Text Matters in Learning to Read

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

Elfrieda H. Hiebert

Text matters in learning to read

Hiebert evaluates several types of texts used for beginning reading instruction by examining the tasks each poses for young readers.

Many varied opportunities to interact with numerous selections of high-quality literature are critical to reading development. But should high-quality literature be the sole material in which children apply their reading knowledge at the earliest stages of reading? Only in the last decade have trade books been the primary material for beginning reading instruction. Until the past decade, the texts of beginning reading instruction had controlled vocabulary. For much of the 20th century, the basis of this control was the frequency of words in written English. At various periods in U.S. education such as the present, texts consisting of phonetically regular words have been proposed as an antidote to the difficulties children experience in learning to read (Flesch, 1957; Grossen, 1997).

In this article, I will examine texts based on high-frequency and phonetically regular words as well as the trade books of current literature-based reading programs. I will consider each type of text by examining the task it poses for beginning readers. What does a beginning reader need to know about written English to be successful with a particular type of text? What will a beginning reader learn about text if consistently presented with a particular type of text? From a task perspective, consistent reading of particular types of texts can be likened to a diet where children eat particular food groups but not others (Fisher & Hiebert, 1990). Through experiences with particular texts, children may be acquiring some nutrients (or skills) and not others. This article addresses the diets provided to beginning readers by different instructional texts. To paraphrase Allington (1994), the three sections

of the article deal with (a) the texts we had, (b) the texts we have, and (c) the texts we need.

The focus of this article on the texts for beginning readers needs to be underscored. Once students have acquired basic word recognition knowledge, selections by students, themes, and contemporary and classic canons should be the basis for choosing texts. But at the very earliest stages of reading acquisition—particularly with students who are introduced to book reading in school—careful attention needs to be paid to the texts of instruction. The texts of instruction are by no means the only exposure that children have to books. A classroom environment in which children are brought to high levels of literacy involves many different types of books and book events (Hiebert & Raphael, 1998). But, while composing only a portion of the books of an early reading classroom, the books used to guide children in independent reading are critical and require careful thought.

The texts we had

Whether the cast of characters consisted of Janet and Mark, Dick and Jane, Alice and Jerry, or another dynamic duo, many generations of children were introduced to reading through texts containing the most frequent words in written English. The register of these texts came to be called "primerese," named after the primers and preprimers that were the first components of reading programs. An example of this type of text comes from the first two pages of the first text of the first preprimer of a popular textbook series of the 1980s, "We Can Go" (Durr et al., 1986).

> I can go. Can you go? Help. Help. I can not go.
> I will help you. You can not help. I can not go.

As the summary of this text in Table 1 shows, it is composed of eight unique words, all of them high-frequency words. This type of text can be traced to the 1930s (Elson & Gray, 1930) when Thorndike's (1903) laws of learning were first applied to beginning reading materials. Well over 50 years later, the law of readiness, which dictated that new content needed to be carefully sequenced with familiar content, was evident in preprimer texts. For example, "It Will Not Go" (Durr et al., 1986), the passage subsequent to "We Can Go," consists of the original eight words and five additional

ones. The law of exercise, which required that new content be repeated, was evident in the 1980s texts in that each of the eight original words in "We Can Go" had been repeated between 16 and 29 times by the end of the last selection of the first preprimer. The need for identical elements, Thorndike's third law, was supported by the ratio of unique to total words, which remained the same across the preprimer and primer passages: one out of every seven or eight words was unique. Because these words convey a story—one child assisting another in learning to roller-skate in "We Can Go"—finishing the text was thought to reinforce successful reading (the law of reinforcement).

Units other than high-frequency words could have been selected as the "stimulus" for learning. The particular choice of high-frequency words emanated from Thorndike's interest in word frequency and research conducted by Gestalt psychologists in the previous decades on the learning of wholes. Although phonetically regular words, rather than high-frequency words, were advocated as the unit of learning from the inception of primerese, this perspective did not gain popularity until the 1950s (Flesch, 1957). The late 1950s and early 1960s saw a number of beginning reading programs where phonetically regular words were the basis for texts (Bloomfield & Barnhart, 1961; Rasmussen & Goldberg, 1964). Here are the first two pages of the first text, "Dad" (Cassidy, Roettger, & Wixson, 1987, p. 15–16), of the first preprimer of a reading program that was advertised as a "phonics" series:

> Dad ran. Ann ran. Dad and Ann ran.
> Dan ran. Nan ran. Dan and Nan ran.

Such texts using phonetically regular words were derived from the same underlying learning theory as high-frequency text. But the target unit consisted of phonetically regular words such as *ran* in "Dad." The unit was not the rime—that is, the vowel-consonant pattern or phonogram. The words *man* and *pan* appeared later in the first stage but *can, fan, tan,* and *van* did not appear in the preprimers of this series. Such phonetically regular texts never came to dominate beginning reading instruction to the degree of programs based on high-frequency words. During every wave of reading reform, however, phonetically regular text has been and continues to be proposed as a

Table 1
Characteristics of texts

Words with rimes[1]			High-frequency[2]		High-content		Other		Entire text (entire unit)		
Unique	Exemplars	Occurrences	Unique	Occurrences	Unique	Occurrences	Unique	Occurrences	Unique	Total	Word density
High-frequency text: "We Can Go"											
			we	3			not	4	8	55	1:7
			can	13					(45)	927	1:21
			go	9							
			I	8							
			you	8							
			help	7							
			will	3							
Phonetically regular text: "Dad;" "Run, Run;" "Adam"[3]											
ad	1	7	and	14					11	89	1:8
an/ann	4	35	to	2					(95)	933	1:10
um	1	10									
am	2	4									
un	1	17									
Literature-based anthology: Who Is Tapping at My Window?[4]											
at	2	2	the	14	bear	1	not	14	22	84	1:4
og	2	2	is	2	pony	1	who	2	(164)	701	1:4
en	2	2	I	14	cony	1	said	14			
ox	2	2	at	2	hare	1	tapping	2			
					window	1	my	2			
							it's	14			
Little books: A Toy Box											
ox	1	3	a	8	toy	2			12	22	1:2
uck	1	1	in	1	space	1			(320)	1,228	1:4
ip	1	1	the	1	ball	1					
ack	1	1			doll	1					
					plane	1					
Dr. Seuss: Green Eggs and Ham[4]											
at	1	3	I	19	green	4	not	12	23	126	1:5
am	3	16	do	12	eggs	4	like	17			
			and	3	house	2	here	3			
			you	4	mouse	2	there	3			
			a	4			anywhere	2			
			in	2			or	3			
							them	12			
							with	2			
							would	5			
Multiple-criteria text: "That Fat Cat"											
an	5	11	the	6			not	1	13	66	1:5
at	6	48									

[1]Rimes with a short vowel are the focus.
[2]High-frequency words are the 25 most frequent words according to Carroll et al. (1971).
[3]Three texts were included in the analysis to achieve seven pages of text.
[4]First seven pages of this text have been analyzed in order to provide equivalent text samples.

primary solution for reading problems (e.g., Flesch, 1957; Grossen, 1997).

Although differing in the criterion for words—high-frequency or phonetically regular—instructional texts for generations of schoolchildren were based on features of words. Perspectives such as those represented by cognitive science, reader response, and sociocultural frameworks drew attention to the influence on beginning reading acquisition of these texts as a whole—their content, text structure, and illustrations. The task that these texts pose for beginning readers must be viewed as a function of the text as a whole, not simply the features of words. Even when considered from the vantage point of features of words, the task presented by high-frequency and phonetically regular text may present a challenge for beginning readers.

Features of words. In the sense that a beginning reader's attention is drawn to the individual word, the task with texts of high-frequency and phonetically regular words was similar. To be successful in reading a text such as "Dad," beginning readers had to recognize words with the *an/Ann* rime. To be successful in reading a text such as "We Can Go," beginning readers had to recognize a core group of high-frequency words. But if children are prone to generalize beyond the particular words in their texts or if teachers are inclined to guide children in generalizing from the words in their texts, the texts provide different opportunities. For both beginning reader and teacher, generalizations about consistencies in the graphophonic system of English are easier with "Dad" than with "We Can Go." Because many high-frequency words have unique letter-sound relationships, it is difficult to generalize beyond specific words. Although a core group of high-frequency words accounts for a large percentage of the words in texts, these words account for only a small percentage of the unique words in texts. Half of the words in texts of third grade and beyond can be accounted for by 109 unique words, but the remainder is made up of about 80,000 unique words (Adams, 1990).

In focusing on high-frequency words, children's attention is diverted from the common and consistent patterns in English—patterns that occur in many words that children must be able to read well. Since even irregular words employ the alphabetic principle, children presumably come to understand the alphabetic character of English as they come to recognize these words. But acquiring this information without guidance about what is common and consistent in written English can be an arduous, haphazard process that, for some children, occurs so erratically that meaningful reading is impossible.

Despite the decades that this perspective held sway, confirmation that the best way to learn high-frequency words was in the context of sentences composed only of high-frequency words was never strong. As the behaviorist stronghold loosened in U.S. psychology and the texts for beginning readers were considered through the lenses of cognitive psychology and linguistics, the obstacles presented by these texts for beginning reading acquisition were understood. As cognitive psychologists studied reading processes, they found that successful readers identified single words quickly and that beginning and poor readers did better when words were in the contexts of sentences or phrases (Lesgold, Resnick, & Hammond, 1985). Successful readers learn to attend to the orthographic features of words, while poor readers continue to require the syntactic and semantic cues of a sentence or phrase to recognize a word. To develop facility in recognizing high-frequency words, beginning readers benefit from occasions where they can study the features of particular words—what distinguishes *here* from *have* (Adams, 1990). Such focused attention is difficult to develop while reading a text, even when the text is made up only of high-frequency words.

The task posed by the phonetically regular texts has not been analyzed in the same fashion as the task of the high-frequency texts. The information on phonetically regular texts comes from program evaluations where children's achievement in programs with phonetically regular text are compared to those with mainstream or basal text series that highlight high-frequency words at the early stages. These treatments rarely consider components other than texts, even though programs typically involve much more than texts. Nor do they analyze how teachers supplement their texts with other materials, including phonics worksheets in mainstream textbook programs

or literature in phonics programs, although first-grade teachers' adaptations are typically extensive (Pressley, Rankin, & Yokoi, 1996). Despite the unanswered questions regarding teachers' adaptations, the findings of the First-Grade Studies were consistent and compelling: Children in programs that emphasized the alphabetic nature of written English had an advantage over children in programs where other features, such as high-frequency words, were emphasized (Bond & Dykstra, 1967).

Despite consistent conclusions such as these, many educators were resistant to phonetically regular text that typically consisted of storylines such as "Dan ran to the fan. Dad had to fan Nan. Dad had to fan Dan" (Rasmussen & Goldberg, 1964). The prominent text used in beginning reading continued to be high-frequency texts, but changes were made to the teachers' editions. For example, the high-frequency word *and* might be suggested as the basis for a lesson on the short /a/, with teachers listing *hand, sand, land,* and *band* on the chalkboard (Chall, 1967/1982). This lesson would not be accompanied by a story about a child writing a message with his or her hand in the sand, at least not with the words *hand* and *sand* stated explicitly.

Phonics instruction disconnected from the texts that children read contributes little to children's use of phonics strategies in recognizing words. Juel and Roper/Schneider (1985) compared two groups of beginning readers who received the same kind of phonics instruction but who read from different books during reading periods: One group read high-frequency texts and the other group read phonetically regular texts. The children who read the phonetically regular texts used letter-sound information beyond the initial letter of a word when confronted with unknown words to a greater extent than children who read the high-frequency texts. They sustained this strategy after the first 6 months of first grade when their texts became less phonetically regular.

Interest in highly decodable texts has been bolstered recently by a program evaluation that compared children's learning in a phonics-based series with children's learning in several versions of literature-based programs (Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998). Even though comprehension performances of students did not differ significantly,

differences on particular measures of word recognition have led to recommendations that highly decodable text be required for beginning reading instruction (Grossen, 1997). Beck (1997), for example, states that "70 to 80 percent decodable would be reliable enough for children to refine their knowledge of the spelling-to-speech mapping system, while 30 to 50 percent is not enough" (p. 17). Such recommendations as well as the percentages of phonetically regular words in published programs represent an "educated conclusion" (Beck, 1997), with Beck observing, "Studies could be done to identify an optimal range" (p. 17). There can be little doubt that information about consistent and common letter-sound patterns is needed for children to learn to read efficiently. There also can be little doubt that there should be opportunities to apply in text the information that is taught and practiced in teachers' lessons. But, beyond these conclusions, there are numerous questions about texts that support the acquisition of a metacognitive stance toward the linguistic systems of written English.

An example of a topic requiring study is the unit of information that needs to be held constant or varied within and across texts. Texts such as "Dad" were based on the assumption that children acquired the alphabetic principle in incremental, carefully segmented steps. Studies from Project Literacy suggest that emphasis on only one pattern at a time may discourage beginning readers from developing a set for the diversity within written English (Gibson & Levin, 1975), a term that was replaced in the 1980s by metacognitive stance. Moreover, exposure to numerous instances of a pattern such as *man, can, van,* and *tan* rather than the repetition of a single instance such as *ran* in "Dad" develops a disposition to apply knowledge of phonics to new words (Juel & Solso, 1981). As these examples show, many issues remain about what linguistic units should be featured in texts for beginning readers.

Features of text. Perspectives from cognitive psychology and linguistics raised questions about features of texts that went beyond the individual word. First, knowledge of children's language acquisition was applied to children's reading acquisition (Goodman, 1968). Children's facility with the syntactic and semantic systems of their spoken language

(Goodman, 1968) is invaluable in learning to read because it allows them to draw on the systems that oral and written language share to figure out what is new about written language—the alphabetic representation of spoken words. Primerese, it was argued, prevented children from drawing on this knowledge. Typical conversations of children learning to roller-skate do not consist of "I can go. Can you go?" or with a child who has fallen down saying, "Help. Help. I can not go." Children are stymied in applying what they know about language when they read such texts.

Researchers demonstrated that texts that used high-content words rather than high-frequency words and varied sentence structure could be easier to read than primerese (e.g., Brennan, Bridge, & Winograd, 1986). For example, "We can" from "We Can Go" might be transformed to "I can skate!" None of these studies, however, examined the effects of revised texts on the reading acquisition of children during the first 6 months of instruction. The youngest children in the studies were second graders (Brennan et al., 1986). While there was little clarification of the degree of control needed by beginning readers, these studies were critical in turning the attention of teachers and teacher educators to the impediments created by primerese.

Schema theory was also used to show how primerese obfuscated the task of reading for beginners (Anderson & Pearson, 1984). When reading a text about roller-skating, children bring a schema or conceptual knowledge about learning to roller-skate. In a text about learning to roller-skate, children would expect to see words such as *roller-skate, fall,* and *helmet.* When the text uses none of this language but relies on words such as *can* and *help,* children are confused about the task of reading. Reading becomes a process of figuring out an alien language that does not connect to children's experiences. Again, children are unable to draw on what they know, in this case, concepts about their worlds.

To summarize, simplifying the text to the lowest denominator of high-frequency words did not facilitate the task of learning to read in the manner that the generation of educational psychologists who advocated this type of text believed. While phonetically regular text is presumed to facilitate acquisition of word

recognition skills better than high-frequency text, numerous questions remain about the kind and amount of phonics information that beginning readers need and the effects of a diet of phonetically regular texts on children's comprehension of and engagement with text.

Texts that used high-content words rather than high-frequency words and varied sentence structure could be easier to read than primerese.

The texts we have

As evidence accumulated on the impediments created by high-frequency texts for young readers, educational policies followed. The theme became "real" literature—text in which vocabulary was not limited to either high-frequency or phonetically regular words. The first literature-based beginning reading programs were created in the late 1980s for selection in California, which mandated literature programs (California English/Language Arts Committee, 1987). By 1993, the change had occurred in all of the major textbook companies. While primerese was prominent in the beginning reading components of Texas-approved textbooks in 1986, an analysis by Hoffman et al. (1994) showed that the texts on the list of approved programs in Texas in 1993 consisted almost entirely of literature.

A visit to many beginning reading classrooms will also show "little book" programs in daily use for reading instruction. These programs consist of many books, short in length (8 to 24 pages rather than the usual 48 or more pages of trade books), presented in a series of levels. Advertisements for the primary reading programs that were adopted in California in its most recent adoption cycle, 1997, indicate that literature-based programs now include little books as well as five or six anthologies for first grade. Even though the function of these little books in literature-based programs awaits analysis, little books are used for beginning reading instruction in

many classrooms and demand the same attention as trade books if the current task for beginning readers is to be understood.

Trade books. The state textbook guidelines in California and Texas (California English/Language Arts Committee, 1987; Texas Education Agency, 1990) called for the elimination of contrived texts and the use of text of literary merit in reading textbooks. A particular genre of literature quickly dominated the early reading components of the literature-based textbook programs—predictable texts. Books that fall into the predictable text genre are characterized by the repetition of a syntactic unit that can range from a phrase to a group of sentences. While this text structure is evident in nursery rhymes and textbooks of a century ago (Stickney, 1885), its recent popularity stems from the publication of Martin's (1967) *Brown Bear, Brown Bear, What Do You See?* and of Martin and Brogan's (1971) philosophy that such texts permit children's successful participation as readers from the start. The first two pages from the first text, *Who Is Tapping at My Window?* (Deming, 1988), of a literature-based textbook program (Pikulski et al., 1993) illustrate a predictable text:

> Who is tapping at my window?
> It's not I, said the cat.
> It's not I, said the rat.

As summarized in Table 1, there are 22 unique words in the first seven pages of this text, with 5 of these words repeated in each episode: "'It's not I', said the ____." Hoffman et al. (1994) found that the number of unique words had increased substantially in first-grade programs from 1986/87 (controlled, high-frequency vocabulary) to 1993 (literature-based programs). They also found that the number of total words had decreased. That is, children were seeing more words, and they were seeing them less frequently.

While all literature for young children does not use a predictable structure, one distinguishing characteristic of literature is the prominence of illustrations (Cullinan & Galda, 1994). The presence of engaging illustrations was a feature that characterized the texts of the 1993 from the 1986 reading programs that appeared on the Texas approved list (Hoffman et al., 1994).

Little books. Little books refer to relatively short texts that are published for classroom reading programs, initially by publishers in Australia and New Zealand, but increasingly as part of mainstream U.S. reading programs. The first two pages of text from the first book, *A Toy Box* (Literacy 2000, 1988), of one little book program consist of two phrases:

> A truck.
> A space-ship.

In that there are five stages to this program that are aimed at first grade, the first stage could be viewed as equivalent to the first anthology of the literature-based program or the preprimer in textbook programs of the past. This first stage of little books is, in turn, divided into five levels, each progressively more difficult than the next. In *A Toy Box*, each phrase begins with the word *A*, followed by a high-content word that names a component of a toy box. The remaining seven texts in the first level use the same format: a phrase or sentence where items in a category are enumerated, such as zoo or farm animals. The illustrations in little books are as salient as those in literature-based texts, although not the products of currently known illustrators.

Behaviorists' solitary focus on words as the basis for the text in beginning reading instruction was a problem as cognitive psychologists and linguists demonstrated. Now the tide has turned. Within the current schemes for choosing texts for beginning readers, the most prominent of which comes from Reading Recovery (Fountas & Pinnell, 1996; Peterson, 1991), characteristics of the naturalness of the language, a close picture-text match, and the predictability of text structure are emphasized. I attend to the task of reading acquisition posed by text features first, followed by a discussion of word features.

Features of text. From the perspective of the shared book experience (Holdaway, 1979; Martin & Brogan, 1971), young children are able to participate as readers from their initiation into reading instruction when they have the scaffolding of the predictable text and an adult to introduce them to the text. But what do we know about predictable texts as a scaffold for learning to recognize many words independently? The research for predictable texts in developing independent reading was limited (e.g., Bridge, Winograd, & Haley, 1983), and even these few studies were narrow in duration and scope. Bridge et al. (1983) reported that

children learned a group of high-frequency words with predictable books as well as children who participated in high-frequency text lessons. But the duration of the study was short. Further, particular activities were done with the predictable books such as matching of phrase and word cards that were not done with the high-frequency texts, which may explain differences in children's learning. Additional work by Ehri and Sweet (1991) indicates that children's attention to individual words in predictable sentences requires some degree of proficiency in word recognition. New words may be learned more quickly when they are separate from the text than in the context of predictable text (Johnston, 1998).

The illustrations are closely linked to the predictable text structures in both the literature-based and little book programs in that the word that changes from episode to episode in the predictable structure is represented in the illustration. But the texts of the two programs differ in the ease with which the category that underlies the predictable structure can be identified and the usefulness of the illustrations in identifying words. In *A Toy Box* and the other texts in the first stage of the little book program, a child can name the members of a familiar category with the aid of illustrations. The high-frequency words are few in number and can easily be remembered, especially when the task is one of labeling illustrations such as "A ball" and "A doll," as in *A Toy Box*.

The texts of the first stage of the literature-based program cannot be read simply by labeling illustrations since the predictable units range from 5 to 17 words (Hiebert & Raphael, 1998) and at least some of the representatives of categories are not familiar such as *wren* and *cony* in *Who Is Tapping at My Window?* (Deming, 1988). In these cases, however, a rhyming pattern is used so that children are aided in figuring out the unfamiliar words (*pony* with *cony, hen* with *wren*). Even the more familiar words, however, may require attention in that *chicken* might be the response to the illustration rather than *hen*.

Although the illustrations of the literature-based program are less useful in some cases because of the unfamiliar vocabulary, both publishers of the literature-based and little book programs have selected or created texts with a close picture-text match, as recommended within the Reading Recovery text selection guidelines for classroom (Fountas & Pinnell, 1996) and tutoring (Peterson, 1991). The reason for this guideline, according to Clay (1985), is that a close picture-text match allows beginning readers to recognize the high-content words that appear infrequently but that are critical to making meaning of the text. In that approximately half of the words that children will see in their texts will appear only once or twice—words such as *spaceship* and *jack-in-the-box* in *A Toy Box*—illustrations can create important scaffolds for beginning readers. Children come to develop independent reading strategies, according to Clay (1985), by using the cross-checking strategy where they test their hypothesis about the word derived from the illustration against the graphic characteristics of the word in the text.

The perspective that scaffolds such as illustrations and predictable structures in texts can allow children to engage in meaningful reading while acquiring fundamental word recognition strategies has been well received by educators. Descriptions of how effective teachers demonstrate to children the appropriate use of these scaffolds and the manner in which children develop in their awareness of graphophonic features and attend less to illustrations have not been forthcoming. On the contrary, some research (e.g., Samuels, 1970) indicates that illustrations act as a distraction for beginning readers. According to Samuels's focal attention theory, prominence of illustrations deters acquisition of automatic word recognition because children can identify words without attending to the graphic features of words. While research on the focal attention theory has not been conducted with the present generation of beginning readers and the highly illustrated books of beginning reading instruction, children who have been reared in a culture dominated by cable television, video, and film are likely to find illustrations as salient as children of earlier generations, if not more so.

On one feature of the texts as a whole, the two programs differ in opportunities for beginning readers: the volume of text. The little book program has 1,225 words in its first stage, spread across 40 texts, while the literature-based program has 701 words across the 6 passages of the first literature anthology for first

grade. Within the literature-based program, teachers' manuals instruct teachers in using a text such as *Who Is Tapping at My Window?* over a week's set of lessons. By contrast, the little book program provides many different texts. While attending to the number of repetitions of words in texts for young children, researchers have never examined how much text children need to receive. Reading many texts rather than a single text per week may be critical for applying word recognition strategies by and sustaining the engagement of beginning readers.

Features of words. Within current schemes of text selection for beginning readers, the only concession to word difficulty is attention to text length (Fountas & Pinnell, 1996; Peterson, 1991). Longer texts presumably are harder texts. But, as can be seen in Table 1, texts that share the same number of pages can vary in the number of words on those pages. The first text in the literature-based program presents the beginning first-grade reader with 22 unique words among its 84 words or a 1:4 word density ratio of unique to total words. While there are fewer unique words in the first text of the little book program, *A Toy Box*, the word density ratio of 1:2 indicates that there is less likelihood that the word will be repeated once the beginning reader has figured it out.

The argument could be made that word density ratios of 1:20 are only necessary when the words are high-frequency words as in the preprimers of the past, but Juel and Roper/ Schneider (1985) reported that number of repetitions of a word predicted children's facility with it in both high-frequency and phonetically regular texts. Did the ratios of the texts of the past represent the opportunities for repetition that children require? If so, the ratios of the present texts—1:2 or 1:4—are substantially discrepant from the ratio of 1:7 of "We Can Go" or the 1:21 of the first preprimer of the high-frequency program (Durr et al., 1986). Repetition of words has not been a primary consideration in the creation or selection of texts, as can be seen in examining more closely the attention given to three types of words in the literature-based and little book texts: (a) high-content words, (b) high-frequency words, and (c) phonetically regular words.

High-content or story-critical words represent the largest group of words within the little books and literature-based programs, accounting for 50% of the unique words in the little book program and 37% in the literature-based program. As well as the argument for emphasizing high-content words in beginning texts because of their picturability, high-content words have been presented as inherently more interesting to young children (Ashton-Warner, 1963). From this perspective, children will remember words such as *dinosaur* and *spaceship* more readily than high-frequency words or phonetically regular words because the high-content words have richer meanings and hold greater interest for young children.

If high-content words provide the foundation for children's word acquisition, particular high-content words would be expected to appear more than once in the first stage of a program. But that is not the case. Only several high-content words appear more than once across the 40 texts of the first stage of the little book program or the six passages that make up the first stage of the literature-based program. Validation that a high-content word can be remembered by beginning readers after seeing the word once in a text is lacking. Even special words that children choose themselves—names of pets or family members—often need to be seen repeatedly for beginning readers to recognize them independently (Ashton-Warner, 1963). It is unlikely that the majority of children will remember words from a category that an adult writer believes is interesting—zoo or farm animals or the participants in the circus—after one occurrence in a text.

The literature-based and little book programs differ in the percentage of high-frequency words: 5% of the literature-based words and 31% of the little book words are among the 25 most frequent words in written English (Carroll, Davies, & Richman, 1971). Despite such differences in percentages, both texts have the same total number of unique high-frequency words: 17 of the 25 most frequent words. The different percentages indicate that the occurrence of these words relative to the other words in the texts varies considerably. One out of every 20 words in the literature-based program is one of the 17 most frequent words, while one out of every 3 words in the little book program is of this type. This percentage of high-frequency words in the little book program does not represent a systematic plan. Two words account for

almost half of the appearances of high-frequency words: *a* and *the*. Their use reflects the labeling structure in the first stage of the little book program where instances of a category such as circus or farm animals are presented with *a* or *the* preceding the noun. Whether beginning readers attend to *a* and *the* when the task involves labeling of illustrations is uncertain. Teachers with a knowledge of word study activities, such as matching of high-frequency words on cards and in books, would have at least some material on which to base instruction. Such activities would be much more difficult with the fewer occurrences of high-frequency words in the literature-based texts.

The pattern of *A Toy Box*, four unique V-C rimes represented within 12 unique words, is typical of the presentation of the phonetically regular words in the first stages of the two programs. Among the 52 words with V-C rimes within the first stage of the little book program, 13 rimes have two or more exemplars (e.g., *cat, pat, rat; dog, frog*) and 18 rimes are represented by one word. The introductory text for the first level of the literature-based program, *Who Is Tapping at My Window?* has four V-C rimes: *at, en, og,* and *ox.* In that there are two words with each rime, this text seems to have been selected for its attention to phonetically regular words. Further, two of these patterns, *at* and *og,* are repeated in at least two of the five subsequent texts of the level. When considered in relation to the other rimes in the first stage—8 other unique rimes that occur in at least two different words and 15 additional rimes occurring in only one word—the consistent and common patterns would be difficult to notice among all of the information provided in the texts. In the first stage of both little book and literature-based programs, the presence of so many different rimes with so few instances of particular rimes is likely an array far too diffuse to attract the beginning reader's attention.
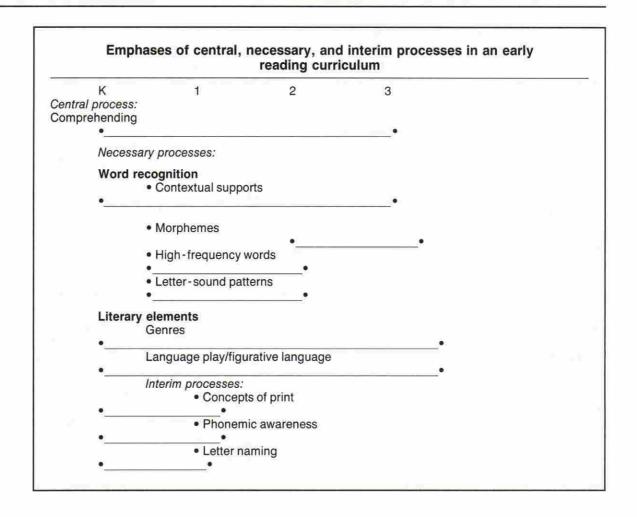
The texts we need

The texts of the present and the texts of the past have been based on a single criterion. A vision of the various processes children need to acquire for successful reading is lacking within both stances. In any perspective on learning, the definition of what needs to be learned is critical. If teachers are to make inroads in the reading acquisition of children in high-poverty schools,

a view of what reading is and its manifestations at various development stages is essential. Only from the vantage point of a model of reading acquisition can we begin to identify the appropriate texts for beginning readers.

The nature of the task for beginning readers. Within all of the standard-setting efforts in the U.S., the beginning reading task has been circumvented by initiating the standard-setting process at Grades 3 or 4. The task for the beginning reader is *not* the same one as the task for the more advanced reader. With Taffy Raphael, I have presented a curriculum (Hiebert & Raphael, 1998; Raphael & Hiebert, 1996) that begins with the central process of comprehending. This central process can be analyzed to finer levels, such as comprehending efferently or aesthetically. Subsumed within the central process of comprehending is the next level of the curriculum—the necessary processes of reading. As the presentation of this curriculum in the Figure shows, the necessary processes of reading vary as readers acquire proficiency. While discussions about literary elements and morphemic (meaning) characteristics of words occur in the beginning reading classroom, attention to these necessary processes is eclipsed by focus on the alphabetic nature of written English or letter-sound relationships, the recognition of frequently occurring function words, and the appropriate uses of graphophonic, syntactic, and semantic contexts in figuring out infrequent but text-critical words. This curriculum places the interim processes of phonemic awareness and letter naming within the necessary processes. For readers who are adept at the necessary processes of word recognition, for example, assessment of the interim processes of phonemic awareness and letter naming is unnecessary.

The view of the task for the beginning reader that emanates from this curriculum is one of developing proficiency with the three necessary processes of applying the alphabetic principle, recognizing high-frequency words, and using the structures of sentences and texts to validate meaning. Although one of these processes may be foregrounded during particular periods of time or within a lesson, attention must be paid to all processes for children to read well.

Identifying the necessary processes is only a first step. The great debate rarely addresses what aspects of phonics or contextual strategies should be taught, but it is at this level that teach-

Emphases of central, necessary, and interim processes in an early reading curriculum

	K	1	2	3

Central process:
Comprehending

Necessary processes:

Word recognition
  • Contextual supports

  • Morphemes

  • High-frequency words
  • Letter-sound patterns

Literary elements
Genres
Language play/figurative language

Interim processes:
  • Concepts of print

  • Phonemic awareness

  • Letter naming

ers have the most questions and that their choices make the greatest difference in children's reading achievement. How many letter-sound correspondences do children need to study before they grasp the alphabetic principle? Is there a point where, after a core group of high-frequency words have been memorized, children will quickly memorize additional words of this type? I will provide an illustration of a curriculum that responds to such questions. My intent is generative rather than prescriptive. I do this in the hope that the presentation of illustrations such as mine will spur educators to share their curricular efforts and reports of children's learning from these efforts.

The curriculum that I describe is a response to the needs of university students involved as America Reads tutors. To aid tutors with the daunting task before them, we have included a simple curriculum in their "tool kit" (Hiebert, Martin, Gillard, & Wixson, 1998). The curriculum includes the most consistent and common phonograms (Wylie & Durrell, 1970), beginning with the short-vowel rimes: *at, an, ap, in, ip, op, ug, ut*. Tutors are guided in selecting texts that provide experiences with these patterns and in initiating associated writing activities where children produce words with these patterns. By conducting conversations around particular patterns, tutors aim to guide children in what is common and consistent about written English.

For the high-frequency words, we have created 5 clusters among the 25 that Carroll et al. (1971) identified as occurring most frequently in written English: (1) *the, am, and, I, was*; (2) *in, is, you, it, that*; (3) *not, are, at, said, they*; (4) *be, of, as, have, this*; and (5) *his, by, one, with, from*. Tutors are advised to look for one or more of the words within a cluster in their choice of books. They are also guided in word study activities with high-frequency words, such as matching activities with cards or dictation of phrases with the high-frequency words.

The third set of strategies within this curriculum involves "monitoring for meaning" or using contextual supports. Without doubt, instruction of this set of strategies is the most difficult for any teacher or tutor working with beginning readers. Using the pragmatic system of written language—the structures of a predictable text, for example, or the illustrations—is the first strategy that children use with texts. If using contextual supports never moves beyond the predictable text and illustrations, children will never be independent readers. At the same time, if children attend only to the words and not to the meaning of text, they will not be proficient readers. Tutors are provided with techniques that encourage children to maintain this fundamental disposition toward text but that direct their attention to the letter-sound correspondences and to high-frequency words. For example, word cards are used to create sentences that a beginning reader needs to read for "sense" or "nonsense."

The nature of text for beginning readers. Texts to support beginning readers' success would give children exposure to the three necessary processes of word recognition. Should features be presented singly as in the exaggerated forms of the texts of the past and present? Should all features be in the same text? Because texts of the latter sort are not available, I will describe the use of single-criterion texts to create multiple-criteria programs. But I will also propose a form of text that exemplifies multiple criteria.

Multiple-criteria programs. A collection of all of the texts, past and present, that have been offered for beginning readers would number in the thousands. Whether the criterion is literary quality, high-frequency words, or phonetically regular words, texts of the past and present for beginning readers have highlighted a particular feature of written English. Teachers need to know how such single-criterion materials can be used to provide optimal experiences for their students.

One option is to use different single-criterion texts, with the aim of providing a comprehensive array of information about written English to beginning readers. For example, one week of lessons might be devoted to application of phonics strategies in texts such as "Dad." The next block of lessons might use little books, such as *A Toy Box*, in order to maintain chil-

dren's attention to the meaningfulness of their reading efforts. Intermittently, a text such as "We Can Go" might be used to expand children's high-frequency word corpus. Many teachers have created programs that draw on several different types of texts. Case studies of such efforts are needed to document the nature of children's reading development in classrooms where knowledgeable teachers have created programs with different balances of text.

Another option with current materials—one that my colleagues and I took in an early intervention program for Title I students—was to sort and sequence little books according to features of word density and phonetic regularity (Hiebert, Colt, Catto, & Gury, 1992). Although our group used little books from numerous programs, I will illustrate the identification and sequencing of texts with the little book program that has been analyzed in this article. Our focus in the intervention, as in the *Tool Kit for Tutors* (Hiebert et al., 1998) that has already been described, was on the most consistent and common rimes (Wylie & Durrell, 1970), not all possible rimes (cf. Foorman et al., 1998). From the various programs that the teachers collected, they were able to identify texts that were particularly appropriate for application of particular rimes and had appropriate word density ratios.

Seeing the word in a text extended the instruction that children received on rimes in the context of their small-group lessons. This instruction revolved around writing rimes because of the opportunities it provides children to test out their hypotheses about letters and sounds concretely. While each text in Table 2 has only one or two exemplars of words with *at*, children would write the word *cat* on an erasable slate and replace the *c* with *m* to create *mat*, with *s* to create *sat*, and so forth. The majority of children in the bottom 40% learned to read well when such an instructional strategy was used consistently over their first-grade year (Hiebert et al., 1992).

To find texts that engage children and give them sufficient experience in applying phonics skills, teachers conducting the intervention had to do considerable juggling. Take, for example, the placement of texts with *at* rimes in the little book program. The *at* rime, a common and consistent rime, does not appear until the second set of texts in the first stage. A text such as *The Pet Parade* (1998) may have two ex-

Table 2
Little books with *at* rime

Text (stage & level of program)	Focus word(s) (number of occurrences)	Number of unique words (word density ratio)
Mud Pie (Stage 1-E)	pat (1)	10 (1:2)
Teeny Tiny Tina (Stage 1-B)	cat (1); rat (1)	14 (1:3)
In My Bed (Stage 1-E)	cat (1)	16 (1:5)
Kittens (Stage 1-D)	mat (1); hat (1)	17 (1:1.4)
What Goes Into the Bathtub? (Stage 1-C)	cat (2)	17 (1:2)
Dressing Up (Stage 1-E)	hats (1)	17 (1:2)
Pet Parade (Stage 1-D)	cat (1); hat (1)	20 (1:2)

emplars of the *at* rime but the presence of 20 unique words and a word density ratio of 1:2 indicate that this text would be challenging for beginning readers. The ideal situation would be to use texts that have more engaging content and language than many of the phonetically regular texts of the past and that provide more opportunities to apply phonics strategies than most, if not all, of the little books and literature-based programs of the present.

Multiple-criteria texts. What might a model be for such multiple-criteria text? A writer whose text is described as appropriate for beginning readers within both popular (Menand, 1997) and professional literature (Anderson, Hiebert, Scott, & Wilkinson, 1985) is Dr. Seuss. His texts, particularly *Green Eggs and Ham* (1960), are identified as models for beginning reading materials. *Green Eggs and Ham* is considerably longer than most texts for beginning readers, but if its first seven pages are considered, its characteristics can be compared with the texts already reviewed.

The content of its first two pages follow:

That Sam-I-am!
That Sam-I-am
I do not like
That Sam-I-am!

As can be seen in Table 1, *Green Eggs and Ham* differs from the present texts of literature-based anthologies and little books in at least two ways. First, the density ratio of 1:5 is closer to the ratios of the texts of the past than is the word density ratios of current texts. Second, Dr. Seuss used high-content words to create rhyme in the text, such as *train* and *rain*. These words are represented by illustrations,

but the picture-text match is not so concrete that children are encouraged to read the illustrations. Dr. Seuss used a form of repetitive text that allows children frequent exposure to high-frequency and, to a lesser extent, phonetically regular words. Unlike the perception that *Green Eggs and Ham* is a vehicle for phonics instruction (Menand, 1997), the basis of the book is high-frequency words.

Although there are many anecdotes about children who learned to read at home with Dr. Seuss's books, the presence of 24 words in the first 7 pages of *Green Eggs and Ham* means that the text would be a difficult one to use with a class of 25 to 30 first graders whose text experiences begin in school. Might the features of *Green Eggs and Ham* with its low density ratio be applied with a focused curriculum of phonetically regular and high-frequency words? To illustrate what might be possible, I have created a text, *That Fat Cat!* based on the curriculum referred to previously (Hiebert et al., 1998), that maintains a low word density ratio and rhyming and rhythmic text:

Fran can pat the cat.

Pat! Pat! Pat!
That fat cat.

Stan can pat the cat.

Pat! Pat! Pat!
That fat cat!

The man can pat the cat.

Pat! Pat! Pat!
That fat cat!

Scat! Scat! Scat!

The rat can NOT pat the cat.
That fat cat!

I make no claims for the literary merit of this text, but it is an example of a text that allows for children to apply knowledge about several different systems of written English. Opportunities for the application of the alphabetic principle are prominent, as they should be at this level. But this text differs from the typical phonics readers of the past and present (e.g., Cassidy et al., 1987) in that multiple exemplars of a rime are presented, encouraging children to apply knowledge about word patterns rather than to memorize words. As the summary of this text in Table 1 indicates, there are 5 exemplars of the *an* rime and 6 exemplars of the *at* rime. Further, this text gives experience with a core group of high-frequency words. Unlike the texts of the little book program, the high-frequency words do not occur serendipitously. In the texts that precede and succeed this text, words from the same clusters of most frequent words would appear. Finally, this text has sufficient repetition and rhythm to ensure sustaining children's engagement over several readings of the text.

Over a decade ago, Anderson et al. (1985) called for inventive writers to use Dr. Seuss as a model for creating engaging texts for beginning readers. This call needs to be extended again but, this time, with a clearer mandate—one that derives from a strong vision of what beginning readers need to learn. Such texts require thought to word density ratios and to the repetitions across as well as within texts of words that share phonetic elements.

For writers to be given a clearer mandate, researchers need to address questions that have been lost in the perennial debates over which methodology is "best." If the children who struggle to learn to read in schools are to learn to read well, the wisdom of teachers who have applied numerous different methodologies will need to be captured. Experiments where the influences of particular features of texts and instructional strategies are examined in depth will clarify issues of how particular children acquire particular processes. Fine-grained analyses will be needed for how teachers converse with students about particular features of texts, while supporting children's engagement and interest in text. Only through our combined wisdom and work as reading educators

will children in our schools be given the texts, instruction, and activities that they require to become the readers they need and want to be.

Author's notes

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305R70004, as administered by the Office of Educational Research and Improvement, U.S. Department of Education. However, the comments do not necessarily represent the positions or policies of the National Institute of Student Achievement, Curriculum, and Assessment, the National Institute on Early Childhood Development, or the U.S. Department of Education, and you should not assume endorsement by the U.S. federal government.

My interest in the study of beginning texts began with the shift in texts in California. Consequently, I have used the texts from the program that was most widely used in California in the textbook adoption cycle that began in 1989. In this article, I have not analyzed the textbook program, including the little book component, of which I am an author because it was not on the California textbook list during the previous adoption cycle. Analyses by Hoffman et al. (1994) of the Texas adoptions (a later copyright) indicate that similarities across programs are substantial, so the same descriptions could be directed to the textbook program on which I have been an author.

Hiebert is a professor in the School of Education at the University of Michigan (CIERA, 610 E. University, 1600 SEB, Ann Arbor, MI 48109-1259, USA) and the director of the Center for the Improvement of Early Reading Achievement, which is the national center of research on early reading. She has worked in the field of early reading acquisition as a teacher and teacher educator for more than 25 years.

References

Adams, M.J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.
Allington, R.L. (1994). The schools we have. The schools we need. *The Reading Teacher, 48*, 14 – 29.
Anderson, R.C., & Pearson, P.D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P.D. Pearson, R. Barr, M.L. Kamil, & P.

Mosenthal (Eds.), *Handbook of reading research* (pp. 255 – 292). New York: Longman.

Anderson, R.C., Hiebert, E.H., Scott, J., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers*. Urbana, IL: Center for the Study of Reading & The National Academy of Education.

Ashton-Warner, S. (1963). *Teacher*. New York: Bantam.

Beck, I.L. (1997, October/November). Response to "Overselling phonics" [Letter to the editor]. *Reading Today*, p. 17.

Bloomfield, L., & Barnhart, C.L. (1961). *Let's read: A linguistic approach*. Detroit, MI: Wayne State University Press.

Bond, G., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly, 2*, 5 – 142.

Brennan, A., Bridge, C., & Winograd, P. (1986). The effects of structural variation on children's recall of basal reader stories. *Reading Research Quarterly, 21*, 91 – 104.

Bridge, C.A., Winograd, P.N., & Haley, D. (1983). Using predictable materials vs. preprimers to teach beginning sight words. *The Reading Teacher, 36*, 84 – 91.

California English/Language Arts Committee. (1987). *English-language arts framework for California public schools (kindergarten through grade twelve)*. Sacramento, CA: California Department of Education.

Carroll, J.B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book*. New York: Houghton Mifflin.

Cassidy, J., Roettger, D., & Wixson, K.K. (1987). *Join the circle*. New York: Scribner.

Chall, J.S. (1982). *Learning to read: The great debate* (2nd ed.). New York: McGraw-Hill. (Original work published 1967)

Clay, M.M. (1985). *The early detection of reading difficulties*. Portsmouth, NH: Heinemann.

Cullinan, B., & Galda, L. (1994). *Literature and the child* (3rd ed). Fort Worth, TX; Harcourt Brace.

Deming, A.G. (1988). *Who is tapping at my window?* New York: Dutton.

Durr, W.K., Pikulski, J.J., Bean, R.M., Cooper, J.D., Glaser, N.A., Greenlaw, M.J., & Schoephoerster, H. (1986). *Bells*. Boston: Houghton Mifflin.

Ehri, L.C., & Sweet, J. (1991). Fingerpoint-reading of memorized text: What enables beginners to process the print? *Reading Research Quarterly, 26*, 442 – 462.

Elson, W.H., & Gray, W.S. (1930). *Elson basic readers*. Chicago: Scott, Foresman.

Fisher, C.W., & Hiebert, E.H. (1990). Characteristics of tasks in two literacy programs. *Elementary School Journal, 91*, 6 – 13.

Flesch, R. (1957). *Why Johnny can't read*. New York: Random House.

Foorman, B.R., Francis, D.J., Fletcher, J.M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*, 37 – 55.

Fountas, I.C., & Pinnell, G.S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Gibson, E.J., & Levin, H. (1975). *The psychology of reading*. Cambridge, MA: MIT Press.

Goodman, K.S. (1968). The psycholinguistic nature of the reading process. In K.S. Goodman (Ed.), *The psycholinguistic nature of the reading process* (pp. 13 – 26). Detroit, MI: Wayne State University Press.

Grossen, B. (1997). *30 years of research: What we know about how children learn to read* (*A synthesis of research on reading from the National Institute of Child Health and Development*). Santa Cruz, CA: The Center for the Future of Teaching and Learning.

Hiebert, E.H., Colt, J.M., Catto, S., & Gury, E. (1992). Reading and writing of first-grade students in a restructured Chapter 1 program. *American Educational Research Journal, 29*, 545 – 572.

Hiebert, E.H., Martin, L.A., Gillard, G.A., & Wixson, K.K. (1998). *Tool kit for tutors*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement.

Hiebert, E.H., & Raphael, T.E. (1998). *Early literacy instruction*. Fort Worth, TX: Harcourt Brace.

Hoffman, J.V., McCarthey, S.J., Abbott, J., Christian, C., Corman, L., Dressman, M., Elliot, B., Matheme, D., & Stahle, D. (1994). So what's new in the "new" basals. *Journal of Reading Behavior, 26*, 47 – 73.

Holdaway, D. (1979). *The foundations of literacy*. Sydney: Ashton Scholastic.

Johnston, F.R. (1998). The reader, the text, and the task: Learning words in first grade. *The Reading Teacher, 51*, 666 – 676.

Juel, C., & Roper/Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly, 20*, 134 – 152

Juel, C., & Solso, R.L. (1981). The role of orthographic redundancy, versatility, and spelling-sound correspondences in word identification. In M.L. Kamil (Ed.), *Directions in reading: Research and instruction* (pp. 74 – 92). Rochester, NY: National Reading Conference.

Lesgold, A.M., Resnick, L.B., & Hammond, L. (1985). Learning to read: A longitudinal study of word skill development in two curricula. In G. MacKinnon & T. Waller (Eds.), *Reading research advances in theory and practice* (pp. 107 – 137) Orlando, FL: Academic Press.

Literacy 2000: Stage 1, Sets A-E. (1989). Auckland, New Zealand: Shortland.

Martin, B. (1967). *Brown bear, brown bear, what do you see?* New York: Henry Holt.

Martin, B., & Brogan, P. (1971). *Teacher's guide to the instant readers*. New York: Holt, Rinehart, & Winston.

Menand, L. (1997, October 6). How to frighten small children: The complicated pleasure of kids' books. *The New Yorker*, 112 – 119.

Peterson, B. (1991). Selecting books for beginning readers: Children's literature suitable for young readers. In D.E. DeFord, C.A. Lyons, & G.S. Pinnell (Eds.), *Bridges to literacy: Learning from Reading Recovery* (pp. 119 – 147). Portsmouth, NH: Heinemann.

Pikulski, J.J., Cooper, J.D., Durr, W.K., Au, K.H., Greenlaw, M.J., Lipson, M.Y., Page, S., Valencia, S.W., & Wixson, K.K. et al. (1993). *The literature experience*. Boston: Houghton Mifflin.

Pressley, M., Rankin, J., & Yokoi, L. (1996). A survey of instructional practices of outstanding primary-level literacy teachers. *Elementary School Journal, 96*, 363 – 384.

Raphael, T.E., & Hiebert, E.H. (1996). *Creating an integrated approach to literacy instruction*. Fort Worth, TX: Harcourt Brace.

Rasmussen, D., & Goldberg, L. (1964). *The bad fan* (Level A, Basic Reading Series). Chicago: Science Research Associates.

Samuels, S.J. (1970). Effects of pictures on learning to read, comprehension and attitudes. *Review of Educational Research, 40*, 397 – 408.

Stickney, J. (1885). *A primer*. Boston: Ginn.

Seuss, Dr. (1960). *Green eggs and ham*. New York: Random House.

Texas Education Agency. (1990). *Proclamation of the State Board of Education advertising for bids on textbooks* (Proclamation 68). Austin, TX: Author.

Thorndike, E.L. (1903). *Educational psychology*. New York: Lemcke & Buechner.

Wylie, R.E., & Durrell, D.D. (1970). Teaching vowels through phonograms. *Elementary English, 47*, 787 – 791.