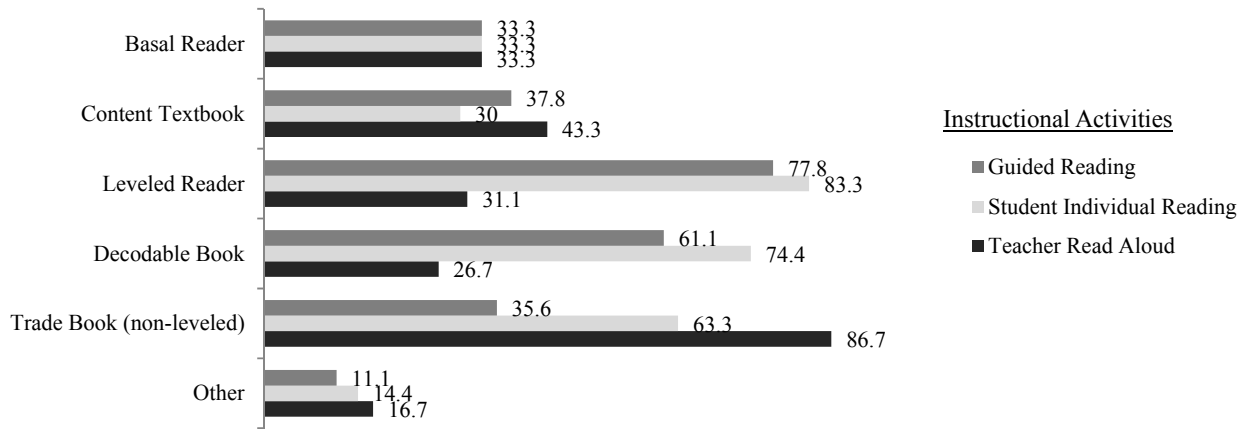


Figure 2. Percent of teachers who reported using various types of texts for three reading instruction activities.

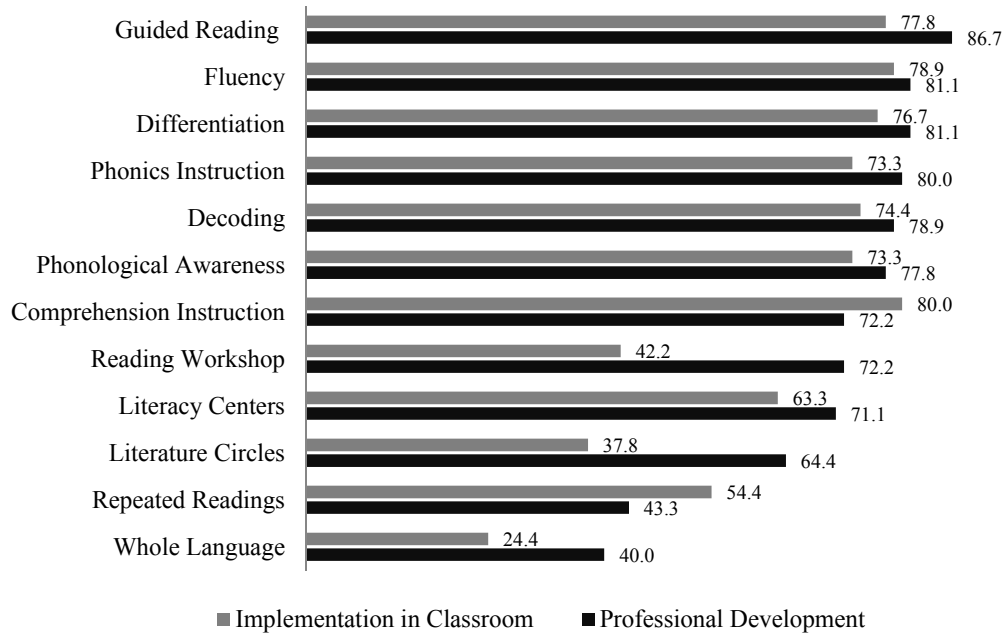


Figure 3. Percent of teachers who encountered prior education on various reading topics.

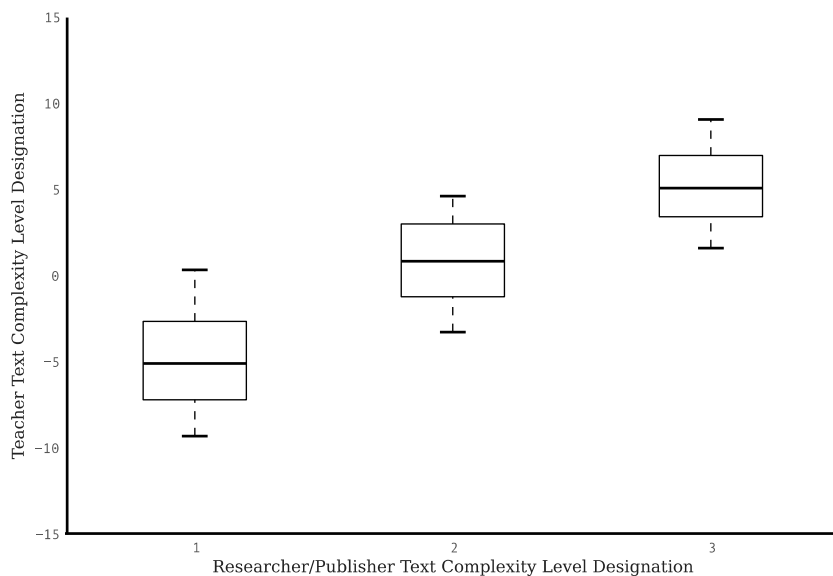


Figure 4. Teacher Text Complexity Level Designation by Researcher/Publisher Text Complexity Designation. On the X-axis, 1 = Easy, 2 = Moderately Hard, and 3 = Hard. Interquartile ranges are noted inside the boxes, with the lower line of the box representing the 25th percentile, the middle line the 50th percentile, and the upper line the 75th percentile. The top “whisker” represents the 90th percentile, and the bottom “whisker” represents the 10th percentile.