Reading & Writing Quarterly, 21: 61–86, 2005

ISSN: 1057-3569 print
DOI: 10.1080/10573560590523667

Routledge
Taylor & Francis Group

TEXT DECODABILITY AND THE FIRST-GRADE READER

Heidi Anne E. Mesmer
College of Education, Oklahoma State University, Stillwater, OK, USA

This study investigated the effects of highly decodable text and coordinated phonics instruction on first graders' word recognition strategies. The quantitative study sought to examine the validity of a major claim about highly decodable text—that it enables readers to apply phonics instruction to a greater extent than less decodable text. All participants received the same fourteen-day phonics instruction. However, participants read either highly decodable or less decodable text following the instruction. Treatment participants reading highly decodable text were found to apply letter/sound knowledge to a greater extent than control participants. They also were more accurate and relied on examiners less for assistance. Treatment and control participants did not differ in self-correction rates. As a replication of an earlier study, this work suggested that readers with knowledge of the alphabetic principle, given the same phonics instruction, will apply it more in a highly decodable context.

Recently, researchers, educators, and policymakers have expressed a renewed interest in the instructional materials used with beginning readers. Several articles exploring text issues have been written (Brown, 1999; Cole, 1998; Hiebert, 1999; Johnston, 1998; Mesmer, 1999; Watson, 1997). The topic "readers and texts" has become an inquiry focus of the Center for the Improvement of Early Reading Achievement (CIERA), and a great variety of instructional materials has flooded the textbook market. Due in part to policies in many states, highly decodable text has become increasingly visible (California Department of Education, 1996; Texas Department of Education, 1997; Virginia General Assembly, 2001).

The topic of text decodability is all the more heated within the context of struggling readers. A number of programs that are typically used with struggling readers, including those in special education programs, rely on

Address correspondence to Heidi Anne E. Mesmer, Oklahoma State University, College of Education, 249 Willard Hall, Stillwater, OK 74074. E-mail: Haemesmer@hotmail.com

highly decodable materials (Wilson, 1996). Nonetheless, opinions vary regarding its use with struggling readers. Some researchers question any use of highly decodable text (Allington, 2001; Allington & Woodside/Jiron, 1998; Hicks & Villaume, 2001; Routman, 1997), while others insist it to be foundational to reading acquisition (Fletcher, Francis, & Foorman, 1997; Groff, 1999; Grossen, 1997; Kameenui & Simmons, 1997; Moats & Hall, 1999; Stein, Johnson, & Gutlohn, 1999). Still others believe that decodability should be considered in choosing instructional materials but should not serve as the only textual scaffold available to readers (Brown, 1999; Cole, 1998; Hoffman, 2001; Mesmer, 1999, Hiebert, 1998).

Unfortunately, opinions and hypotheses dominate discussions because only a small amount of research informs current knowledge. The National Reading Panel (2000), Snow, Burns, & Griffin (1998), and others (Allington, 1997; Pikulski, 1998) have noted this gap. In response to calls for study, researchers have analyzed reading materials (Hoffman, Patterson, Assaff, Sailors, Mast, & McCoy, 1993; Menon & Hiebert, 1999; Stein et al., 1999) or investigated how readers respond to texts of varying levels of decodability (Hoffman, Roser, Paterson, Salas, & Pennington, 2000; Mesmer, 2001b). In 1998, Pikulski pinpointed the need for an examination of readers' responses to text by asking, "How many studies have been published in which the nature of beginning reading texts is systematically varied and the effects measured?" (p. 30). The present study responds to this need.

Decodability, like other textual scaffolds, may have some benefits. However, it may also produce side effects. Ultimately, classroom teachers need to understand when they might use texts with some level of decodability and when these texts are not useful. This study sought a sophisticated understanding of the use of decodability as a textual scaffold. Decodability is not a long-term solution: like the training wheels on a bicycle, it is designed to facilitate future independence.

BACKGROUND

This section will briefly review the current literature. The first portion examines definitions of highly decodable text, and the following section describes its theoretical purposes. The third section analyzes studies that examined the effects of decodable text on readers, and the final section presents a developmental model for the present study.

What is Decodable Text?

Most define decodable text by the presence of the following two features: a proportion of words with phonically regular relationships between letters and sounds, and a degree of match between the relationships represented in text and those that the reader has been taught (Allington & Woodside-Jiron, 1998; Groff, 1999; Hiebert, 1998; Stein et al., 1999; Willows, Borwick, & Hayvren, 1981).

Phonic Regularity

In most studies, phonic regularity refers to letter/sound patterns that are structurally simple, follow common letter/sound associations, and recur in many different words (Ehri, 1998; Hiebert, 1998; Hoffman, 2001; Hoffman et al., 1993; Juel & Roper-Schneider, 1985; Menon & Hiebert, 1999). Phonic regularity has been assessed by examining texts using many methods, such as rating scales, rimes, and bigram versatility.

Rating scales have been used extensively to determine phonic regularity. The following represent some tools:

1. a three-level rating scale based on Venezky's (1970) work (Juel & Roper-Schneider, 1985)
2. an eight-level system (Menon & Hiebert, 1999)
3. The Scale for Text Accessibility and Support-Grade 1 or STAS 1 (Hoffman et al., 1993, 2000).

(See Mesmer, 2001a, for a more detailed discussion of these systems).

Given the many ways of accounting for phonic regularity, several unifying conclusions can be made about what regularity is and how it relates to decodability. First, regular patterns, be they rimes or bigrams, recur in many different words. Second, the most regular patterns have one-to-one letter/sound correspondences (cat, trap, stump) or predictable two-to-one patterns (seat, boat, blue). Third, highly decodable text is usually not distinguished by a great number of regular words but instead by the clustering of words with similar word parts together within specific passages (Hiebert, 1999; Juel & Roper-Schneider, 1985). Thus, in highly decodable text, specific words may not be repeated more often, but certain universal word parts are (rimes, bigrams, and specific patterns).

Lesson-to-Text Match

Currently decodable text is characterized by a degree of match between letter/sound relationships represented in text and those that the reader has been taught. Here, this feature is called lesson-to-text match (LTTM), but in other works it has been called "instructional consistency," "wholly decodable," and "text related," (Hoffman, 2001; Reutzel & Daines, 1987; Stein et al., 1999). In most studies, LTTM is expressed as a percentage of words matching phonics lessons specified in teacher's manuals. In seven studies, researchers specifically evaluated this variable (Barr, Dreeben,

& Wiratchai, 1983; Beck, 1981; Beck & Block, 1979; Meyer, Greer, & Crummey, 1987; Reutzel & Daines, 1987; Stein et al., 1993, 1999). Consistently, they found highly decodable text to contain more words matching phonics lessons. The purpose of the LTTM feature is to specify what has been taught, because it can be independently measured much more easily than learned information.

There are three main limitations of the LTTM feature. First, in previous studies conceptualizations of LTTM took into account lessons specified in teacher's editions despite the fact that actual phonics lessons will differ (Ball & Cohen, 1996; Pressley, Rankin, & Yokoi, 1996). Like Memer (2001b), the present research examines actual phonics instruction as it coordinates with words in text. Second, in previous studies, the nature of the lessons is not specified even though the type of phonics instruction, use of modeling, opportunity for practice, and quality of feedback all influence how "decodable" words are to the reader. For this reason, this study's methodology specifies exactly the nature of lessons so that judgments can be made about the instruction's potential for rendering words decodable. A final limitation of the LTTM feature is that no optimal percentage of LTTM has been established by research (Allington, 2001). A major purpose of this study is to report the percentage of words coordinated with lessons and to also report the effects on readers.

Both features of the decodability definition, phonic regularity and LTTM, contain restrictions. Using the analogy of readability, the limitations of decodability may be better understood. Readability is established by examining word- and sentence-level features that are independent of the reader. However, there is clearly a reader/text interaction that also influences the readability of a book. This reader/text interaction is based on individual differences that are difficult, if not impossible, to measure. Like readability, decodability can be independently but not comprehensively measured by calculating regularity and LTTM. Just as educators have accepted readability as a gross measure of text difficulty, so also can they accept decodability as a gross measure of word level difficulty at specific literacy stages. As a whole, the decodability construct is functional and draws on over twenty years of research.

What Are Theoretical Purposes of Decodable Text?

To become proficient, beginners must read great amounts of text, yet they are significantly challenged by the task of identifying words in print. Textbook authors have employed a variety of textual scaffolds to assist readers as they identify words in text. All textual scaffolds, be they connected to decodability or not, represent a manipulation of the reading material at some level. Materials that do not contain textual scaffolds are written with aesthetic purposes as the primary authorial force. Thus, the purpose of decodability, like other scaffolds, is to assist the literacy acquisition process.

Decodability, however, prompts readers to use specific strategies to identify words. As a textual scaffold, decodability has two purposes. First, it offers a context through which readers can use letter/sound strategies (Adams, 1990; Gough, 1997; Stahl, Duffy-Hester, & Stahl, 1998). Theoretically, students reading highly decodable text should be able to apply phonics instruction due to the coordination between letter/sound information and words in print (Adams, 1990; Barr, 1972; Bryne, 1991; Gough & Juel, 1991). In contrast, when students learn isolated phonics rules or complete worksheets, they are not likely to generalize decoding strategies to reading. Second, the regularity in decodable text along with the lesson-to-text match focuses attention on letter/sound strategies. Readers who apply letters and sounds during connected reading must be attending consciously to them (Adams, 1990). Decodable text encourages this conscious attention.

The Effects of Decodable Texts and Instruction on Word Recognition Strategies

Four examinations of readers' responses to decodable text can be found in the literature (Foorman, Frances, Fletcher, Schatschneider, & Mehta, 1998; Hoffman et al., 2000; Juel & Roper-Schneider, 1985; Mesmer, 2001b). A quasi-experimental study by Juel and Roper-Schneider (1985) compared the abilities of two first grade groups ($N = 93$), one reading in a highly decodable text and the other in a high-frequency text. Both groups received the same supplemental, scripted phonics program. The groups did not differ in their abilities to read words in their own basals, but the decodable group performed better in reading pseudo-words at November/December and February assessments. For this reason, the current study was timed during the first two-thirds of the first grade year.

In a larger study, Foorman et al. (1998) compared three groups of at-risk first and second graders ($N = 285$), one of which received direct phonics instruction coordinated with decodable text. The other two groups received either an embedded or implicit phonics program. Researchers measured decoding, receptive vocabulary, phonological processing, spelling, and comprehension. The readers in the highly decodable text group had superior performance on both the decoding and comprehension measures, with the differences being most robust in decoding.

A smaller study contrasted readers' abilities to apply phonics instruction in a decodable and control context (Mesmer, 2001b). This study eschewed pseudo-word assessments in favor of decoding abilities used during connected reading. For fifteen days, all participants received the same phonics instruction but were assigned to read either highly decodable or control texts. Students reread texts, and data were collected using running records. Oral reading errors were rated for the match between the pronounced and written word. The treatment group had more errors in which two or more phonemes matched the print than the control group.

Hoffman et al. (2000) also examined student behaviors as related to predictability, decodability, and text-leveling systems. These researchers rated little "books" for difficulty as well as decodability and predictability. Participants were introduced to sets of books via three treatments, and their accuracy, fluency, and reading rates were measured. Participants' accuracy was positively and significantly correlated with more highly decodable texts, indicating that decodability may have assisted students in recognizing words. Participants' fluency was negatively correlated with less decodable text. Thus, as words became less decodable, participants were less fluent.

According to these studies, young readers may be influenced in very specific ways and at very specific times by text decodability. Readers seemed to be most influenced during the first portion of first grade, a period that in fact may be more precisely connected to word recognition phases. Readers using decodable text also appear to pay greater attention to letter/sound information, both in a global sense as well during contextualized reading. Furthermore, even when readers have been given the same phonics instruction, those who read highly decodable text appeared to apply this instruction more than control groups. Interestingly, three of these studies included significant numbers of English Language Learners (ELLs). A contribution of the present study is that the sample is different demographically.

A Model Based on Word Recognition Phases

Ultimately, decodability is a tool, one that can be used correctly or incorrectly, appropriately or inappropriately. Having discussed *how* decodability may work in the previous section, this section examines *when* it may work. As mentioned earlier, the theoretical purpose of decodability is to assist readers in applying letter/sound strategies. With this purpose in mind, the question remains: "When, developmentally, do readers need the most assistance to apply letter/sound knowledge?"

The word recognition literature specifies a series of developmental phases through which readers pass as they become automatic. In these stages, readers are distinguished by the extent to which they apply the alphabetic principle. Prealphabetic readers do not use letter/sound correspondences to recognize words (Ehri, 1991, 1994, 1995, 1998; Ehri

& McCormick, 1998). Partial alphabetic readers understand the alphabetic principle and gradually apply their knowledge by using beginning or final consonants. As readers move into full alphabetic reading, they connect all letters in a word with the phonemes that the letters represent. At the consolidated alphabetic phase, readers internalize multiletter units and patterns. Finally, in the automatic phase, readers retrieve most of their words by sight and are free to focus most of their attention on meaning.

Once readers understand and begin to apply their alphabetic knowledge at the partial alphabetic stage, they often struggle to fully apply it in the next phase. However, moving students into the full alphabetic phase is a crucial goal for first-grade teachers (Ehri, 1998). Highly decodable text may be particularly useful in helping young readers transition from the partial alphabetic phase into full alphabetic reading. At this transition, readers know enough about letter/sound correspondences to begin applying them but not enough to handle the full range of complexities in English. In addition, they need help to decode vowels, the most complex letter/sound relationships in English.

PURPOSE AND QUESTIONS

The study had the following four major purposes:

1. to examine the theoretical claims about text decodability
2. to identify students who would benefit from highly decodable text
3. to contrast reading behaviors of groups (accuracy, self-corrections, need for assistance, repetitions, and insertions)
4. to make improvements to an earlier study.

The study's design chiefly investigated whether or not use of highly decodable text resulted in a greater application of letter/sound information. Both treatment and control groups received the same phonics instruction, but the treatment group read highly decodable text and the control group read less decodable text. The essential question was, "When treatment and control groups were given the same phonics instruction, did treatment readers apply this instruction more as they read a highly decodable text than control readers who read less decodable text?"

This major query was examined through an outcome variable in which readers' oral reading errors were rated for their match to the printed stimuli. Each error was assigned a point for each phoneme matching the print in the correct sequence. For example, the word "ship" has three phonemes. The error "lip" has two corresponding phonemes and would receive a two rating; the error "show," a rating of one; and the error "got," zero. Errors with two or more phonemes matching the print were called graphically similar. This division corresponds with the developmental juncture of the model. A reader who correctly pronounces zero or one phonemes is most likely a pre- or partial-alphabetic reader using either no alphabetic information or only initial sounds. A reader pronouncing at least two or more phonemes correctly is more likely to move into the full alphabetic phase by using more than simply initial or final sounds. The treatment group was expected to have a higher average number of errors with two or more phonemes matching the print.

The second purpose of the study was directly related to the type of student who might benefit from highly decodable text. This purpose was exploratory in nature but deemed vital given state mandates for the use of highly decodable text in first grade and the preponderance of highly decodable text in programs for struggling readers. Using a statewide screening measure, the study selected students who were beginning readers demonstrating an understanding of the alphabetic principle and some knowledge of sight words. Students who did not demonstrate these skills at a minimal level or who had clearly surpassed beginning reading stages were not selected.

The third purpose was to build on earlier work by contrasting reading behaviors (Hoffman et al., 2000). Through data collected on running records, the researcher examined word accuracy, the number of self-corrections, repetitions, and the number of words provided by the assessor. This query was related both to participants' autonomy as they read highly decodable text as well as the possible side effects of decodability. Treatment participants were expected to be more autonomous as evidenced by higher percentages of word accuracy, fewer appeals to the examiner to pronounce words, and higher self-correction rates. As a side effect, they were expected to repeat more often due to their increased focus on decoding.

Finally, in three crucial ways, this research improved upon a previous study. First, it was timed during the first one-third of the first grade year, when previous researchers had found highly decodable text to have its most pronounced effects. Second, the decodable and control texts were more uniform because the same publishers produced both. Third, a new scale, containing eight levels instead of three, was used to measure phonic regularity. The following questions guided the study:

1. When compared to control participants receiving the same phonics instruction, did treatment participants have more errors in which two or more phonemes matched the printed word?
2. When compared to control participants receiving the same phonics instruction, did treatment participants have higher percentages of word

accuracy, fewer occasions in which they relied upon examiner assistance to pronounce words (tolds), more self-corrections, and more repeated words or phrases?

METHODS

Sample

One school in a metropolitan area in the southeast participated in the study. The school belonged to an urban district with approximately fifty elementary schools. The school was labeled a "model" school by the system, and many parents petitioned to have their children attend the school. This fact made the population of students at this school different than other schools, with higher poverty rates. The population at this school consisted of 47% African-Americans, 51% Caucasian, and 2% other; 26% of the students received free or reduced priced lunches.

Students from three first-grade classrooms participated in the study. In these three classrooms, teachers shared similar practices. Teachers used the morning message, shared reading, journaling, writer's workshop, and independent reading. Teachers also used a variety of reading materials, including literature, big books, and "little books," but did not make great use of highly decodable text. Prior to the intervention, the researcher met with the teachers and discussed the purpose of the study. Teachers agreed to the intervention and were eager for the additional small group instruction that their students would receive.

Fifty-one participants were screened for this study and 23 participated in the study. Participants had summed scores between 22 and 44 on the statewide screening measure. Seventeen students had scores that were too high to qualify for the study, five students had scores that were too low, one student discontinued due to excessive absences, and five students were eliminated due to the need to maintain equal numbers across classrooms. Those who participated in the study understood the alphabetic principle, could apply this to initial and final consonants, and had acquired some sight word knowledge. They had neither begun to fully analyze words nor had spellings that included many vowels.

The sample consisted of ten females and thirteen males. Eleven participants were Caucasian and twelve were African-American. An equal number of children from each classroom participated in the study (eight per classroom), with the exception of one classroom, from which one participant was lost due to excessive absences. In each classroom, the eight participants were divided into two small groups, one receiving the treatment and one the control. The intervention took place outside of the regular classroom, with the researcher meeting separately with treatment and control groups. A total of six small groups existed across the three classrooms.

Materials

The Phonics Readers Levels 2–4 by Sundance Publishers were chosen as the treatment text because they were highly decodable on three measures—LLTM, regularity, and number of syllables. This pronounced decodability made them especially viable for a short-term intervention. The control books included the *Alpha Kids Readers* Levels 2–5, also published by Sundance publishers. These texts could be best described as "little books." They contained short sentences, many high-content words, close print–picture match, natural language, and some predictable elements.

Analysis of Decodability

During the total intervention, each group read fourteen books. The content words in the treatment and control text were compared on LLTM, phonic regularity, total number of running words, number of syllables, and repetitions. Content words are nouns, verbs, adjectives, and adverbs, whereas structure words are pronouns, determiners, quantifiers, prepositions, intensifiers, coordinate and adverbial conjunctions, conjunctive adverbs, relative pronouns, and auxiliary and linking verbs (Clark & Clark, 1977). Content words were counted as opposed to structure words because of the significant differences between these word-types, both in purpose and in frequency. To understand the essence of a text, readers must recognize these content words. For each feature, words were analyzed and book set means compared using t-tests ($\alpha = .05$).

In the areas of LTTM, specific phonic regularity categories, and number of syllables, the highly decodable and control texts differed significantly (see Table 1 for specific decodability statistics). Note that the LTTM takes into account *only* those sounds taught during the research intervention.

Procedures for measuring the LLTM followed those used in earlier studies (Mesmer, 2001b). Each content word was analyzed for its match to the letter/sound focus of lessons delivered by the researcher. For each book, an average LTTM was obtained by comparing the number of content words matching lessons to the total number of content words in the book. For example, every short /e/ word that appeared in the text when the short /e/ was the focus of the lessons was counted as matching. Later, when the sound focus shifted to short /a/, both short /a/and short /e/ words were counted as matching. The treatment text contained about 40% of words matching the lessons, whereas the control text contained only about 8%.

TABLE 1 Decodability Statistics for Treatment and Control Books

	Treatment		Control			
Feature	*M*	*SD*	*M*	*SD*	*T*	*p*
Percent words matching lessons/book (LLTM)	40.3	(16.3)	7.7	(2.11)	6.5	.0000
Number of content words/book	87.9	(34.3)	52.5	(29.5)	2.8	.009
Number of total words/book	122.5	(45.8)	92.5	(42.9)	1.7	.07
Number of repetitions/word/book	4.4	(0.9)	3.1	(0.7)	3.8	.0009
Number of syllables/word	1.1	(0.3)	1.3	(0.0)	−2.5	.01
Percent of content words/book in following categories:						
Regularity 1 (1 & 2 letter words)	5.1%	(5.9)	2.5%	(3.9)	1.3	.19
Regularity 2 (short vowel o-v-c)	54.2%	(18.5)	15.4%	(12.7)	6.3	.00
Regularity 3 (blends & diagraphs)	18.1%	(14.0)	10.4%	(2.9)	1.6*	.13
Regularity 4 (silent e)	2.1%	(2.8)	12.1%	(13.3)	−2.6*	.02
Regularity 5 (vowel digraphs)	0.0%	(0.3)	12.8%	(10.2)	−4.3*	.00
Regularity 6 (r- and l-controlled)	8.7%	(8.0)	4.5%	(3.7)	1.7	.09
Regularity 7 (diphthongs)	5.0%	(4.1)	12.9%	(10.7)	−1.6*	.03

Rather than reporting Level 8, the percent of multisyllable words, the table reports the average number of syllables per word. Thus, Level 8 is not reported, and the percentages for Levels 1–7 do not collapse to 100.

*indicates that a test for unequal variances was conducted.

Phonic regularity was evaluated using the scale developed by Menon & Hiebert (1999). This scale included eight levels and was chosen because of its specificity in categorizing words. Level 1 words included short one- and two-letter words, and Level 2 words contained short vowel c-v-c and v-c patterns. Level 3 words contained clusters (tr, str, sl) and digraphs (th, sh, ch) but not words ending in r, l, or gh. Level 4 words contained the silent e pattern, and Level 5 words vowel digraphs (ea, ee, oa). Level 6 words contained r-controlled words and words ending in ll. Finally, Level 7 words included diphthongs (oi), and Level 8 words were multisyllabic.

The mean percent of words per book was reported for each word type. The treatment and control books did not differ significantly at all levels of phonic regularity. The most notable differences occurred at Level 2 (short vowels): as might be expected, the decodable text contained many more Level 2 words in comparison to the control text. The decodable text also contained fewer silent e, vowel digraph, and diphthong words.

The words for the two sets of books were analyzed for the number of syllables each contained. For each book, an average number of syllables per content word was measured. As was found in previous studies, all book sets differed significantly on the number of syllables per word, with the decodable books containing fewer syllables per word than the control books.

Number of Content Words, Running Words, and Repetitions

The number of words contained in a text can place demands on the reader, and for this reason affect the difficulty level of the text. Therefore, each set was compared on the number of content words per book set. When the number of content words was compared, the text sets differed with the control books, on average containing about 52 words compared to the 87 words in the decodable text. Depending on one's perspective, this difference might favor either the control or decodable readers. Assuming that the reading of predictable, less decodable texts is usually driven by memory, the depressed number of content words in the control set would theoretically have enabled the control group to read more accurately because they would have had to remember fewer words. From this perspective, the difference favored the control group. However, the fewer number of content words could have disadvantaged the control group due to the less amount of time spent reading and the fewer opportunities to practice reading words containing the patterns taught.

For this reason, a second analysis was conducted comparing the total number of running words per book as opposed to only content words. In this analysis, the two book sets did not differ significantly. Collectively, fourteen highly decodable texts contained exactly 1724 words, and the control books exactly 1298 words. Within-set books varied widely in numbers of words, as indicated by the high standard deviations for both total number of content words and total number of running words.

Finally, each word within each book was analyzed for the number of times that it was repeated. The book means were than converted to group means for both decodable and control sets, yielding an average number of repetitions per word per book. On average, words in the highly decodable set were repeated one time more than those in the control set. Although repetition can characterize both highly decodable and less decodable text, repetition in highly decodable text appears to be higher.

Design and Procedures

Due to the numerous questions about highly decodable materials as well as calls for the manipulation of text, this study had an experimental design. The independent variable was text, either a highly decodable text or a less decodable text. The dependent variables included the degree of application of letter/sound information, word accuracy, percentage of errors self-corrected, number of repetitions, and number of words provided by the researcher. Both treatment and control participants received this phonics instruction. The phonics instruction intervention lasted fourteen days, and each participant spent twenty minutes per day in this instruction. Although all participants received the same phonics instruction, this instruction took place within their prospective treatment groups, outside of the regular classroom. Following each lesson, participants read either highly decodable or less decodable text, depending on group. Maintaining internal validity was a primary concern in the study, and the following procedures increased control:

1. selecting students from three different classrooms to control for teacher effects.
2. balancing the number of treatment and control participants within each classroom.
3. counterbalancing the times during the day that groups met for lessons.
4. selecting participants within a specific score range to decrease the sample's heterogeneity.
5. controlling the phonics lessons.

Preintervention Procedures

Students were assessed by their classroom teachers using a statewide literacy screening measure. Administered to first graders in fall and spring of each year, this screening gives teachers diagnostic information about their students and identifies students who are at risk for reading failure. All students in the first grade are administered a twenty item, preprimer word list and a spelling test. The spelling assessment approximated students' applied knowledge of letter/sound information. Students are assigned points for correct letter/sound representation in words based on various features of the English language (e.g., initial consonants, final consonants, blends, digraphs, short vowels, and long vowels). In addition, a bonus point is assign if the word is spelled correctly. The total number of points that a student can earn for the beginning of the year spelling is 57. The raw scores for both the word list and spelling list are summed. If the summed score is below 28, then other diagnostic measures are administered, including phonological awareness, concepts about print, and letter knowledge. Readers with scores at 28 or below qualify for additional assistance.

Within classrooms, participants were randomly assigned to treatments. Four participants from each classroom were in the treatment group and four were in the control group. When the total number of qualifying participants within a classroom was not equal to eight, participants were ordered by their summed scores alternately eliminated from the top and then the bottom of the score range until only eight participants remained.

Intervention Procedures

The researcher delivered the daily twenty-minute lesson. The lesson components were ordered to scaffold participants' alphabetic learning. The lesson was based on a model called "Hear-it, Read-it, Spell-it," with which students first practiced hearing a target vowel sound with phonemic segmentation activities, practiced reading that sound in words, spelled the sound in a dictated word, and then read books. For the treatment group, the words in the books contained the target sound, and for the control group, there was no control on words in text (See Appendix A, Lesson Matrices).

Instruction within each activity progressed from teacher modeling to group guided practice to independent practice. During independent practice, each participant individually completed one to two items within a given activity (e.g., reading words, segmenting words, or reading pages). The researcher gave participants specific, corrective feedback for independent practice items. When items were not correct, the researcher asked a scaffolding question. If the student was not successful after the scaffolding question, then the researcher modeled the correct procedure and asked the student to repeat. If the student was still not successful, then the researcher completed the item with the student.

Phonemic Segmentation

The phonemic awareness activities specifically focused on segmenting words because this skill is crucial for later decoding (Muter & Snowling, 1998; Nation & Hulme, 1997; Snow et al., 1998). In addition, phonemic segmentation corresponds with the skills that a full-alphabetic reader must have in order to be successful (Ehri, 1998). The researcher first demonstrated how words could be stretched out and then modeled a segmenting procedure for tapping the table each time a new sound was heard. After the word was segmented, it was pronounced completely. Toward the end of the lesson sequence, when students were learning both short e and short a, they were asked to differentiate between two words containing these sounds using the methods described above. The Lesson Matrices detail the specific words used for modeling, guided practice, and independent practice.

Reading Words

During this part of the lesson, a strategy for blending sounds together in words was modeled. The process for reading the words involved three parts. Covering the last letter in the word, participants articulated the first and second letter/sounds and then blended these together. After blending these two sounds, they added the ending sound to the word and finally articulate the blended word.

Dictated Writing

At the end of each lesson, participants wrote one word containing focus sounds. During this segment of the lesson, the researcher said the word in isolation and then used the word in a sentence to ensure participants' familiarity with it. After this point, the researcher stretched the word out, slowly enunciating each sound, and then asked the participants to write the word. While the participants wrote the word, the researcher alternately said the word in a standard fashion and stretched the word. All of the writing was recorded in little journals.

Reading a New Book

The last segment of the lesson included introducing and reading a new book. Prior to reading the book, each participant was asked to look through the book at the pictures, thinking about the content of the story. After this, a conversation followed, posing possible predictions about the story. During the first reading of the book, the entire group was asked to read the book chorally. This independent reading gave participants practice in applying the skills that they had just learned. Depending on the group's success, the researcher read along with the participants or stayed silent. At the second reading, each child was asked to volunteer to read a page of the book.

Approximately 90% of the time, all participants volunteered to read a page. When a child did not volunteer, the researcher asked the participants to read a page and volunteered to read along with him/her. At points when a participant could not supply the correct word, the researcher waited three seconds and then supplied the word.

Rereading

Before each lesson, one participant from each group individually reread the book from the previous day with a graduate assistant. Each participant orally reread a familiar book every fourth day. During this time, oral reading errors were collected using running record procedures (Clay, 1993). To consistently collect types of oral reading errors, standard procedures were established. Because "Tolds" were counted as errors, participants were given ample opportunity to figure a word out and encouraged to use strategies to help them. Participants were told a word if after two-second wait, a generic prompt, and then a five-second wait they could not correctly pronounce it. The running records and the lessons were not performed by the same person to avoid any bias in providing words to students. Repetitions, insertions, omissions, and substitutions were denoted using standard markings (Clay, 1993).

Post Intervention Procedures

After the intervention was completed, researchers coded the data. First, the average word accuracy rate for each running record was calculated. Insertions, omissions, substitutions, and tolds were all counted as errors. Repetitions were not counted as errors because, in many cases, readers repeated words or phrases to monitor themselves. Second, researchers rated each substitution for its graphic similarity as described earlier. After rating each substitution, a count was tabulated of the number or errors with ratings of 0–1 and those greater than or equal to 2. The number of self-corrections also was tabulated. Finally, any running record with 85% word accuracy was not used in this analysis. This resulted in two pieces of data, one from each group, being omitted.

FINDINGS

Analysis

Due to the design and two-sample comparison, Student's t appeared to be the most appropriate path for data analysis. However, the assumptions of this test were examined, and an inspection of the data indicated that t-tests would not be sufficient for several reasons. The distributions of many of the sampling means were not normal and often moderately skewed. The t-test is relatively robust to a violation of the normal distribution assumption if the total sample is over 40, which it was not. In addition, the sample means did not have equal variances. The t-test is also robust to this violation when sample sizes are equal, which they were not.

For these reasons, the non-parametric equivalent to the t-test was the Wilcoxon Rank Sum Test. In a recent comparison of Student's t-test and this test, the Wilcoxon better maintained Type 1 errors at the nominal rate for non-normal distributions (MacDonald, 1999; Zimmerman, 1994). In addition, with non-normal distributions, the Wilcoxon gave better power, an advantage that was increased when sample sizes were unequal (MacDonald, 1999). The Wilcoxon places the scores of each participant in rank numerical order while retaining the group identity (treatment of control). Then the ranks are summed by group. Higher rank sums indicated higher overall scored. As specified in APA, the sum of ranks (W) are reported along with related p values. In addition, z scores are reported.

The baseline equivalence of participants was established using the statewide assessment described earlier. Table 2 shows the mean scores and standard deviations on the spelling assessment, word list, and the summed scored. This information is reported to make the findings more understandable. However, the means were not analyzed parametrically: neither the word list scores (treatment $W = 111.5$, control $W = 119.0$, $p = .93$) nor

TABLE 2 Participant's Mean Scores on the Word List and Spelling Assessments

	Word List Max = 20		Spelling Max = 57		Summed score Benchmark score = 28	
	M	*SD*	*M*	*SD*	*M*	*SD*
Treatment	12.9	(5.12)	19.7	(5.3)	32.6	(6.28)
Control	12.5	(4.17)	20.4	(4.6)	33.0	(6.46)

the spelling scores (treatment $W = 107.5$, control $W = 123.5$, $p = .87$) nor the summed scores (treatment $W = 104.5$, control $W = 126.5$, $p = .71$) differed significantly. On average, participants had made some inroads into reading and knew some sight words but did not appear to be full alphabetic readers.

The major question of the study related to the degree to which alphabetic information was applied during text reading. This was measured by the GSE variables. The GSE ≤ 1 represented substitution errors in which zero or one phoneme matched the printed word, and the GSE ≥ 2 category represented substitution errors in which two or more phonemes matched the printed word. Total GSEs or Total Substitutions represented the combination of these categories. Participants had an average of 3.5 running records. Table 3 shows the data for the aforementioned categories.

TABLE 3 Sum of the ranks, z scores, and p values for GSE < 1, GSE > 2, and Total GSEs categories

	Treatment $N = 11$	Control $N = 12$
Mean errors with 0–1 rating (GSE < 1)	4.4 (5.4)	6.2 (5.9)
W(sums of the ranks)*	94.5	136.5
z	−1.0	
Mean errors with 2+ rating (GSE $\geqslant$ 2)	5.2 (2.3)	3.1 (1.9)
W (sums of the ranks)†	139.1	91.5
z	2.0	
Mean total substitutions (total GSEs)	9.6 (6.9)	9.3 (6.8)
W (sums of the ranks)‡	112.5	118.5
z	.14	

Mean errors represents the mean number of errors in a given category over 3.5 points of measurement per participant.

*p value = .26.

†p value = .03.

‡ p value = .88.

Essentially, the data showed that the groups had the same total number of substitution errors as indicated by Mean Total Substitutions. However, the treatment group had significantly more errors in the GSE ≥ 2 category, with a rating of 2 or greater. The reverse was expected for the GSE ≤ 1 category, with ratings of 0–1, but this was not the case. Therefore, the answer to the first question, "Do treatment participants have more errors in which two or more phonemes match the printed word?" is yes. By this outcome variable, treatment participants applied letter/sound information to a greater extent than control participants.

How did readers in the treatment group compare to control readers in other behaviors? Were they more autonomous as they read? Table 4 reports the means and standard deviations on word accuracy, tolds, self-corrections, repetitions, and insertions along with the Wilcoxon Scores (sum of the ranks), the z scores, and the *p* values. The parametric data not only give the reader an understanding of the general tendencies of the sample but also support the decision to use non-parametrics. Groups differed in three reading behaviors: word accuracy, number of tolds, and repetitions. The treatment group was slightly more accurate than the control group, relied on the examiner less or pronounce unknown words, and had more repetitions. Thus, the treatment

TABLE 4 Sum of the ranks, Z scores, and p values for Reading Behaviors

	Treatment $N = 11$	Control $N = 12$
Mean percent of words read accurately (SD)	95.6 (2.3)	92.6 (4.2)
W (sums of the ranks)*	136.0	95.0
z score	1.7	
Mean number of tolds (SD)	0.2 (0.2)	1.7 (1.8)
W (sums of the ranks)†	79.5	151.5
z score	–2.2	
Mean percent of errors self-corrected (SD)	20.0 (14.2)	19.5 (23.1)
W (sums of the ranks)‡	119.5	111.5
z score	.6	
Mean number of repetitions (SD)	2.3 (1.0)	1.5 (1.2)
W (sums of the ranks)§	132.5	98.5
z score	1.5	
Mean number of insertions (SD)	0.5 (0.3)	0.6 (0.8)
W (sums of the ranks)¶	114.7	117.0
z score	.23	

*$p = .03$; †$p = .02$; ‡$p = .26$; §$p = .05$; ¶$p = .79$.

participants were more accurate, less likely to appeal to the examiner to pronounce words, but also more likely to repeat words or phrases. However, in other reading behaviors the treatment group did not respond as predicted. Note that the treatment group did not differ with respect to numbers of self-corrections. Thus, the treatment group was more autonomous in some ways but not in self-correcting.

This study is limited in several respects. First, the intervention phase was fourteen days; clearly, a minimal amount of time using any type of highly contrived texts is desirable. However, it is likely that the transition from the partial to full-alphabetic phase scaffolded by decodable text lasts longer than fifteen days. The need to control for actual instruction, a unique study feature, partially justified this limitation. Future studies may extend the intervention phase of the study. In addition, the small number of participants limited the study. This study does not establish the long-term influences of decodability.

IMPLICATIONS

This work enhances current knowledge about the benefits and side effects of highly decodable text in three ways. First, the findings support one of the major theoretical claims about decodable text—that it results in a greater application of letter/sound information during text reading. These findings replicate similar data from a previous study (Mesmer, 2001b). Second, the results here document the actual behaviors of reading highly decodable text. Although highly decodable text lived up to one of the theoretical assertions made about it, it did not result in readers who were necessarily more autonomous in all reading behaviors. Finally, the screening and selection of participants for this study suggested that readers must have some knowledge of the alphabetic system in order to demonstrate the behaviors found in this study. In addition, many readers in the first one-third of first grade did not need the intervention's short vowel focus and highly decodable text, at least according to the screening instrument used here.

Living up to Theoretical Claims

The major finding suggests that highly decodable text does live up to theoretical claims made about it. The phonic regularity and lesson-to-text match in highly decodable text resulted in greater application of letter/sound knowledge to recognize words. In this study, readers of highly decodable text applied letter/sound knowledge during text reading to a greater degree than treatment readers did. Although both readers of highly decodable text and less decodable text had similar numbers of substitutions, the quality of treatment readers' substitutions showed more attention to letter/sound information. Because the purpose of phonics instruction is application during text reading, the finding indicates that decodability facilitates this application.

As the primary finding of the study, this result intersected with other work. This finding was supported by a very similar study (Mesmer, 2001b) as well as by the Juel & Roper-Schneider (1985) study. In the latter study, readers of more highly decodable text outperformed a control group on a generic decoding measure at the beginning and middle of the year, whereas readers in the Mesmer study applied letter/sound knowledge more during connected reading. Thus, when readers received the same phonics instruction, as they did in both of these studies, they were likely to internalize and apply that instruction if they had exposure to more highly decodable text.

To a teacher working with readers, be they struggling or normally-developing, the implication is clear—if you want readers to generalize letter/sound information taught in a synthetic phonics lesson to connected reading, then you must give them text that has some phonic regularity and coordination with phonics instruction. In this study, the coordination with phonics instruction (LTTM) was 40% over the fifteen-day intervention. One contribution of the present study is that it examined the effects of decodable text when it matched with actual, as opposed to scripted, instruction. Thus, educators employing decodability scaffolds must make certain that the text coordinates with phonics lessons that are being delivered in the classroom. It is possible for a stand-alone phonics program and a series of highly decodable texts to operate separately and thus to have no benefit for the students. In this case, the phonics instruction may focus on one sound while the text words emphasize another.

The specific phonic regularity features that resulted in this increased application of phonics instruction are discussed in greater detail below. The highly decodable text in this study was characterized by very specific patterns. On average, about 54% of the words in this study's highly decodable texts were short-vowel c-v-c words (e.g., pin, bag, mat). This base of words coordinated with the phonics lessons. The fact that the books featured short vowel words made coordinating instruction easier. For a teacher, the easily identifiable sound focus in highly decodable text is an added advantage. One can retrieve a set of more highly decodable books, quickly identify the sound focus, and then concentrate on teaching that sound. This focused instruction is especially important for readers who are not breaking the code and need the benefits of very consistent teaching to generalize behavior.

Words Accuracy, Tolds, Repetitions, and Highly Decodable Text

In this study, the treatment group displayed reading behaviors that indicate both benefits and potential side effects of using highly decodable text. The major benefit appeared to be that the treatment group was more autonomous in its reading behaviors in two ways: it was slightly more accurate than the control group, and it relied less on the examiner to provide words. Although both groups read within acceptable word accuracy limits for instructional level text (90% or above), the treatment group was more accurate. This slight edge in word accuracy may have been the result of the treatment group using the decodability scaffolds to help it identify words. Therfore, one effect of using highly decodable text was a slight increase in accuracy; this increased accuracy occurred with readers who understood and could apply the alphabetic principle.

The second piece of evidence supporting increased self-confidence was the lower number of tolds for the treatment group. This difference was more pronounced than the word accuracy differences and seemed to indicate that the treatment group again felt more empowered by the match between the instruction that they had received and the textual scaffolds in place. The control group, on the other hand, may have been frustrated by the plethora of patterns that they confronted and more likely to give up on applying phonics instruction. Although the situation clearly seems irrational, this is the exact position we place readers in when they are trying to make sense of the alphabetic code and have little contextual opportunity to do so. In this day of heavy emphasis on phonics instruction, the study affirms the need to consider text along with phonics instruction.

If a teacher has the goal of increasing student accuracy, attention to letter/sound information, and independence, then using highly decodable text and coordinating phonics instruction may be effective for readers with basic alphabetic knowledge. The implication may be even more important for struggling readers, some of whom reportedly have this knowledge yet reach an obstacle when they need to more fully apply it (Brown, 1999; Hicks & Villaume, 2001; Mesmer, 1999). However, the children in this study did not read highly decodable text exclusively; their classroom libraries and instruction provided them with a variety of other materials. Thus, a balance of texts existed throughout the study. This implies that the use of highly decodable text need not be exclusive—in fact, exclusive use may not be desired at all.

In two ways, highly decodable text did not have the intended effects. First, the treatment group did not self-correct at greater rates. Certainly a characteristic of skilled reading is the ability to self-correct, and better decoders usually do this at higher rate (Ehri, 1998). In fact, self-correcting is perhaps the pinnacle of autonomous reading. Although the treatment group showed some independence in not relying on the examiner to pronounce words, its independence did not extend to greater self-monitoring. Both groups self-corrected about 20% of their errors, or about 1 out of 5. While this is not the optimal 1:3 ratio, it is respectable. The self-correcting behaviors may not have differed because both treatment and control groups may have made use of the scaffolds present within each of the different texts. Thus, the control group may have relied on the print/picture match, predictability, and other features to assist them as they self-corrected. On the other hand, the treatment group depended on phonic regularity and coordination with their phonics lessons.

The second effect of using highly decodable text was an increased level of repetition for the treatment group. Although neutral, this finding can be analyzed as potentially negative or positive. Repetition can be viewed as a self-monitoring behavior, one in which skilled readers participate in order to self-monitor. However, excessive repetition can accompany disfluency. When analyzed from this perspective, the finding implies the possible propensity for readers of highly decodable text to be less fluent.

As a behavior accompanying disfluency, the increased repeating behavior may be accounted for developmentally. Several researchers have documented a decreased fluency as readers initially learn to decode (Chall, 1983). Therefore, the increased repetition may have been a predictable developmental result of increased attention to decoding. The key to using highly decodable text lies in identifying the development of the reader and the teaching goals. Some developing readers actually compromise accuracy for fluency and may need to focus on more the latter temporarily at the expense of the former. If the teaching goal of greater application of letter/sound information *temporarily* outweighs the goal of fluency then the teaching compromise may be acceptable.

Sampling and Timing

The study's developmental model was not validated by this research; nonetheless, it was useful to the study in three ways. First, it assisted in setting screening standards for participants. The participants in this study had a basic understanding of the alphabetic principle, could apply this to spelling, and could recognize some preprimer words. Although the screening did not specifically verify the participants' use of initial consonants to decode, previous research has established that use of initial consonants in spelling and use of initial consonants in decoding appear simultaneously (Stahl, McKenna, Kovach, & Eakle, 2001). These participants were at least partial alphabetic readers but not operating effortlessly in consolidated alphabetic phases. Whether or not these participants were transitioning from partial alphabetic to full alphabetic reading as specified in the model is not clear. Future research should fully test and validate this model.

The second benefit of the developmental model was that it allowed for an application of study results to struggling readers regardless of grade level. Although this study took place with first graders, the participants could be located along a developmental continuum upon which readers of any age might be located. The extension to struggling readers is that given the same abilities, they may also demonstrate behaviors similar to those observed in this study.

Finally, the model has a pivotal policy implication: the key to using highly decodable text is determining students' development *along a continuum of reading behaviors*. If highly decodable text were used with readers who did not have a firm understanding of the alphabetic principle, then it is possible that the same results would not have been found. Policies stipulating the use of highly decodable text in the first grade assume that readers have the prerequisite skills previously mentioned. But in this study, five of those screened did not; in addition, seventeen would not have benefited from the intervention because their skills seemed to surpass the requisite alphabetic knowledge. Therefore, policies about decodable text should specify that assessment data be used to determine readers' development and that such data may be used in deciding whether or not decodability is an appropriate textual scaffold. Educators should keep in mind that a major goal of early reading instruction is increasingly sophisticated application of the alphabetic principle until readers are accessing words automatically. Part of this progression is reading full alphabetic reading in the first grade (Ehri, 1998). In addition, a variety of texts, including highly decodable text, should be available to teachers.

Given its short intervention length and other limitations, this work adds only a piece to the puzzle over a highly disputed material. It does suggest how readers might actually interact with decodability scaffolds. No one material, be it controlled, predictable, or decodable, will work with all readers all the time. Future research should focus on whether or not decodability results in readers who are more able to handle a range of materials containing *no* textual scaffolds as they mature. In addition, investigations of the use of highly decodable text with struggling readers should be undertaken, as well as examinations validating the study's developmental model.

REFERENCES

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, Mass: MIT Press.

Allington, R. (1997). Overselling phonics: Five unscientific assertions about reading instruction. *Reading Today, 15*, 1–15.

Allington, R. (2001). *What really matters for struggling readers*. New York: Addison Wesley Longman.

Allington, R. & Woodside Jiron, H. (1998). Decodable text in beginning reading: Are mandates and policy based on research? *ERS Spectrum, 16*, 3–13.

Ball, D. L. & Cohen, D. K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*, 6–8.
Barr, R. (1972). The influence of instructional conditions on word recognition errors. *Reading Research Quarterly, 7*, 509–529.
Barr, R., Dreeben, R., & Wiratchai, N. (1983). *How schools work*. Chicago: The University of Chicago Press.
Beck, I. L. (1981). Reading problems and instructional practices. In G. A. MacKinnon and T. G. Waller (Eds.), *Reading research: Advances in theory and practice*, (Vol. 2, pp. 53–93). Newark, DE: Academic Press.
Beck, I. L. & Block, K. K. (1979). *An analysis of the dimensions that affect the development of code-breaking ability in eight beginning reading programs*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center.
Brown, K. (1999). What kind of text—for whom and when? Textual scaffolding for beginning readers. *The Reading Teacher, 53*, 292–307.
Bryne, B. (1991). Experimental analysis of the child's discovery of the alphabetic principle. In L. Rieben and C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 75–84). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.
California Department of Education. (1996). *Guide to the California Reading Initiative of 1996*. Sacramento: Author.
Chall, J. S. (1983). *Stages of reading development*. New York: McGraw Hill.
Clark, H. C. & Clark, E. V. (1977). *Psychology and language*. New York: HBJ.
Clay, M. M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
Cole, A. (1998). Beginner-oriented texts in literature-based classrooms: The segue for a few struggling readers. *The Reading Teacher, 51*, 488–501.
Ehri, L. C. (1991). Learning to read and spell words. In L. Rieben and C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 57–74). Hillsdale, NJ: Lawrence Erbaum Associates Publishers.
Ehri, L. C. (1994). Development of the ability to read words, In R. Barr, M. Kamil, P. Mosenthall, and P. Pearson (Eds.), *Handbook of reading research*, (Vol. 2, pp. 383–417). New York: Longman.
Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*, 116–125.
Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. In J. L. Metsala and L. C. Ehri (Eds.). *Word recognition in beginning literacy* (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Ehri, L. C. & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, *14*, 135–163.
Fletcher, J. M., Francis, D. J., & Foorman, B. (1997). Only 15–20 percent. *Reading Today*, *5*, 2–18.
Foorman, B., Francis, D., Fletcher, J., Schatschneider, C., & Metha, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *The Journal of Educational Psychology*, *90*, 37–55.
Gough, P. (1997). Critical balances: Early instruction for lifelong reading. *Reading online* [On-line serial]. Available at www://readingonline.org/critical/houston/gough.
Gough, P. B. & Juel, C. (1991). The first stages of word recognition. In L. Rieben and C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 47–56). Hillsdale, NJ: Lawrence Erlbaum Associates Publishers.

Groff, P. (1999). Decodable words in reading textbooks: Why they are imperative. The National Right to Read Foundation. Available at http://www.nrrf.org/27_decode_txtbooks.htm. Accessed October 20, 2004.

Grossen, (1997). *Thirty years of research: What we know about how children learn to read*. Santa Cruz, Calif.: The Center for the Future of Teaching and Learning.

Hicks, C. P. & Villaume, S. K. (2001). Finding our own way: Critical reflections on the literacy development of two Reading Recovery Children. *The Reading Teacher*, *54*, 398–412.

Hiebert, E. (1999). *Text matters in learning to read.* CIERA Report 1-001. Ann Arbor, Mich.: Center for the Improvement of Early Reading Achievement.

Hoffman, J. V. (2001). Decodable texts for beginning reading instruction: Leadership or politics in California and Texas? *The California Reader*, *34*, 2–8.

Hoffman, J. V., McCarthey S.J., Abbott, J., Christian, C., Corman, L., Curry, C., Dressman, M., Elliott, B., Maherne, D., & Stahle, D. (1993). So what's new in the new basals? A focus on first grade. *Journal of Reading Behavior*, *26*, 47–73.

Hoffman, J. V., Roser, N. L., Patterson, E., Salas, R., & Pennington, J. (2000). Text leveling and "little books" in first grade reading. *Journal of Literacy Research*, *33*(3), 507–528.

Johnston, F. (1998). The reader, the text, and the task: Learning words in first grade. *Reading Teacher*, *51*, 666–675.

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

Kameenui, E. J. & Simmons, D.C. (1997). Decodable texts and language of dichotomy: A response to Allington. *Reading Today*, *15*(2), 18.

MacDonald, P. (1999). Power, Type I, and Type III error rates of parametric and nonparametric statistical tests. *The Journal of Experimental Education*, *67*, 367–379.

Menon, S. & Hiebert, E. H. (1999). *Literature anthologies: The task for first grade readers.* Ciera report# 1-009. Ann Arbor, Mich.: Center for the Improvement of Early Reading Achievement

Mesmer, H. A. (1999). Scaffolding a crucial transition using text with some decodability. *The Reading Teacher*, *53*, 130–142.

Mesmer, H. A. (2001a). Decodable text: A review of what we know. *Reading Research and Instruction*, *40*, 121–142.

Mesmer, H. A. (2001b). In J. Hoffman, D. Schalbert, C. Fairbanks, J. Worthy, and B. Maloch. (Eds). Examining the theoretical claims about decodable text: Does text decodability lead to greater application of letter/sound knowledge in first grade readers? *50th Yearbook of the National Reading Conference*, 444–459. National Reading Conference: Oak Creek, WI.

Meyer, L. A., Greer, E. A., & Crummey, L. (1987). An analysis of decoding, comprehension, and story text in four first-grade reading programs. *Journal of Reading Behavior, 6*, 69–98.

Moats, L. & Hall, (1999). *Straight talk about reading: How parents can make a difference during the early years.* Chicago, IL: Contemporary Books.

Muter, V. & Snowling, C. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly*, *33*, 320–337.

Nation, K. & Hulme, C. (1997). Phonemic segmentation, not onset–rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly*, *32*, 154–167.

National Reading Panel, (2000). *Report of the National Reading Panel: Teaching Children to Read.* Washington, DC: National Institute of Child Health and Human Development.

Pikulski, J. J. (1998). Business we should finish. *Reading Today, 15*(5), 30.

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

Reutzel, D. R., & Daines, D. (1987). The text-relatedness of reading lessons in seven basal reading series. *Reading Research and Instruction*, *27*, 26–35.

Routman, R. (1997). Routman thanks Allington. *Reading Today*, *15*(5), 30.

Snow, C., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children.* Washington, DC: National Academy of Sciences Press.

Stahl, S., Duffy-Hester, A., & Stahl, K. A. D. (1998). Everything you wanted to know about phonics (but were afraid to ask). *Reading Research Quarterly*, *33*, 338–335.

Stahl, S., McKenna, M., Kovach, J. J., & Eakle, J. (2001). Phonological awareness, word recognition, and spelling: Rethinking the relation. Paper presented at the Annual Meeting of the National Reading Conference, December, San Antonio, Texas.

Stein, M., Johnson, B., & Gutlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Remedial and Special Education*, *20*, 257–287.

Texas Department of Education. (1997). *1997 Proclamation of the State Board of Education Advertising for Bids on Instructional Materials.* Austin, TX: Author.

Venezky, R. L. (1970). *The structure of English orthography.* The Hague: Mouton.

Virginia General Assembly. (2001). *House Joint Resolution No. 794.* Richmond, VA: Author.

Watson, D. (1997). Beyond decodable texts—supportive and workable literature. *Language Arts*, *74*, 635–643.

Willows, D. M., Borwick, D., & Hayvren, M. (1981). The content of school readers. In G. E. MacKinnon and T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (vol. 2, pp. 95–175). New York: Academic Press.

Wilson, B. A. (1996). *The Wilson Reading System.* Millbury, Mass.: Wilson Language Training Corp.

Zimmerman, D. W. (1994). Not on the influence of distribution shape on parametric and nonparametric tests. *The Journal of General Psychology*, *121*, 391–401.