

WHAT RESEARCH SAYS ABOUT TEXT COMPLEXITY AND LEARNING TO READ

Richard L. Allington ■ Kimberly McCuiston ■ Monica Billen

Inherent in the English Language Arts Common Core State Standards (CCSS) is the notion that students should be reading “complex works of literature” (National Governors Association Center for Best Practices [NGA Center] & Council of Chief State School Officers [CCSSO], 2010, p. 3). The greatest increases in text complexity occur for grades 2–5 (Hiebert & Mesmer, 2013). This development stems from the push for college-ready individuals and unambiguous evidence that school texts had been decreasing in complexity over time (NGA Center & CCSSO, 2010, p. 2).

However, the NGA Center and CCSSO did not provide empirical evidence that this is the case. Both studies reviewed for the CCSS actually indicated a curvilinear relationship for the complexity of school texts over time. Both studies (Chall, 1977; Hayes, Wolfer, & Wolfe, 1996) reported that texts began becoming less complex around the 1930s, and this trend extended through the 1950s. But in the late 1960s, school texts began increasing in complexity, a pattern of change

that has continued through today. Thus, at least in grades 3 and 6, school texts have higher levels of complexity today than at any time in the past 65 years (Gamson, Lu, & Eckert, 2013). This finding undermines the argument for the need to increase text complexity in the elementary grades made by the NGA Center and CCSSO and leaves us wondering about the role complex text should have in the elementary classroom.

Throughout this article, we discuss the research on text complexity. We begin by grounding the importance of this topic in current events; namely, the implementation of the CCSS. From there, we briefly

Richard L. Allington is a professor of literacy studies at the University of Tennessee, Knoxville, Tennessee, USA, and past president of the International Reading Association; e-mail rallingt@utk.edu.

Kimberly Flanders McCuiston is a PhD candidate in teacher education with a concentration in literacy studies at the University of Tennessee, Knoxville, Tennessee, USA; e-mail kmccuist@utk.edu.

Monica Billen is a PhD candidate in teacher education with a concentration in literacy studies at the University of Tennessee, Knoxville, Tennessee, USA; e-mail mbillen@utk.edu.

discuss the historical background of text complexity and text appropriateness. Next, we cover research that has been published in regards to oral reading accuracy as a measure of text complexity in both the classroom setting and intervention setting. We touch on the influence that text difficulty, or complexity, may have on student engagement, vocabulary knowledge, and self-regulatory practices. We also highlight the problems that we have found with the available research on oral reading accuracy. Additionally, we discuss the research we have found that links the use of hard texts to reading achievement. To conclude, we add our thoughts and assertions from this body of literature.

Text Complexity and the CCSS

School texts have *not* been decreasing in complexity. After analyzing hundreds of grade 3 and grade 6 school readers published from 1905 to 2004, Gamson and colleagues noted, "The blanket condemnation made by the CCSS authors that school reading texts have trended downward in difficulty in the last half-century is inaccurate" (p. 388). Especially at grade 1 and grade 3, significantly more complex texts are being used today than ever before. At grade 6, the texts have shown "remarkable stability" over time, but the readers used in grade 6 today are at least as complex as readers published at any previous period.

Hiebert and Mesmer (2013) analyzed the fourth-grade text used on the

2006 National Assessment Educational Progress (NAEP) reading assessment. They noted that the NAEP text was at a Lexile level of 620, around the proposed midpoint CCSS span for grades 2 and 3. Two-thirds of the fourth graders performed below Proficient and one-third performed below the Basic level of proficiency on that NAEP reading assessment. The fourth graders performed dismally using second- and third-grade-level texts following the guidelines found in the CCSS framework! Thus, the proportion of fourth graders who will perform adequately at Lexile levels 790 to 820 on the new CCSS assessments will, indeed, be quite small.

Attempts to perform well on such assessments will require new materials, a lot of professional development, and greater resources. However, the United States Department of Education provides additional dollars only for more testing. Left unaddressed by the CCSS framework are the two-thirds of fourth graders today who cannot meet the older, traditional reading-level standards. We fail to see how reporting that even more children are failing to meet the standards set by the CCSS authors can be considered progress or even beneficial, especially given that fourth-grade reading performance on the NAEP assessments has reached an all-time high (National Center for Educational Statistics, 2013).

Even less comprehensible is why the NGA Center and CCSSO targeted the elementary grades to produce annual growth of 163 Lexile levels,

while middle and secondary grades are expected to produce only 53 Lexile levels of annual growth (Hiebert & Mesmer, 2013). The CCSS put the onus of developing readers who can read considerably more complex texts on teachers of kindergarteners to teachers of students through grade 5. The CCSS authors could have indicated expected improvement in reading texts of greater complexity and spread the load equally across the 13 years of public education. But that is not what the NGA Center and CCSSO did. They frontloaded their demands for increasing the difficulty of school texts on elementary schools, leaving middle and high schools largely to carry on with students who would now be able to read increasingly complex school texts.

All of this has been accomplished by misrepresenting what the research on text complexity indicated and without a shred of research evidence that the CCSS will improve reading achievement. It is almost enough for one to wish the CCSS authors and the USDE were offering suggestions based in scientifically reliable, replicable research! This, then, begs the question: what does research say about text complexity?

Historical Background of Oral Reading Accuracy and Text "Appropriateness"

We have 70 years of evidence that children are more likely to learn to read and to learn content when the text can be read with a high level of accuracy and comprehension. It all began in the 1940s, when Emmet Betts developed his levels of text complexity based on oral reading accuracy and comprehension of what had been read (Betts, 1946). This was a direct measure of text complexity noting, simply, how accurately a student read a text. Precisely how he set the levels is a bit of a mystery

"Children are more likely to learn to read and to learn content when the text can be read with a high level of accuracy and comprehension."

“Betts recommended that children first read the test passage silently before being asked to read it aloud.”

(Shanahan, 1983) and cannot be said to have involved extensive research.

According to Betts (1946), students are reading at their independent level when they can read with at least 99% word-reading accuracy and 90% comprehension. His standards for the instructional level are slightly lower: accuracy rates between 95% and 98% word recognition and between 75% and 89% comprehension. Betts characterized a student’s frustration level by either word recognition below 90% or comprehension below 50%. In establishing all of these levels, Betts recommended that children first read the test passage silently before being asked to read it aloud.

Over the years, a number of scholars have questioned Betts’s levels, with some finding evidence to support his recommendations (Ehri, Dreyer, Flugman, & Gross, 2007; Fuchs, Fuchs, & Deno, 1982; Gambrell, Wilson, & Gantt, 1981; Gickling & Armstrong, 1985; Jorgenson, Klein, & Kumar, 1977; Morris, Bloodgood, Perney, Frye, Kucan, & Trathen, 2011; O’Connor, Bell, Harty, Larkin, Sackor, & Zigmond, 2002; Treptoe, Burns, & McComas, 2007) and others finding evidence that Betts’s levels needed to be adjusted (Johns & Magliari, 1989; Pikulski, 1974; Powell, 1970).

Also, a number of scholars have, over the years, provided reviews of oral reading accuracy research and noted the various problematic issues in research designs (Allington, 1984b; Halladay, 2012; Leu, 1982; Weber, 1968; Wixson, 1979). Thus, the issues of using oral

reading accuracy levels as standards for the complexity of texts that are to be read in reading instructional settings has received much attention, but not much of that attention has appeared in the past two decades.

In addition to reviews of the research, there has also been a constant stream of argument against having children read aloud, especially in assessment situations (Huey, 1908; Hunt, 1970; Smith, 2012). The general argument offered is that oral reading places additional performance demands on readers as compared to silent reading. Additionally, our goal is to create children who can read silently and understand the text. Mosenthal (1977) demonstrated that oral and silent reading produced different outcomes, and he suggested that the two modes of reading were distinct processes and evidence gathered with one mode should not be used to generate statements about reading in the other mode. Still after debate and argument, Betts’s original oral reading accuracy criteria have largely stood the test of time as an acceptable procedure for determining the difficulty, or complexity, of a text that is, or might be, used for reading instruction in elementary schools.

Oral Reading Accuracy in Reading Acquisition: The Classroom

Several studies have been conducted that provided evidence that helps us better understand the importance of oral reading accuracy during classroom instruction. The studies we detail were

year-long studies of elementary-grade classrooms, studies where oral reading accuracy was measured and linked to reading development.

Anderson, Evertson, and Brophy (1979) reported an experimental study of first-grade reading instruction where teachers and their students were assigned to either a treatment or a control group. The treatment group teachers were simply given a 22-point written handout of guidelines to follow when working with reading groups. Most of the guidelines focused more on management than on instruction, although the principles did deal with responding to children as they read. Not surprisingly, the researchers found that treatment teachers implemented a larger number of instructional behaviors keyed on the handout than did control group teachers and that the students in treatment teachers’ classrooms made greater progress in developing reading proficiency. Anderson et al. also reported that the average number of errors made while reading aloud was negatively correlated with reading growth. Specifically, they wrote, “A high rate of success may be especially important when students are asked to read a passage aloud. The more mistakes made during the average reading turn, the lower the achievement” (p. 216). Commenting on the most effective classrooms, they reported that about 75% of all reading turns in these

“The average number of errors made while reading aloud was negatively correlated with reading growth.”

“The most effective teachers structured oral reading opportunities where most children read with 100% accuracy!”

classrooms were completed without oral reading errors. In other words, the most effective teachers structured oral reading opportunities where most children read with 100% accuracy!

A second, and similar, study was carried out by a team of researchers from the Far West Laboratory for Educational Research and Development (Berliner, 1981; Fisher & Berliner, 1985; Fisher et al., 1978). One difference from the Anderson et al. (1979) study was that observations were completed for second- and fifth-grade classrooms. The researchers reported that the average number of errors made while reading aloud correlated negatively with achievement (second grade, $r = -.36$; fifth grade, $r = -.20$). They also reported negative correlations between high error rates and engagement. In other words, oral reading errors were linked, as in Anderson et al. (1979), with smaller amounts of reading growth. However, the correlations at fifth grade were much smaller than at second grade, leading the authors to conclude that “For younger children, less difficult materials led to higher achievement gain, whereas for older children the difficulty level of the materials was not related to achievement gain” (p. 787). Thus, it may be that reading texts at high levels of oral reading accuracy is more important for beginning readers than for older and more accomplished readers.

These studies of elementary-grade reading instruction found that oral reading accuracy is an important factor in determining reading development. Higher levels of oral reading accuracy

during reading lessons produced larger gains in reading development than did reading with lower levels of accuracy. These studies support using texts that fit Betts’s original oral reading accuracy criteria.

The third and final study also examined elementary school readers. Morris et al. (2011) assessed two cohorts of subjects enrolled in grades 2–6 across four consecutive years on a variety of informal reading assessment tasks, including oral reading accuracy. They found that the student oral reading accuracy data “provided empirical support for traditional but seldom-tested performance criteria in reading diagnosis; for example, oral reading accuracy = 95%” (p. 226). Thus, the “craft knowledge,” as Morris and colleagues label it, of using texts that can be read with at least 95% accuracy is again supported by the evidence.

Oral Reading Accuracy in Reading Acquisition: Struggling Reader Interventions

A second area of research has focused on oral reading accuracy during intervention lessons for struggling readers. In these cases, the oral reading accuracy during classroom reading lessons was not documented; only the accuracy of reading while participating in the reading intervention. Ehri and her colleagues (2007) studied the effects of a specific tutorial program for first-grade language minority students. They found that certified teachers produced greater gains than did paraprofessionals

and also that oral reading accuracy during the intervention lessons was important. They posited, “The reading achievement of students who received... tutoring appeared to be explained primarily by one aspect of their tutoring experience—reading texts at a high level of accuracy, between 98% and 100%” (p. 441). Teachers provided more such experiences than did paraprofessionals, either because teachers were better at selecting appropriate texts or because they were better at introducing selected texts to children. Regardless, once again, elementary-aged struggling readers experienced greater reading growth when provided with primarily high-accuracy oral reading experiences.

Working with older (grades 3–5) struggling readers, O’Connor and her colleagues (2007) provided a daily tutorial for 18 weeks, with the level of text difficulty used as the key variable. All students had to achieve a reading level below the beginning second-grade level and had to read at a rate below 80 words per minute on second-grade-level text to be included in the intervention. Students (about half classified as pupils with a learning disability) were randomly assigned to be tutored using grade-level classroom texts or using selected reading-level-matched texts. Text complexity ranged from 3.5 grade level to 7.0 grade level—roughly one and a half

“Higher levels of oral reading accuracy during reading lessons produced larger gains in reading development...”

years to three and a half years above the reading levels of the participants. Texts for the reading-level treatment group were judged by oral reading performances, and texts that could be read with 90% to 95% accuracy were selected for this group. These texts, then, were largely at each student's instructional level as determined using Betts's criteria. A third group of similar students served as the control group and did not participate in the tutorial interventions.

Tutoring sessions for both the grade-level and the reading-level groups were organized so that students spent 5 minutes of each tutoring period focused on phonological or word work activities, 20 minutes focused on text reading, and 5 minutes focused on comprehension support activities. Students from the same classrooms were assigned to different treatment groups whenever possible in order to minimize classroom teacher effects.

Students in both tutorial treatment groups produced higher levels of reading gains than students in the control (no tutorial) group. But there were differential effects of the two treatment groups depending on the degree of reading difficulties students had exhibited at the beginning of the intervention. When the poorest readers with the lowest words correct per minute (WCPM) scores are compared with better reading participants who demonstrated higher oral reading WCPM rates, the reading-level tutoring produced significantly larger gains than grade-level tutoring on word identification, word attack, and fluency. While students in the grade-level group outperformed students in the control group, the reading-level students outperformed both groups. In other words, students who read texts at their instructional reading level outperformed students who read texts above their instructional reading level.

“Elementary-aged struggling readers experienced greater reading growth when provided with primarily high-accuracy oral reading experiences...”

O'Connor et al. (2007) concluded, “The proposition that poor readers will make stronger comprehension gains by reading in grade level texts with appropriate support (e.g., assisted reading) was not borne out here” (p. 483). This study contradicts the assertion of Susan Pimentel, one key author of the CCSS, who stated, “There is no research that says, if an eighth grader is reading at a fifth-grade level, and you give him fifth-grade texts, he will catch up” (quoted in Rothman, 2012, p. 4). There might not be a specific study using eighth-grade students and fifth-grade-level texts, but the O'Connor et al. study did demonstrate that struggling readers will be more likely to “catch up” when tutored with texts at their instructional level than if their tutors had provided them with grade-level texts.

Like the studies of text difficulty in classroom reading lessons reviewed above, both of these intervention studies indicate that the difficulty of texts that students are expected to read is an important factor when the focus is on struggling readers and that more difficult texts produce less growth than texts matched to Betts's original instructional-level criteria.

Influences of Text Difficulty Engagement

Several studies have provided evidence as to why having students read texts with at least 95% accuracy may result in larger gains in reading achievement than having those students read texts

that can be read only with lower levels of accuracy. Jorgenson (1977) found that when students were placed in texts easier than their predicted instructional level, they were observed to be better behaved and more independent. Similarly, Gambrell et al. (1981) reported that when students were reading with accuracy levels above 95%, they were on task during 42% of the lesson period. At the same time, students reading texts with lower levels of accuracy were on task only 22% of the time. Both Gickling and Armstrong (1985) and Treptoe et al. (2007) also reported that students reading texts at the 95% accuracy level or higher were more likely to be on task and more likely to demonstrate greater comprehension of the material they had read than when reading text where their oral reading accuracy fell below the 95% level.

Vocabulary

Consider also that independent reading seems to be the source of most vocabulary acquisition (Nagy, Anderson, & Herman, 1987; Nagy, Herman, & Anderson, 1985; Stahl, 1999; Swanborn & DeGlopper, 1999). It isn't that every unknown word encountered while reading is learned but that most of the words adults know were acquired while reading independently, not from vocabulary lessons. Anderson (1996) summarized the research on learning word meanings while reading independently by noting, “The overall likelihood [of learning the meanings of new words through

“Students who read texts at their instructional level outperformed students who read texts above their reading level.”

reading] ranged from better than 1 in 10 when children were reading easy narratives to near zero when they were reading difficult expositions” (p. 61). Learning the meanings of unknown words only 10% of the time wouldn’t be an effective strategy if good readers didn’t read millions of words each year. But good readers acquire an ever-increasing vocabulary, which makes learning more new words even more likely, if only because their larger store of known vocabulary enhances the power of virtually every context encountered while reading. This enhanced context means that learning the meaning of an unknown word becomes more likely. However, if a text has many unknown words, the context will be too weak to facilitate even word recognition, let alone learning the new word’s meaning. Of course, for new vocabulary to be learned, the reader must encounter at least a few unknown words. That seems right where Betts’s original criteria—at least 95% oral reading accuracy—places the reader.

Self-Regulating Behaviors

Leslie and Osol (1978) examined readers’ self-corrections and use of decoding strategies as they read texts of varying levels of difficulty. Text difficulty mattered on both: “The results of the present study support the recommendation that children be instructed from materials which they can read orally with at least 95% accuracy. It seems that readers use self-correction strategies and attempt to use graphic clues to decode unknown words when they read material with at least 95% accuracy” (p. 444).

Self-correction is an important self-regulating strategy, one that good readers develop early in their reading development. As Clay noted in her 1969 study, high-progress primary-grade readers spontaneously corrected 1 of every 3 oral reading errors, while the low-progress readers corrected only 1 of every 20 oral reading errors. Self-correction is an essential aspect of the self-teaching hypothesis. The argument, made best by Share (1995), is that much of what good readers demonstrate is acquired through self-teaching while engaged in voluntary, independent reading. However, when texts are too difficult, self-teaching is suppressed. Stanovich (1992) describes the situation faced by struggling early readers after establishing that struggling readers engage in reading far less frequently than do good readers:

Further exacerbating the problem of differential exposure is the fact that less-skilled readers find themselves in materials that are too difficult for them (Allington, 1977, 1983, 1984[a]; Bristow, 1985; Forell, 1985; Gambrell et al., 1981). The combination of deficient decoding skills, lack of practice, and difficult materials results in unrewarding early reading experiences that lead to less involvement in reading-related activities. Lack of exposure and practice on the part of

the less-skilled reader delays the development of automaticity and speed at the word recognition level. Thus, reading for meaning is hindered, unrewarding reading experiences multiply, and practice is avoided or merely tolerated without real cognitive involvement. (p. 328)

Thus, it would seem that the best research evidence currently available supports the use of texts that can be read with at least 95% accuracy for instructional purposes. In sum, these studies all found that students reading texts with at least 95% accuracy were more likely to be on task during classroom reading lessons and more likely to understand what had been read. Both of these factors linked to oral reading accuracy would serve students well in developing their reading proficiencies. The more off-task behavior and lower comprehension exhibited by students reading texts with low accuracy interferes with learning and, in these cases, with learning to read.

Hard Texts and Reading Achievement

Having reviewed the studies relevant to using texts that can be read with at least 95% accuracy, we now discuss two studies where students were expected to read harder texts that produced lower oral reading accuracy levels. The first study, conducted by Stahl and Heubach (2005), involved whole-class reading instruction from grade-level readers. The lesson design included use of Oral Recitation Lesson format, which included the teacher reading the text aloud first, then

“High-progress primary-grade readers spontaneously self-corrected 1 of every 3 oral reading errors, while low-progress readers corrected only 1 of every 20 errors.”

the students rereading the text multiple times in varying formats. The steps included a story map introduction, teacher reading the basal story aloud, repeated reading of the basal story, partner reading of the basal story, choice reading every day (20 minutes), and home rereading of the basal story. Students still having difficulty were provided with two additional reading supports: echo reading, where an adult read the text as the student attempted to follow along, or a segment of the complete text was read repeatedly to foster fluency.

With this level of support, these first-grade (and then second-grade) students did progress in reading, making an average of 1.88 and 1.77 grade levels growth in each of the two years the project was in place. As Stahl and Heubach (2005) noted,

Students were able to benefit from reading material at these lower levels of accuracy because of the higher levels of support they were given for the reading through the routines of the program. In this program, students were supported in their reading by having multiple exposures to the same material, by having stories read to them, by exposure to the vocabulary prior to their own reading, by reading the story at home one or more times, possibly by echo reading, and by partner reading... The instructional reading level for a given child is inversely related to the degree of support given to the reader. (p. 55)

The second study, reported by Morgan, Wilcox, and Eldredge (2000), employed “dyad reading” for 15 minutes a day for 95 days (or roughly 24 hours of dyad reading activity). During the 15-minute period, a better reader read aloud with a weaker reader. The better readers led the oral reading performance, with the weaker reader also reading but perhaps following the lead of the better reader. There were three levels of text difficulty used. One group

“Students reading texts with at least 95% accuracy were more likely to be on task...and more likely to understand what had been read.”

used texts two levels above the reading level of the struggling readers, a second group used texts three levels above, and a third group used texts four levels above. All struggling readers tested at the non-reader, pre-primer or primer levels of reading development at the beginning of the study.

The group using texts two levels above their reading level made the greatest gains. On an end-of-study task that involved students reading a trade book, the group that had read texts at two levels above the reading level of the struggling readers again had the strongest performance on reading accuracy, fluency, and comprehension. Less progress was made when using the most difficult texts, but even the texts two levels above probably could not have been read with 95% accuracy by the struggling readers. However, the one-on-one assisted reading support provided by the skilled reader partner proved sufficient support to allow the progress. The authors noted that in such assisted reading contexts, teachers might use texts that are more difficult than the texts usually deemed appropriate for struggling readers.

As far as we can tell, these are the only two studies where “frustration-level” texts have been used productively. However, in both cases, struggling readers were given enormous additional supports with either multiple rereadings or assisted reading. Neither of these supports is typical of classroom reading instruction. Perhaps if classroom lessons were altered such that these levels

of support were available every day for every reader, then it might be beneficial to use texts that can be read at accuracy levels below 95%. This must be counterbalanced, though, against the studies noted above, which reported that the incidences of oral reading errors were negatively related to reading development in typical classroom environments.

Finally, we found no studies comparing the outcomes of providing students with texts that could be read with at least 95% accuracy with the outcomes obtained by using harder texts. Both of the harder texts studies remind us of the sorts of “studies” that commercial publishers use to try and convince customers that their program really works. However, almost no publisher ever provides a comparative study of their program with a different program. None provide truly randomized experiments that demonstrate the superiority of their products. We know that one-to-one assisted reading can facilitate development of reading proficiency. But would providing that same assisted reading practice in texts that could be read with at least 95% accuracy have produced the same levels of growth? Or perhaps even greater growth? Providing students with teacher read-alouds, lots of partner reading activity, plus home rereading and echo reading of texts at the 95% accuracy level might also produce reading gains as great or greater than the gains found when using more difficult texts. For answers to such questions, we will have to hope that someone does the research.

“Texts with a little bit of a challenge is important if reading growth is expected.”

Easier Texts and Achievement

Carver and Liebert (1995) demonstrated that 24 hours of reading texts a year or two below the reading levels of the students produced almost no positive achievement effects with third-, fourth-, and fifth-grade students. This study was done during the summer months, where students were expected to read for two hours every day. Whether this six-week program was simply too brief or whether 24 hours of reading is insufficient to produce measurable reading growth, we will never know. But Carver and Liebert suggested that reading texts with a bit of a challenge is important if reading growth is to be expected. We are in general agreement with their assertion, if only because in our (Allington et al., 2010) study of summer voluntary self-selected reading, significant reading achievement effects were observed for students given the opportunity to participate in a book fair just before the end of the school year. In this study, some students from low-income families were simply provided the opportunity to attend a book fair and self-select 15 books for summer voluntary reading. When compared to the control group of students who selected no summer books, reading achievement of the students who self-selected books was significantly higher. However, greater gains were observed for students who chose texts at rather than text below their reading achievement levels.

Studies of Oral Reading Accuracy: Problematic Aspects

While a number of studies supporting the use of texts that can be read with at least 95% oral reading accuracy are available, the set of studies on the effects of oral reading accuracy on reading development provide hardly any consistent scheme for research designs. The first area of concern is how an error is defined. In many cases, there is no description of how oral reading accuracy rates were determined. In a few cases, such details are given, but there seems to be little consistency in defining an error. For instance, in some studies, only misreadings of words that were not self-corrected by the student were scored as errors. In other studies, such self-corrections might be counted as two or more errors (one error for the initial misreading and a second error for the self-correction/repetition).

A second area of concern is the difficulty of the text on which students were tested. In some cases all readers, good and poor, read the same texts. In other studies, students read multiple texts which typically varied in complexity. When all students read the same text, it makes comparisons easier but problematic because the texts invariably are harder for struggling readers than they are for good readers. This is important because all studies do depict increasing text difficulty as the source for an increasing number of oral reading

errors. Some studies indicated that more difficult texts produce greater numbers of nonsense word errors and greater numbers of omission/no response errors. Thus, if text difficulty is held constant, then readers of differing levels of proficiency make different sorts of errors as well as differing numbers of errors (Biemiller, 1979; Stanovich, 1992). It would seem that we need more studies that compare readers of different proficiency levels reading texts at comparable levels of difficulty as measured by rates of oral reading accuracy.

In addition to the level of text difficulty, there is another area of concern: the nature of the classroom reading curriculum. In general, studies have reported different error patterns in beginning readers taught in a code-emphasis curriculum than are observed in beginning readers taught with a meaning-emphasis curriculum (Barr, 1972, 1974; Cohen, 1975; Dank, 1977). At the same time, most studies reporting oral reading accuracy levels fail to describe the reading instructional materials the students are reading, either in the classroom or in the intervention. It shouldn't be surprising that beginning readers exposed to a code-emphasis reading curriculum produce more nonsense word errors and more no response errors. Similarly, it shouldn't be surprising that students exposed to a meaning-emphasis reading curriculum produce more semantically and syntactically acceptable errors. The key

“Texts that can be read with 95% or greater accuracy are directly, and in some studies causally, related to improved reading achievement.”

question, which is largely unaddressed, is how students in either type of instructional program refine their strategy use and become proficient readers.

Finally, how reading adequacy is judged is also problematic. Some studies linked oral reading error rates to achievement on standardized tests but did not report on the immediate comprehension of the texts read during instruction. Other studies evaluated on-task behaviors while reading texts of different levels of difficulty. Only a very few studies provided data on reading lessons over the course of a school year and then linked error rates in oral reading to student achievement growth. In short, attempting to summarize what the research says about text complexity and learning to read is in itself an exercise in complexity.

The evidence accumulated suggests that texts that can be read with 95% or greater accuracy are directly, and in some studies causally, related to improved reading achievement. Texts that are read with either significantly lower or higher levels of accuracy fail to produce positive effects as large as the “just right” texts—“just right” being texts that can be read with 95% accuracy or higher.

Conclusion

Our review of available research on text complexity and learning to read demonstrates two important points. First, elementary textbooks have not been getting less complex as the CCSS authors asserted and, thus, the complexity levels of texts used in the elementary grades is not the source of the problem that the CCSS are supposed to solve. Hence, increasing the complexity of the texts used in elementary schools as the best strategy for enhancing reading achievement, as the CCSS authors recommend, lacks a base in the research evidence available.

Second, a number of studies have demonstrated that texts used for instruction that can be read with at least 95% accuracy produce greater gains than harder texts. Two studies have demonstrated that progress can be made using harder texts, but only if substantial individual support is given by an adult or a better reader. However, the instructional supports used in both studies seem impractical given the demands already placed on classroom teachers. Additionally, we do not know if the same gains would have been made with easier texts or if the gains were simply the result of giving students a lot of practice with a lot of support.

While the original criteria set by Betts (1946) were not established on any strong research base, research evidence accumulated over the past 70 years indicates that Betts’s original oral reading performance criteria are accurate and reliable, at least in primary-grade classroom reading lessons and in reading interventions for struggling readers in the elementary grades. At the same time, the question of the appropriateness of the 95% accuracy level across upper grade levels remains largely unstudied. Most studies to date have used elementary school students as subjects and have not gathered evidence over a long term (e.g., grades 1–4), but they have established that progress in developing reading proficiency over a shorter term is best supported by using “just right” texts.

We argue that the research evidence available suggests that before more complex texts are used in reading lessons, two things should occur. First, research should be conducted to provide us with better evidence on the potential for positive effects of using more complex texts in our reading lessons. Second, we need better evidence of instructional scaffolding that might be best used to facilitate

“We recommend that elementary teachers...adhere to the traditional oral reading accuracy criteria.”

just how more complex texts can be used to enhance reading development.

We contend that in order for students to become proficient readers who are engaged in text while self-regulating and building vocabulary knowledge, the text must appropriately match the student’s reading level. We fear that the push from the CCSS to promote the use of more complex texts will result in decreased reading engagement and less time spent reading, with a potential decline in reading achievement the ultimate outcome. We recommend that elementary-grade teachers continue to adhere to the traditional oral reading accuracy criteria of instructional texts that can be read at 95% accuracy or higher until the outcomes of research on both issues is available.

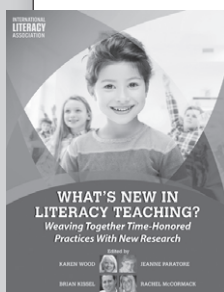
REFERENCES

- Allington, R.L. (1977). If they don’t read much, how they ever gonna get good? *Journal of Reading*, 21(1), 57–61.
- Allington, R.L. (1983). The reading instruction provided readers of differing abilities. *The Elementary School Journal*, 83(5), 548–559.
- Allington, R.L. (1984a). Content coverage and contextual reading in reading groups. *Journal of Reading Behavior*, 16(1), 85–96.
- Allington, R.L. (1984b). Oral reading. In P.D. Pearson (Ed.), *Handbook of reading research* (pp. 829–864). New York, NY: Longman.
- Allington, R.L., McGill-Franzen, A.M., Camilli, G., Williams, L., Graff, J., Zeig, J., ... & Nowak, R. (2010). Addressing summer reading setback among economically disadvantaged elementary students. *Reading Psychology*, 31(5), 411–427.

- Anderson, L.M., Everson, C.M., & Brophy, J.E. (1979). An experimental study of effective teaching in first-grade reading groups. *The Elementary School Journal*, 79(4), 193–223.
- Anderson, R.C. (1996). Research foundations to support wide reading. In V. Greaney (Ed.), *Promoting reading in developing countries* (pp. 55–77). Newark, DE: International Reading Association.
- Barr, R. (1974). The effect of instruction on pupil reading strategies. *Reading Research Quarterly*, 10(4), 555–582.
- Barr, R.C. (1972). The influence of instructional conditions on word recognition errors. *Reading Research Quarterly*, 7(3), 509–529.
- Berliner, D.C. (1981). Academic learning time and reading achievement. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 203–225). Newark, DE: International Reading Association.
- Bristow, P.S. (1985). Are poor readers passive readers? Some evidence, possible explanations, and potential solutions. *The Reading Teacher*, 39(5), 318–325.
- Betts, E.A. (1946). *Foundations of reading instruction: With emphasis on differentiated guidance* (4th ed.). New York, NY: American Book Company.
- Biemiller, A. (1979). Changes in the use of graphic and contextual information as functions of passage difficulty and reading achievement level. *Journal of Reading Behavior*, 11(4), 307–319.
- Carver, R.P., & Liebert, R.E. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly*, 30(1), 26–48.
- Chall, J.S., Conard, S.S., & Harris, S. (1977). *An analysis of textbooks in relation to declining SAT scores*. New York, NY: Advisory panel on the Scholastic Aptitude Test score decline.
- Clay, M.M. (1969). Reading errors and self-correction behaviour. *British Journal of Educational Psychology*, 39(1), 47–56.
- Cohen, A.S. (1974). Oral reading errors of first grade children taught by a code emphasis approach. *Reading Research Quarterly*, 10(4), 616–650.
- Dank, M. (1977). What effect do reading programs have on the oral reading behavior of children? *Reading Improvement*, 14(1), 66–69.
- Ehri, L.C., Dreyer, L.G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414–448.
- Fisher, C.W., & Berliner, D.C. (1985). *Perspectives on instructional time*. New York, NY: Longmans.
- Fisher, C.W., Filby, N.N., Marliave, R., Cahen, L.S., Dishaw, M.M., Moore, J.E., & Berliner, D.C. (1978). *Teaching behaviors, academic learning time, and student achievement: Final report of Phase III-B, beginning teacher evaluation study*. San Francisco, CA: Far West Laboratory for Educational Research and Development.
- Forell, E.R. (1985). The case for conservative reading placement. *The Reading Teacher*, 38(9), 857–862.
- Fuchs, L.S., Fuchs, D., & Deno, S. (1982). Reliability and validity of curriculum-based informational reading inventories. *Reading Research Quarterly*, 18(1), 6–25.
- Gambrell, L.B., Wilson, R.M., & Gantt, W.N. (1981). Classroom observations of task-attending behaviors of good and poor readers. *Journal of Educational Research*, 74(6), 400–404.
- Gamson, D.A., Lu, X., & Eckert, S.A. (2013). Challenging the research base of the Common Core State Standards: A historical re-analysis of text complexity. *Educational Researcher*, 42(7), 381–391.
- Gickling, E.E., & Armstrong, D.L. (1985). Levels of instructional difficulty as related to on-task behavior, task completion, and comprehension. *Journal of Learning Disabilities*, 11(9), 32–39.
- Halladay, J.L. (2012). Revisiting key assumptions of the reading level framework. *The Reading Teacher*, 66(1), 53–62.
- Hayes, D.P., Wolfer, L.T., & Wolfe, M.F. (1996). Schoolbook simplification and its relation to the decline of SAT-Verbal scores. *American Educational Research Journal*, 33(2), 498–509.
- Hiebert, E.H., & Mesmer, H.A.E. (2013). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher*, 42(1), 44–51.
- Huey, E.B. (1908). *The psychology and pedagogy of reading*. New York, NY: MacMillan.
- Hunt, L.C. (1970). The effect of self-selection, interest, and motivation upon independent, instructional, and frustration levels. *The Reading Teacher*, 24(1), 146–151.
- Johns, J.L., & Magliari, A.M. (1989). Informal reading inventories: Are the Betts criteria the best criteria? *Reading Improvement*, 26(2), 124–132.
- Jorgenson, G.W. (1977). Relationship of classroom behavior to the accuracy of the match between material difficulty and student ability. *Journal of Educational Psychology*, 69(1), 24–32.
- Jorgenson, G.W., Klein, N., & Kumar, V.K. (1977). Achievement and behavioral correlates of matched levels of student ability and materials difficulty. *Journal of Educational Research*, 71, 100–103.
- Leslie, L., & Osol, P. (1978). Changes in oral reading strategies as a function of quantities of miscues. *Journal of Literacy Research*, 10(4), 442–445.
- Leu, D.J. (1982). Oral reading error analysis: A critical review of research and application. *Reading Research Quarterly*, 17(3), 420–437.
- Morgan, A., Wilcox, B.R., & Eldredge, J.L. (2000). Effect of difficulty levels on second-grade delayed readers using dyad reading. *Journal of Educational Research*, 94(2), 113–119.
- Morris, D., Bloodgood, J.W., Perney, J., Frye, E.M., Kucan, L., Trathen, W., ... & Schlagal, R. (2011). Validating craft knowledge: An empirical examination of elementary-grade students' performance on an informal reading assessment. *The Elementary School Journal*, 112(2), 205–233.
- Mosenthal, P.B. (1977). Psycholinguistic properties of aural and visual comprehension as determined by children's abilities to comprehend syllogisms. *Reading Research Quarterly*, 12(1), 55–92.
- Nagy, W., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237–270.
- Nagy, W., Herman, P.A., & Anderson, R.C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233–253.
- National Center for Educational Statistics. (2013). *NAEP 2013 reading: A report card for the nation and the states*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from http://nationsreportcard.gov/reading_math_2013
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.
- O'Connor, R.E., Bell, K.M., Harty, K.R., Larkin, L.K., Sackor, S.M., & Zigmond, N. (2002). Teaching reading to poor readers in the intermediate grades: A comparison of text difficulty. *Journal of Educational Psychology*, 94(3), 474–485.
- Pikulski, J.A. (1974). A critical review: Informal reading inventories. *The Reading Teacher*, 28(2), 141–153.
- Powell, W.R. (1970). Reappraising the criteria for interpreting informal inventories. In D.L. DeBoer (Ed.), *Reading diagnosis and evaluation* (pp. 100–109). Newark, DE: International Reading Association.
- Rothman, R. (2012). The complex matter of text complexity. *Harvard Education Letter*, 28(5), 1–2.
- Shanahan, T. (1983). A critique of P. A. Kilgallon's study: A study of relationships among certain pupil adjustments in reading situations. In L. Gentile, M. L. Kamil & J. Blanchard (Eds.), *Reading research revisited* (pp. 577–582). Columbus, OH: Merrill.
- Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151–218.
- Smith, F. (2012). *Understanding reading: A psycholinguistic analysis of reading and learning to read* (6th ed.). New York, NY: Routledge.

WHAT RESEARCH SAYS ABOUT TEXT COMPLEXITY AND LEARNING TO READ

- Stahl, S.A. (1999). *Vocabulary development: From reading research to practice*. Newton Upper Falls, MA: Brookline.
- Stahl, S.A., & Heubach, K. (2005). Fluency-oriented reading instruction. *Journal of Literacy Research*, 37(1), 25–60.
- Stanovich, K.E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 307–342). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Swanborn, M.S.L., & DeGlopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, 69(3), 261–286.
- Treptoe, M.A., Burns, M.K., & McComas, J.J. (2007). Reading at the frustration, instructional, and independent levels: The effects on students' reading comprehension and time on task. *School Psychology Review*, 36(1), 159–166.
- Weber, R.M. (1968). The study of oral reading errors: A review of the literature. *Reading Research Quarterly*, 4(1), 96–119.
- Wixson, K.L. (1979). Miscue analysis: A critical review. *Journal of Reading Behavior*, 11(2), 163–179.



© 2015
 ISBN 978-0-87207-166-7
 Nonmembers: \$19.95
 Members: \$15.95

What's New in Literacy Teaching? Weaving Together Time-Honored Practices With New Research

KAREN WOOD, JEANNE PARATORE,
 BRIAN KISSEL, & RACHEL McCORMACK
 Editors

Learn the latest literacy teaching strategies with this collection of articles from the *What's New in Literacy Teaching?* ILA E-ssentials series—with the addition of two exclusive new chapters on classroom discourse and reading assessment—that will pump up your practice AND meet your PD needs.

INTERNATIONAL
LITERACY
 ASSOCIATION

ORDER NOW! literacyworldwide.org/ebookwnlt

800.336.7323 (U.S. and Canada) | 302.731.1600 (all other countries)