Changing Readers, Changing Texts: Beginning Reading Texts from 1960 to 2010

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

TextProject Article Series
March 2014

TextProject, Inc.
SANTA CRUZ, CALIFORNIA
Abstract

Over the past 50 years, substantial changes have occurred in the texts used for beginning reading instruction. This article analyzes four prominent perspectives that have most influenced beginning reading instruction texts over this period. These changes in texts are examined in relation to changes over time in the nation’s demographics and the experiences that relate to young children’s literacy learning, including media exposure, availability of children’s books, and time spent in non-parental care. These analyses are followed by an examination of the research underlying three assumptions that shape the reading instruction found in all current core reading programs: (a) earlier is better, (b) word repetition is not a factor in learning to read, and (c) one size fits all. The review ends with a call for the design and selection of texts for the students who most depend on schools to become literate.
Changing Readers, Changing Texts:

Beginning Reading Texts from 1960 to 2010

Currently accepted definitions of reading comprehension all reflect a process that involves three interrelated elements—a reader, a text, and an activity, or purpose for reading—that shape and are shaped by the social and cultural context of the reading act (e.g., RAND Reading Study Group, 2002). Of these three elements, texts have been viewed to be most malleable and, as a result, often have been the focus of interventions and policies intended to promote reading achievement, particularly for beginning readers. Indeed, for several decades, state education policymakers have focused their attention on these beginning reading texts, issuing a series of mandates about specific features they must contain.

This article focuses on the texts intended for use with beginning readers, describing changes that have occurred in them over the past 50 years and examining these changes relative to the students who are expected to achieve literacy by using them. Changes in readers are more difficult to describe (and certainly to engineer) than are changes in textbooks, but the strengths and needs of students should be the foundation on which educational practice is based. Therefore, the first section of this article describes several societal and cultural factors that have affected young children’s schooling and literacy learning over the 50-year period. The second section addresses the major turning points in the evolution of beginning reading texts over this period. The third section examines the text-student match for different groups of readers. The article concludes with a proposal for developing texts that can improve that match for students who are now failing to attain proficient literacy levels.

Changes in Readers

The most obvious change in beginning readers over time has been in the demographics of
a school entry cohort. However, changes in children’s preschool experiences also potentially influence literacy development.

Changes in demographics

As the data in Figure 1 make clear, the demographics of an American grade cohort have changed remarkably over the past 50 years. Reflecting the nation’s strict national-origins quota-based immigration policies of the late 1920s, most children in the class of 1960 were born to parents who themselves had been born in the United States. This picture began to change with the passage of the Immigration and Naturalization Act of 1965, which abolished the quota system and led to a dramatic increase in immigration to the U.S. from the countries of Latin America and Asia. This trend continued through the 1990s and early 2000s. By 2007, the number of immigrants was three times the level it had been in 1970.

Table 1 shows what these changes in immigration patterns have meant for American schools. In 1960, 86% of the first-grade cohort was classified as white; in 2006, that number had declined to 56.5%. As the table shows, the largest shift in demographics has been in the increased percentage of students who are Hispanic. Differences in immigrant status and ethnicity can be related to diversity in children’s native languages, poverty levels, and literacy patterns in their homes and communities. These variables, in turn, can affect the literacy opportunities that children have before school entry. However, the manner in which selected demographic variables affect literacy knowledge cannot be compared across the 50-year period. This is because large-scale studies of young children’s literacy were begun only in the late 1990s, when the National Center for Education Statistics initiated two studies, the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) (Denton & West, 2002) and the

---

1 Throughout the paper, age cohorts will be identified according to the year of their entry into first grade. For example, the class of 2010 refers to the cohort that entered first grade in the fall of 2010.
Early Childhood Longitudinal Study, Birth Cohort of 2001 (ECLS-B) (Flanagan, McPhee, & Mulligan, 2009). It is the latter study that provides the most relevant information about the influences on kindergartners’ literacy knowledge of demographic variables. The scale score for literacy in this study combines letter recognition, letter-sound knowledge, recognition of words, phonological awareness, receptive and expressive vocabulary knowledge, and knowledge of print conventions. Of a total of 85, the average score is 43.9 and the standard deviation is 14.2. Table 2 summarizes data on the relationship of this measure to six variables: gender, race/ethnicity, poverty status, home language, family type, and age.

Although poverty status and race/ethnicity (especially for American Indian and Alaska Native children) influence levels of literacy knowledge, it is the age-related patterns that are most striking. From age five on, every additional six months of age is reflected in one-third of a standard deviation of higher literacy knowledge. The variance from age five to six is equivalent to the effects of a home language other than English.

Changes in children’s preschool experiences

From 1960 to 2010, the experiences of children as preschoolers changed in several significant ways germane to literacy learning: (a) increased exposure to literacy-related media (television and movies), (b) greater availability of children’s books, and (c) increased time spent in non-parental care.

Media exposure. The type and extent of media exposure of the first-grade cohort of 1960 differed significantly from previous generations. Whereas young children in the 1930s and following saw occasional movies, most children after 1960 grew up with televisions in their homes. American households with televisions increased from 9% in 1950 to 65% in 1955 and 87% in 1960. By 1975, the percentage had stabilized in the high 90s (U.S. Census Bureau, 2011),
but the amount of television programming aimed at children in the mid- to late-1950s was already substantial, with shows such as *Lassie* and host/puppet shows such as *Kukla, Fran and Ollie*.

From its earliest days, television had been marketed as a means for increasing children’s learning, although equally vocal claims had been made of its potential negative effects (Schramm, 1968). In that large-scale studies of young children’s learning were rare at that time, the effects of television viewing were not established with the class of 1960. With the launch of *Sesame Street* in 1969, however, documentation of the effects of educational programs on television began. *Sesame Street* was the pioneer of what became a burgeoning industry of educational programming for preschoolers (Fisch, 2004). Viewing of *Sesame Street* and its successors such as *Blue’s Clues* has been associated with increased literacy, numeracy, and general knowledge (Mielke, 2000). Long-term effects of frequent viewing of *Sesame Street* for literacy, including more leisure reading as adolescents, also have been reported (Huston, Anderson, Wright, Linebarger, & Schmitt, 2000).

Since the arrival of videos and cable television in the early 1980s, opportunities for young children to watch educational programs have increased beyond those available to early viewers of *Sesame Street*. In the late 1990s and early 2000s, the landscape changed again with a deluge of digital products for the preschool market. The benefits of many of these products (e.g., the Baby Genius series) have not been validated, but there can be no argument that they have become commonplace in the lives of today’s preschool children.

Because of a lack of data on children’s school-entry literacy knowledge, media effects on that knowledge remains uncertain, although available evidence points to at least some positive benefits of media exposure on a variety of cognitive and socioemotional outcomes (Troseth,
Saylor, & Archer, 2006).

*Availability of children’s books.* The low-cost Golden Books that had become widely available in the 1940s meant that children in the 1960s had access to books in greater abundance than did their parents and grandparents. But the Golden Books were only the beginning, as the titles of children’s books published annually went from approximately 1,500 in 1955 to almost 3,000 in 1980 to about 9,000 in 2000 (Bogart, 2010). With the publication of Dr. Seuss’s *Cat in the Hat* (Geisel, 1957), a new type of children’s books entered the marketplace: Books to support independent reading skills. Following the success of Dr. Seuss, similar books proliferated as part of trademarked programs (e.g., I can read, Easy readers).

The ECLS-K examined the effects on children’s literacy performance of their access to books at home and of read-alouds by parents (Denton, West, & Walston, 2003). Although parent read-alouds predicted literacy performance on kindergarten entry, the number of available books predicted end of first-grade performance. Levitt and Dubner (2005) chose this pattern as the focus of a chapter in their popular book, *Freakonomics*, to illustrate how statistical analyses uncover anomalies in human behavior. One of the explanations for the greater influence of number of books—and one not raised by Levitt (an economist) and Dubner (a journalist)—may lie in the influence of time spent with books on a broader and more extended variable—vocabulary (Paris, 2005).

*Time spent in non-parental care.* Without question, more preschoolers are spending more time in non-parental care settings in the 2000s than they did in the 1950s. But obtaining trustworthy information to contrast these patterns is obfuscated by factors such as lack of data from the earlier period. Current available data show, however, that at least for the year prior to kindergarten, approximately 74% of four-year-olds spend at least part of a day in non-home
settings (Flanagan et al., 2009).

Table 2 includes information on kindergartners’ entry-level literacy knowledge as a function of spending time in different early care/education settings in the year prior to kindergarten entry. For a child receiving only parental early care and a child in non-parental early care settings, the difference is .35 of a standard deviation but the effect of spending the year in Head Start rather than in parental early care is substantially less (.04 of a standard deviation). A year spent in a non-parental early care setting—especially if it is Head Start—does not approach the difference made by a child’s age on kindergarten entry.

Conclusions: Changes in readers

The presence of literacy-focused media, more books for children, and more time spent in preschool learning environments raises the question: Are children more literate than their counterparts 50 years ago? Once again, the lack of long-term data makes it impossible to provide a definitive answer to that question. However, the data in Table 3 do give a glimpse of the literacy knowledge of the 1998 kindergarten class. As these data show, children are learning letter names; but even with greater availability and numbers of books, increased time spent in preschool settings, and more exposure to literacy-focused media, few children learn to read prior to first-grade. What does appear to have changed over the past 50 years are the expectations of parents, educators, and the public about formal reading instruction for young children. In general, the dominant advice from schools and experts to parents in the 1950s and early 1960s was “hands off” in formally teaching children to read (Durkin, 1966). In the recent report of the National Early Literacy Panel (NELP) (Lonigan & Shanahan, 2009), formal instruction in literacy-related skills is assumed in kindergarten and identified as a critical part of the preschool curriculum as well. These changes in expectations as well as those found in texts are discussed in
the following section.

Changes in Texts

Currently, only a handful of publishers—three publishers produce the five programs that dominate the market—are the primary vendors for the core reading programs that compete nationally for district- and statewide adoptions. The programs sold by these publishers vary little from one another in either philosophy or components (Foorman, Francis, Davidson, Harm, & Griffin, 2004). For this reason, the analyses in this section focus on the texts contained in one program. This specific program was chosen for two reasons: (a) of the two core reading programs reviewed by Chall (1967/1983) in her landmark study, it is the only one still published, and (b) it is typical of current core reading programs (Foorman et al., 2004; Hiebert, 2005).

Publishers typically produce a new edition, or copyright, of a program every four to five years (Chall & Squire, 1991). One copyright can be almost identical to a previous one, as was the case with the Scott Foresman Basic Readers from 1940 (Gray, Baruch, & Montgomery, 1940) through the mid-1960s (Robinson, Monroe, & Artley, 1962). Copyrights can differ markedly, however, when states such as California and Texas, which adopt textbooks for statewide use, mandate the inclusion of new features in those textbooks. The copyrights of the program chosen for these analyses represent four watersheds in beginning reading texts: (a) copyright 1962 (Robinson et al., 1962), which typifies the stress placed on the repetition of high-frequency words in beginning reading texts from 1930 through the late 1980s; (b) copyright 1993 (Allington et al., 1993), which contains texts that reflect the whole language movement’s push for the use of authentic literature; (c) copyright 2000 (Afflerbach et al., 2000), which marks the emergence of decodable texts to accompany the strong phonics emphasis in instruction; and (d) copyright 2008 (Afflerbach et al., 2008), which shows the changes in expectations about formal
literacy instruction for young children. Beginning reading programs contain many components (e.g., word and sentence cards, ancillary sets of books, numerous workbooks), but their core are the five or six books, called readers or anthologies, that contain grade-specific reading selections. It is the anthology selections that are the focus of these analyses.

The indices used by Chall (1967/1983) in her analyses are the basis for the comparison of text features of the four copyrights presented here. The first set of indices relates to cognitive load: (a) type-token ratio, or the number of unique words relative to total words in a text, and (b) number of words that appear only once in a text. The second set addresses linguistic content. Each unique word is analyzed on two dimensions: (a) word frequency, which is defined as the percentage of words that are among the 300 most-frequent words in written English (Zeno, Ivens, Millard, & Duvvuri, 1995), and (b) grapho-phonic complexity, which is measured by the percentage of words that contain simple vowel patterns (i.e., CV, CVC, CVC-e, and CVVC) (Juel & Roper/Schneider, 1985). For each of the four copyrights, linguistic and cognitive data are summarized for the 10 texts that form the first and last units of the first-grade program.

1960-1990

Prior to 1930, books such as the Stickney Readers (Stickney, 1885/1985) consisted of selections that contained both high-frequency words and numerous phonetically regular words (e.g., sun, sat, bee). In 1925, Gray articulated a different perspective that would become the foundation for texts over a 60-year period. Appearing first in the Elston Readers (Elson & Gray, 1930), this text is often remembered for its characters (Dick and Jane and their family members, friends, and pets). Gray’s interest, however, lay in applying Thorndike’s (1903) laws of learning—readiness, exercise, identical elements, and reinforcement. By repeating a core group of high-frequency words, Gray believed that the laws of learning were applied in a form that
ensured successful reading acquisition. Since Thorndike’s laws of learning said nothing about what needed to be repeated in learning to read, Gray turned to Thorndike’s (1921) analysis of word frequencies, which showed that approximately 300 words accounted for almost 50% of all words in written English. In Gray’s model, these words became the unit of repetition and pacing. Because many of these frequently used words have variant letter-sound patterns (e.g., the, of, one, some), Gray’s instruction encouraged beginning readers to use a “look-say,” or memorization strategy rather than sounding out words.

As the word-features data in Table 4 show, 50% of the words in the 1962 copyright’s entry-level, first-grade texts fell into the 300 most-frequent group. The first unit of that copyright did have a substantial number of words with common vowel patterns, reflecting the names of characters (e.g., Dick, Spot, Jane) or high-frequency words (e.g., can, did, had), but these were not words chosen to represent particular letter-sound patterns. New words were introduced at the rate of 10 new words for every 100 words of text, which means that most words were repeated at least 10 times. No word appeared a single time in the first-grade entry text, but the exit-level text had a small percentage (7) of single-appearing words.

From 1930 when this model was first used in beginning reading programs (Elson & Gray, 1930; Gates & Huber, 1930) to 1990, when no mainstream reading programs used the model, the look-say approach was criticized (e.g., Flesch, 1955). These criticisms escalated with the increased scrutiny of America’s educational system following the Soviet Union’s successful Sputnik launch in 1957 (e.g., Bloomfield & Barnhart, 1961). But it was Chall (1967/1983) who made the issue prominent with the publication of Learning to Read: The Great Debate. Based on a summary of reading-acquisition studies and an analysis of textbook features, Chall concluded that the look-say approach failed to give beginning readers the code, or phonics emphasis needed
for a propitious start in reading. Stringent adherence to this approach had produced texts, Chall argued, that moved far too slowly and tediously. She argued for texts that offered a more strenuous pace and an enlarged vocabulary with more phonetically regular words.

A frequent assumption is that Chall’s (1967/1983) critique precipitated substantial changes in beginning reading programs. And it is true that the 1971 copyright of the Scott Foresman program (Goodman et al., 1971) did change course, dropping the Gray (1925) model and increasing the number of unique words (but not the number of phonetically regular words). When that program failed in the marketplace (Chall & Squire, 1991), however, the next copyright (and those of other publishers) returned to the look-say instructional approach. The substantial change after Chall’s critique was in the teacher’s manuals that accompany reading programs, which almost doubled in size as publishers responded to recommendations for increased phonics instruction. Rarely, though, did the content of these phonics lessons connect to the words in student reading selections (Beck & McCaslin, 1978).

1990-2000

As frameworks grounded in cognitive science and psycholinguistics were brought to the analysis of text in the 1970s and early 1980s, criticisms of the look-say model became compelling and frequent. The Commission on Reading summarized the cognitive science research in *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985). This report had two distinct themes that influenced the next generation of beginning reading texts. The first theme summarized Chall’s (1967/1983) updated analysis of the 1970s copyrights of texts, which she concluded remained focused on high-frequency words rather than phonetically regular words. The Commission recommended that beginning texts “should be interesting, comprehensible, and give children opportunities to apply phonics (Anderson et al., 1985, p. 118).
It identified *Green Eggs and Ham* (Geisel, 1960) as an exemplar of beginning text. Although this recommendation was based on beginning reading research, it was research from the report’s Extending Literacy chapter that influenced beginning reading texts most directly. That chapter summarized research on the obstacles to comprehension created by manipulations of text to comply with readability formulas. Not a single study had been conducted on the ability of completely uncontrolled text to affect reading acquisition positively when California education officials (California English/Language Arts Committee, 1987) and, soon after, those in Texas (Texas Education Agency, 1990), specified that programs accepted for their next textbook adoptions had to contain only “authentic” literature, or literature that was not constrained by readability formulas. By the early 1990s, the anthologies of all major reading programs consisted of authentic literature. Among teachers and the public, this instructional perspective was known as *whole language*.

In that these state mandates did not distinguish between texts for beginning reading and those for subsequent grade levels, the vocabulary of the first-grade texts was no longer controlled according to the Gray (1925) model. Instead, the new beginning reading texts used predictable text structures, in which key words are substituted within repeated phrases or sentences. A prototypical predictable text, *Brown Bear, Brown Bear* (Martin, 1967), differed substantially from *Green Eggs and Ham* (Geisel, 1960) in the amount and repetition of vocabulary. The type-token ratio of the entry-level unit in the 1993 copyright is 29:100 (see Table 4) and that of *Green Eggs and Ham* is 7:100, which means that the texts of the 1993 entry-level unit had four times as many unique words as *Green Eggs and Ham*. Given that most high-frequency words are function words (e.g., *I, can*), these words appeared in the whole language texts, but 46% of the unique words appeared a single time in the first 10 texts of the 1993 entry-level text. With no constraints
on the number of new, unique words or a need to repeat them, the rate of introducing new words from the entry- to exit-level texts is flat. The only accommodation is in length, with entry passages substantially shorter than exit passages (although substantially longer than entry passages in 1962).

2000 to 2008

The look-say model lasted for approximately 60 years as the driving philosophy for first-grade texts. The whole language perspective would last for approximately a decade as a vehicle of policy. When the results of the first state-by-state comparison of the National Assessment of Educational Progress (NAEP) were released in 1996 (Campbell, Donahue, Reese, & Phillips, 1996), the poor performances of California’s fourth graders were interpreted as a reflection of whole language policy (Levine, 1996). This precipitated a widespread demand for a return to reading instructional methods that stressed “the basics”—most specifically, phonics.

Because of state textbook adoption cycles, Texas (Texas Education Agency, 1997) preceded California (California Reading/Language Arts Committee, 1999) in mandating that programs provide increased phonics instruction accompanied with decodable texts. From among numerous types of texts that emphasize decodability, Texas chose the individual phoneme-grapheme as the focus of text creation, meaning that a program had to contain lessons and accompanying texts for each of the 44 phonemes in English. The texts were considered to be fully decodable, if the phoneme-graphemes of all words not taught as high-frequency or sight words had been presented in the current or a previous lesson (Stein, Johnson, & Gutlohn, 1999). This was deemed to ensure that children would have the potential to read all words in a text accurately. Both Texas and California set percentages for words in a text that needed to fit the potential-for-accuracy criterion: 80% (Texas Education Agency, 1997) and 90% (California
The data in Table 4 indicate that the percentage of phonetically regular words in reading texts increased from 24% in 1993 to 42% in 2000. Despite this difference in the kinds of words in texts, the 2000 copyright was similar to the 1993 copyright in the cognitive load, with the percentage of single-appearing words and type-token ratios much the same in both copyrights. Since decodability was defined as the match between the content of lessons and of students’ texts, repetition of individual words in the students’ texts was not a factor in evaluating texts for beginning readers in 2000.

**2008-present**

The most recent reading program copyrights (Afflerbach et al., 2008; Bear et al., 2008; Beck et al., 2008) retain the same types of words (i.e., decodable) and a similar rate of introduction as did the copyrights of the early 2000s. What has changed from 2000 to 2008 is when students are asked to read the first texts of the programs. Formal reading instruction in the core reading programs now begins in kindergarten.

Analyses of core programs from the late 1980s show no reading texts for kindergartners. At most, the reading readiness workbook used in kindergarten had some pages that could be folded into a small booklet (Hiebert & Papierz, 1990; Morrow & Parse, 1990). In an observational study of kindergarten reading instruction, Durkin (1987) reported that the majority of teachers emphatically stated *no* when asked whether reading should be taught in kindergarten. They believed that children should be taught to match letters to sounds but should not be involved in formal reading instruction.

In the early 1990s, publishers added big books to be used for read-along sessions to the kindergarten components of their reading programs (Allington et al., 1993). A set of decodable
texts was also included in the early 2000 kindergarten program copyrights (Afflerbach et al., 2000; Cooper et al., 2002), but these books were not the primary focus of instruction. By the 2007 Texas textbook adoption (Texas Education Agency, 2005) and the 2006 call for new texts in California (California Board of Education, 2006), formal reading instruction in kindergarten was a given.

The data on the kindergarten components of the 2008 copyright in Table 4 show that texts for exit-level kindergarten are highly decodable, with 80% of the words having simple vowel patterns (Juel & Roper/Schneider, 1985). In the entry level texts, accommodations include a high percentage of words that start with the featured initial consonants (e.g., muffin, mittens) and do not follow the CVC pattern that dominates the exit-level kindergarten texts. Texts for the kindergarten exit-level are also substantially shorter than were texts for first-graders in previous copyrights. At the same time, expectations for first graders’ reading abilities are higher, as is evident in the substantially longer average entry-level text.

Changing Readers, Changing Texts: The Match for the Students Who Depend on School to Become Literate

Without doubt, the factors most directly affecting levels of students’ literacy development over the past 50 years arise from the increased diversity within a school-entry cohort. In many American schools today, a high percentage of students speak English as a second language. A high percentage of students also live in low-income communities. Presumably, policymakers are responding to the needs of these students—and others who are most in need of additional help—when they issue new mandates for changes in beginning reading texts. The question, then, is: How well do the textbooks produced in response to these mandates initiate students, especially those from low-income and culturally diverse communities, into literacy?
The lack of long-term data on first graders’ performances complicates any efforts to correlate features of first-grade texts with student performance over time. However, the ECLS-K indicates how well kindergartners and first graders were performing at the end of the 1990s (Denton & West, 2002). The data in Table 3 indicate almost a third (31%) of the entering kindergarten cohort knew letter names and beginning sounds. Another third knew letter names, while the final third could name neither letters or identify beginning letter sounds, at least at the criterion level.

There also are no data to indicate how students’ initial status in kindergarten predicts later status but data on end-of-first-year reading performance show a robust relationship to grade-four reading (Juel, 1988). The distribution on the NAEP at grade four—33% (below-basic), 34% (basic), and 33% (proficient/advanced) (National Center for Education Statistics, 2009)—closely resembles kindergartners’ performances on letter naming/beginning sound matching on school entry—33% (not attaining either), 36% (attaining letter naming), and 31% (attaining both letter naming and beginning sounds). The measure that represents the criterion task for grade one is reading connected text. On this task, 52% of exiting first graders fail to reach criterion. Table 4 summarizes the features of the unique words on the connected-text reading task of the ECLS-K. These features can be compared to those of current texts. The text that 52% of exiting first graders could not read to criterion is at comparable levels to the instructional texts of grade-one entry level in 2000 and kindergarten exit level in 2008.

If over half of the students cannot read the entry-level first-grade text at the end of the year, a reasonable hypothesis is that the approximately half of the students in a cohort spend their first two years of reading instruction (kindergarten and first grade) with texts that are inaccessible. Students in the basic group show varying degrees of success with the text. For well
over half of a cohort, however, the texts that they were given at the beginning of first-grade (and now in kindergarten) are still unreadable at the end of grade one.

Improving the Reader-Text Match for Students

Who Depend on Schools to Become Literate

As these analyses show, there is a clear need for texts that better match the needs of beginning readers. Such texts can be developed, but even the best texts cannot overcome the challenges created by several prominent assumptions that underlie the beginning reading instruction of all current core reading programs—assumptions that run counter to existing theoretical and empirical scholarship. Therefore, before proposing a model for developing improved texts, this article will address each of these assumptions: (a) earlier is better, (b) word repetition is not a factor in learning to read, and (c) one size fits all.

Reading instruction assumptions of current beginning reading programs

Earlier is better. Correlations between early literacy skills, such as letter naming and letter-sound matching, and subsequent reading performance have been the bases for frequent recommendations that the best approach for boosting low-achieving children’s reading achievement is to improve their literacy-related skills of prior to kindergarten entry (e.g., Claessens, Duncan, & Engel, 2009; Lonigan & Shanahan, 2009). The use of correlational data as the justification for the downward acceleration of the reading curriculum raises two questions: (a) Does the earlier acquisition of literacy-related skills translate into higher reading levels for children in subsequent years? And (b) Do the gains from teaching literacy-related skills to low-achieving preschoolers and kindergartners balance human and economic costs of this instruction?

In interpreting data on early childhood status and later performance, a distinction must be
made between relative and absolute performances. First, evidence is strong that children’s performances relative to their peers stay relatively stable over time (Juel, 1988). Relative status to peers, however, does not mean that initially low-performing students have not attained particular standards or that initially higher achieving peers have attained higher standards. Currently, even though two-thirds of an age cohort begins kindergarten with criterion-level letter-naming knowledge (Denton & West, 2002), a significant percentage of these students are not attaining adequate levels as fourth graders on the NAEP (National Center for Education Statistics, 2009). Further, substantial investments in literacy-related instruction of four-year-olds and primary-level students have not resulted in commensurate gains in reading achievement in higher grades (Gamse, Jacob, Horst, Boulay, & Unlu, 2008; Jackson et al., 2007).

Systematic data on the effects of earlier school entry and reading instruction on absolute levels of performance are limited, but two recent studies shed some light on the topic. Over a six-year period, Suggate, Schaughency and Reese (2008) assessed students who had school entry ages (SEA) of either five or seven years on a range of measures of reading and pre-reading skill development. Hierarchical linear modeling accounted for receptive vocabulary, parental income and education, school community affluence, classroom instruction, home literacy environment, reading self-concept, and age. Findings showed that by ages 10 to 11, an initial advantage in reading skills associated with earlier SEA was no longer apparent.

In a second study, Suggate (2009) reanalyzed data from the reading portion of the 2006 Programme for International Student Achievement (PISA) study. Performances of 15-year-old students across 55 countries were examined, controlling for social and economic differences and, Suggate claims, linguistic differences were ameliorated by the large sample size. Students’ SEA was not a significant predictor of countries’ mean reading achievement, but students in countries
with an earlier SEA tended to have larger variance in reading achievement at age 15. In speculating on the widening achievement gap in countries with earlier SEA, Suggate questions whether there are psychological or developmental costs to early literacy instruction. That is, does earlier initiation into literacy instruction for students who are not prepared for this instruction lead to lower levels of engagement and performance in reading?

Because of a lack of evidence on the psychological effects of earlier instruction, Suggate’s (2009) question cannot be answered in American school settings. Reflecting on Chall’s (1983) reading model, however, the fit between current forms of early literacy instruction and children’s developmental capabilities can be questioned. Chall identified the first of six stages of reading development as Stage 0 and viewed this stage to precede formal reading instruction. Chall’s Stage 0 consisted of the typical activities then associated with kindergarten, such as listening to books read aloud, playing with puppets, and scribbling and drawing. The instruction for kindergarteners in the teacher’s manuals of current copyrights of core reading programs (Afflerbach et al., 2008; Bear et al., 2008; Beck et al., 2008) is a far cry from those Stage 0 activities and are typical of the formal reading instruction Chall described as Stage 1 where children are expected to read books and are taught lessons about words and parts of words. Kindergarten texts in current programs have been written to emphasize the target grapheme-phoneme correspondence of the lesson, modeled after the Open Court Reading program, which is illustrated by the text: “Tim spins. Tim dips. Tim has a hat. Tim spins his hat. Tim hits a pit...and sits. Tim is sad.“ (Adams et al., 2000).

Such texts are likely fathomable for children with the hundreds of hours of Stage 0 experiences that Adams (1990) describes as typical in the preschool lives of medium/high-income children. For the children who have not manipulated magnetic letters, listened to
hundreds of bedtime stories, and scribbled messages, however, texts with *pits, dips, and spins* may be challenging to comprehend. Young children’s attitudes and beliefs can be difficult to capture with traditional interviews and surveys, but MacKinnon (1959), after observing hundreds of hours in first-grade classrooms, described young children with little conventional literacy on school entry as confused and bewildered when their initial reading experiences consisted of didactic phonics exercises. MacKinnon did not describe how long this confusion lasted nor did he conduct long-term observations or interviews to determine long-term effects. The students in MacKinnon’s sample were first graders. Immersion into Stage 1 instruction with heavy diets of decodable texts and without Stage 0 experiences can hardly be the way to initiate young children into successful, lifelong literacy patterns.

*Word repetition is not a factor in learning to read.* After decades of working with beginning reading anthologies that feature the tedious repetition of words, it is understandable that teachers were glad to be rid of such rigidly controlled texts. Ill-conceived implementation, however, does not negate the underlying principle that novices require at least a modicum of repetition in the content they are being asked to learn. Programs for the English learning of individuals who are already proficient literacy users in their native languages emphasizes repetition of English vocabulary (Nation, 2004). Similarly, research on deliberate practice of expertise in domains that range from surgery to computer programming underscore the need for repetition of content (Ericsson & Lehmann, 1996). Yet, in texts used for reading acquisition, little attention is paid to repetition of vocabulary. The issue of repetition in learning to read does not require debate. If repetition of content was not a factor in learning to read, 100% of exiting first graders would be able to read the simple text of the ECLS-K, not 48%. The question is not whether repetition makes a difference but, rather what content to repeat for learners at differing
points in the developmental progression.

The original algorithm for repetition of words in beginning reading texts was based on Gates’s (1930) research. Gates concluded that, for first graders of average ability, the optimal number of repetitions was 30. Consequently, beginning reading programs of the time (Elston & Gray, 1930; Gates & Huber, 1930) required that every word be repeated 30 times—a pattern that remained in place for almost 60 years. Unexamined was the generalizability of the finding as a function of the words that Gates had studied—the highest frequency words of written English. As evident in the 10 most frequent words— the, of, and, to, a, in, is, that, it, was—many of these words are highly abstract, and many also have variant vowel patterns.

Even for novices learning their first handful of words, all words do not require the same amount of repetition to be memorable. Children’s interest and personal meaningfulness account for the fact that names (e.g., their own names, Mommy, Daddy) are among the first words they remember (Hiebert, 1983). In addition to meaningfulness, factors that influence word learning are concreteness of the word (Sadoski & Paivio, 2000) and the presence of highly frequent grapho-phonemic units (Laxon, Gallagher, & Masterson, 2002; Leslie & Calhoon, 1995; Martinet, Valdois, & Fayol, 2004) and/or morphemes (Carlisle & Katz, 2006). Sufficient research exists to conclude that recognition of words varies as a function of word features (Landauer, Kireyev, & Panaccione, 2011).

Repetition of content is also a function of where learners are in the developmental progression. Human beings have the ability to generalize, once a critical amount of knowledge has been gained. Share (1995) describes this phenomenon as self-teaching, where readers use existing knowledge to acquire new vocabulary. When almost half of the words occur a single time in beginning reading texts—even with phoneme-graphemes that have been taught in
lessons—young children may have a difficult time employing a self-teaching strategy.

Attempts to establish rules for repetition that generalize across all words (Gates, 1930) or word parts (e.g., California Board of Education, 2006; Texas Education Agency, 1997) fail to recognize that different features of words, including component parts such as the individual phoneme-grapheme (Thompson, Cottrell, & Fletcher-Flinn, 1996), influence the rate of learning and retention, and that readers’ existing knowledge also is a powerful influence on rate and retention. A priority of federal research programs needs to be on increasing information on levels of repetition required for different types of words and for students at different points in the learning progression.

One size fits all. The evidence for the one-size-fits-all assumption found in reading programs comes from teacher reports and newspaper accounts (e.g., Davis, 2011), not from the archival literature. Esquith (2004), a teacher in inner-city Los Angeles, writes how the mandate extended to his fifth-grade class: “every child at a particular grade level is supposed to finish each unit at the same pace. A child in the fifth grade at one school should be on the same unit as a fifth-grader at a school across town.” (p. 38).

The origins of this assumption are unclear. One source may lie in perceptions of equity. That is, giving appropriate materials to initially lower performing students reflects lower expectations. A related perspective is that, if students are pushed, they will rise to the occasion and read harder material. This latter perspective is evident in the recommendation of stretch texts within the Common Core State Standards (CCSS Initiative, 2010). The existing research, however, points in the opposite direction. When texts are too difficult for students, their comprehension suffers (Mesmer & Hiebert, 2011). Their attention to the task declines as well (Gambrell, Wilson, & Gantt, 1981). For beginning readers who don’t have the self-teaching
strategies required to generalize knowledge (Share, 1995), instructional time spent in texts where few words are known can be particularly futile (Johnston, 2000).

Esquith (2004) raises another possibility for the “all on the same page” mandate: the lack of preparation and knowledge among beginning teachers. Given that the most inexperienced teachers typically teach children in inner-city schools, this edict is often enforced in these settings. After describing the recognition of administrators that one size does not fit all, Esquith states: “these officials have pointed out to me that for many of our young teachers, a carefully scripted series of reading lessons will help, as the teachers are new and don’t know what to do.” (Esquith, 2004, p. 38).

*Developing an Integrated Model of Text*

Typically, a review of research ends with a call for more research, and this one is no exception, calling for research that can lead to the development of a model of text that integrates the word-level features in the three types of beginning reading texts of the past 50 years.

As this article has shown, each change in texts over the years has focused on a single text feature. When each text type was introduced as an innovative reform, it was viewed as a competitor for the existing or previous text types. In actuality, all three of these elements, together with elements related to syntax, discourse/genre, and program, require consideration in a comprehensive model of text for beginning readers (Mesmer, Cunningham, & Hiebert, 2010). An overview of the word-level component of this comprehensive model illustrates both the complexity of beginning texts and also the integration of perspectives that have often been treated separately in the design of texts.

In vocabulary research, words are regarded as multi-dimensional (Nagy & Hiebert, 2010). Additional variables are considered as the vocabularies become specific to content areas
(e.g., conceptual complexity), but the primary features of frequency, letter/syllabic structure, and familiarity are typically used to determine the ease with which students will learn vocabulary (see, e.g., Landauer et al., 2011). Not surprisingly in that these variables represent consistent and critical dimensions of all words, the variables of vocabulary research are the same as the variables in the word recognition literature. Within the study of vocabulary at higher levels, however, these variables are not viewed to be competing with one another as the variable that requires attention to the exclusion of other variables.

Indeed, evidence shows that all three features of words—meaningfulness, frequency, and grapho-phonemic-morphemic structures—influence the speed with which words are recognized and the number of repetitions required to learn a word and that these variables interact with one another (Laxon et al., 2002; Leslie & Calhoon, 1995; Martinet et al., 2004). Thus, meaningful words that occur with at least moderate frequency and that have consistent and common word patterns are learned and remembered more efficaciously than unfamiliar words that are rare and have less consistent and common word patterns (Hargis, Terhaar-Yonkers, Williams, & Reed, 1988). To fit the criterion that Graves, Juel, and Graves (2004) identify for word recognition vocabulary—that a word be in a child’s oral vocabulary—the dimension of familiarity is the first to be addressed in choosing words for texts. When the answer is affirmative, words are filtered through the lenses of frequency and grapho-phonemic-morphemic structure. For example, words such as reality (as in reality show) and idol (as in American Idol) might be familiar to many children currently but these words are neither frequent in written language to give beginning readers leverage nor is the internal structure (i.e., phoneme-graphemes, morphemes, syllables) sufficiently common and consistent to support generalizable knowledge and strategies.

There is both theory (e.g., Ehri, 1991) and empirical evidence (e.g., Jenkins, Peyton,
Sanders, & Vadasy, 2004; Menon & Hiebert, 2005) that serves as the basis for the study of multi-dimensional models of beginning text. However, a richer and more detailed scholarship is needed to ensure that texts are designed for particular periods of reading development. Research initiatives that make the gathering of this scholarship a priority are needed on the part of federal and foundation agencies, if the substantial portion of an American age cohort who now receive inaccessible texts as kindergartners and first graders is to be served.

Conclusion

Over the 50-year period covered in these analyses, the changes in beginning readers and in the texts that initiate them into formal reading have been massive. Over this same period, a seismic shift has also occurred in when children begin formal reading instruction. In brief, as the diversity of students in American schools has increased over the past 50 years, the texts used in beginning reading instruction have gotten increasingly more complex. And, students are expected to read these texts at an even earlier age.

Granted, these changes all present policymakers and textbook publishers with highly complex issues to address. Too often, however, both have responded to the complexity with simplistic solutions that have, in fact, exacerbated the obstacles for the students who are most in need. In the case of the textbook adoptions by the nation’s two largest states, the solutions have often gone from one extreme to another—with the solution of one decade negated by the solution of the subsequent decade. Even more troubling than this vacillation between extremes is the solution of beginning formal reading instruction in kindergarten—or preschool (Lonigan & Shanahan, 2009), thus replacing a period of rich literacy experiences, the stage 0 that Chall (1983) identified, with earlier and earlier immersion in code-dominated instruction. For children who depend on schools to acquire literacy, this eliminates a crucial induction period.
The dramatic changes in philosophy of reading programs simply fail to address the major underlying issues. English is a complicated amalgam of languages. A simplistic solution to the teaching the English language code will not work. Words are multi-dimensional, not either “high-frequency” or “phonetically regular” or meaningful. There is no word in English that lacks an alphabetic base. Words differ in their frequency and in the frequency of their component parts. And, at their core, words convey meaning. All of these dimensions, as well as the ways in which words form sentences to tell stories or convey information, matter. As long as simplistic solutions to the complexities of written English and of the public school classroom are mandated, we can anticipate that an increasing number of American students will be left with inadequate literacy levels for the digital-global age.
References


Mahwah, NJ: Lawrence Erlbaum.


Texas Education Agency (1990). *Proclamation of the State Board of Education advertising for bids on textbooks*. Austin, TX: Author.


Texas Education Agency (2005). *Proclamation of the State Board of Education advertising for bids on textbooks*. Austin, TX: Author.


U.S. Census Bureau, Population Division, Education & Social Stratification Branch, Fall school enrollment of the civilian noninstitutional population 5 to 34 years old, by level of school, age, color, and sex, for the United States: October 1961, Table 5. Retrieved on April 13, 2011, [http://www.census.gov/population/www/socdemo/school/p20-117.html](http://www.census.gov/population/www/socdemo/school/p20-117.html)


U.S. Department of Education, Office for Civil Rights, 1986 State Summaries of Elementary and Secondary School Civil Rights Survey; and National Center for Education Statistics, Common Core of Data survey. (This table was prepared May 2000.) Table 44. Retrieved on April 13, 2011, [http://nces.ed.gov/programs/digest/d00/dt044.asp](http://nces.ed.gov/programs/digest/d00/dt044.asp)


http://nces.ed.gov/programs/digest/d08/tables/dt08_041.asp


### Table 1

Racial/Ethnic Distributions of American Children/Youth

<table>
<thead>
<tr>
<th></th>
<th>American Indian/Alaska Native</th>
<th>Asian/Pacific Islander</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>White</td>
<td>Black</td>
</tr>
<tr>
<td>1960</td>
<td>86</td>
<td>14</td>
</tr>
<tr>
<td>1970</td>
<td>85</td>
<td>15</td>
</tr>
<tr>
<td>1980</td>
<td>70</td>
<td>16</td>
</tr>
<tr>
<td>1990</td>
<td>62.9</td>
<td>17.1</td>
</tr>
<tr>
<td>2003</td>
<td>60.3</td>
<td>16.8</td>
</tr>
<tr>
<td>2006</td>
<td>56.5</td>
<td>17.1</td>
</tr>
</tbody>
</table>

aData in this table and Figure 1 come from reports of the U.S. Census Bureau (1961, 1970, 2011) and reports of the U.S. Department of Education (2000, 2003-2004, 2008).
Table 2

ECLS (Birth Cohort of 2001): Contrasts in Kindergarten-Entry Literacy Performances

<table>
<thead>
<tr>
<th>Overall Factors</th>
<th>Specific Comparison</th>
<th>Difference (Scale Scores)</th>
<th>Proportion of standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Girls-Boys</td>
<td>44.9-43 = 1.9</td>
<td>.13</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>White-Black</td>
<td>46.4 - 41.1 = 5.3</td>
<td>.37</td>
</tr>
<tr>
<td></td>
<td>White-Hispanic</td>
<td>46.4 – 39.4 = 7</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>White-Asian</td>
<td>46.4 – 51.9 = -5.5</td>
<td>-.39</td>
</tr>
<tr>
<td></td>
<td>White-American Indian/Alaska Native</td>
<td>46.4 – 37.1 = 9.3</td>
<td>.65</td>
</tr>
<tr>
<td></td>
<td>White – All ethnicities</td>
<td>46.4 – 44.2 = 2.2</td>
<td>.15</td>
</tr>
<tr>
<td>Poverty status</td>
<td>At or above - below</td>
<td>46.0 – 37.3 = 8.7</td>
<td>.61</td>
</tr>
<tr>
<td>Home Language</td>
<td>English – Non-English</td>
<td>44.8-40.4 = 4.4</td>
<td>.31</td>
</tr>
<tr>
<td>Family type</td>
<td>Two parent – single parent</td>
<td>45.2 – 40.1</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>Two parent - other</td>
<td>45.2 – 38.9</td>
<td>.44</td>
</tr>
<tr>
<td>Age</td>
<td>&lt;5 – 5/5.5</td>
<td>35 – 39.7 = -4.7</td>
<td>-.33</td>
</tr>
<tr>
<td></td>
<td>&lt;5 - 5.5/6</td>
<td>35 – 45.8 = -10.8</td>
<td>-.76</td>
</tr>
<tr>
<td></td>
<td>&lt;5 – 6+</td>
<td>35 – 50.4 = -15.4</td>
<td>-1.08</td>
</tr>
<tr>
<td>Primary care</td>
<td>Nonparental - parental</td>
<td>44.8 – 39.8</td>
<td>.35</td>
</tr>
<tr>
<td>(year prior to kindergarten entry)</td>
<td>Center: non Head Start – Head Start</td>
<td>47.2 – 40.3</td>
<td>.49</td>
</tr>
<tr>
<td></td>
<td>Home: nonrelative – relative</td>
<td>45.4 – 39.5</td>
<td>.42</td>
</tr>
</tbody>
</table>
Table 3

Literacy performances of students in ECLS-K Cohort (1998) at Fall & Spring of Kindergarten and Spring of First Grade

<table>
<thead>
<tr>
<th>Literacy Outcome</th>
<th>Fall (Kindergarten)</th>
<th>Spring (Kindergarten)</th>
<th>Spring (First Grade)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter recognition</td>
<td>67</td>
<td>95</td>
<td>100</td>
</tr>
<tr>
<td>Beginning sounds</td>
<td>31</td>
<td>74</td>
<td>98</td>
</tr>
<tr>
<td>Ending sounds</td>
<td>18</td>
<td>54</td>
<td>94</td>
</tr>
<tr>
<td>Sight words</td>
<td>3</td>
<td>14</td>
<td>83</td>
</tr>
<tr>
<td>Words in context</td>
<td>1</td>
<td>4</td>
<td>48</td>
</tr>
</tbody>
</table>
Table 4

Features of Beginning Reading Texts From Four Copyrights of a Core Reading Program

<table>
<thead>
<tr>
<th>Copyright</th>
<th>Program Level</th>
<th># of words per text</th>
<th>New, Unique Words per 100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(X)</td>
</tr>
<tr>
<td>1962</td>
<td>Entry Gr. 1</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Exit Gr. 1</td>
<td>378</td>
<td>8</td>
</tr>
<tr>
<td>1993</td>
<td>Entry Gr. 1</td>
<td>79</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>Exit Gr. 1</td>
<td>385</td>
<td>20</td>
</tr>
<tr>
<td>2000</td>
<td>Entry Gr. 1</td>
<td>83</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Exit Gr. 1</td>
<td>334</td>
<td>19</td>
</tr>
<tr>
<td>2008</td>
<td>Entry K</td>
<td>33</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Exit K</td>
<td>110</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Entry Gr. 1</td>
<td>131</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Exit Gr. 1</td>
<td>353</td>
<td>23</td>
</tr>
<tr>
<td>ECLS</td>
<td></td>
<td></td>
<td>75</td>
</tr>
</tbody>
</table>
Figure 1

Continents of Origin: Foreign-Born Individuals in U.S.