AN ANALYSIS OF TWO READING INTERVENTION PROGRAMS

How Do the Words, Texts, and Programs Compare?

ABSTRACT

In this study, the student texts and teacher guides of two reading intervention programs for at-risk, first-grade students were analyzed and compared: Fountas and Pinnell's *Leveled Literacy Intervention* (LLI) and Scott Foresman's *My Sidewalks* (MS). The analyses drew on the framework of available theory and research on beginning texts developed by Mesmer, Cunningham, and Hiebert in 2012. This framework includes attention to word-level, text-level, and program-level features. The student texts of the two programs had similar average percentages of single-appearing words and words that can elicit a mental picture (concrete words); however, LLI texts featured more repetition of words, a slightly higher percentage of highly frequent words, and a considerably higher percentage of multisyllable words. MS texts contained a higher percentage of phonetically regular words and a higher lesson-to-text match between phonics elements in teacher guides and the words in student texts. Instructional implications and future research directions are discussed.

Maria S. Murray
STATE UNIVERSITY OF NEW YORK AT OSWEGO

Kristen A. Munger
STATE UNIVERSITY OF NEW YORK AT OSWEGO

Elfrieda H. Hiebert
TEXTPROJECT AND UNIVERSITY OF CALIFORNIA, SANTA CRUZ

SINCE response to intervention (RTI) gained momentum after the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 2004, school districts looking to support their at-risk readers have been faced with myriad choices among commercially available intervention materials. Publishers of educational materials compete for billions of dollars in revenue annually

 0013-5984/2014/11404-0002 $10.00

("Pearson Leads the Pack," 2010), and educators can become overwhelmed or confused when faced with the wide variety of materials marketed for at-risk readers. The focus of this study was to identify the differences and similarities in the word-level, text-level, and program-level features of two widely used, commercially available intervention programs offered explicitly for beginning readers identified as at risk for reading failure. We compared first-grade versions of the intervention programs' texts and teacher guides, since first grade is a critical year for establishing successful beginning reading skills (Juel, 1988).

We became aware of the two intervention programs compared in this study when they first appeared in schools in which we were working as teacher educators and researchers. Initial reviews of the programs showed that they share components such as student texts and teacher guides featuring phonics lessons, but there appeared to be distinct differences in the student texts. One program's texts featured many phonetically regular words, whereas the other program's texts were labeled "leveled" and appeared to include more repetition of words and syntactic patterns. Knowing of a decades-long tradition of descriptive studies comparing attributes of classroom reading programs according to their philosophical viewpoints—meaning versus code emphasis (Beck & McCaslin, 1978; Chall, 1967/1983; Foorman, Francis, Davidson, Harm, & Griffin, 2004; Meyer, Greer, & Crummey, 1987; Stein, Johnson, & Gutlohn, 1999)—we became interested in how commercially available intervention programs used in RTI settings may be connected to these viewpoints as well. Wanting to know more about the nature of the words featured in the programs' texts and their relation to the instruction provided in the programs, we decided to inventory and compare characteristics that have been shown in the literature to be important for beginning reading acquisition. This information is valuable to the educators and researchers who choose these intervention programs from among a multitude of options available in today's marketplace. As the dimensions of variation in these programs are not completely obvious to the untrained eye, we believe reporting variations will lead to more informed choices about what each intervention can provide to at-risk beginning readers.

We begin by providing a theoretical framework that organizes the features of texts having empirical or theoretical support in the literature. We briefly examine the prominent viewpoints relating the features of beginning reading texts to both descriptive and intervention studies addressing these texts. We finish by presenting our review of the word-level, text-level, and program-level analyses of texts in these programs.

A Theoretical Framework for Comparing Beginning Reading Texts

The assumption that a particular type of text can best promote reading for all students is typified in the "reading wars" (Pearson, 2004) between those advocating a code-emphasis approach to reading instruction and those advocating a meaning-emphasis approach (e.g., Bond & Dykstra, 1967; Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998). Reflecting on reviews of research, Mesmer, Cunningham, and Hiebert (2012) suggested that the types of texts currently used for beginning reading instruction reflect mandates of state legislatures and advocacy of special interest groups more than evidence from theory or research. They argued for

a reexamination of the research and design of the texts of reading acquisition and proposed the rudimentary components of a framework to initiate this process.

The Mesmer et al. (2012) framework addresses features of text elements related to words in individual texts, words occurring across multiple texts, and program instruction of the words. Related to this work, we propose that at the word level, multiple features of words can influence instructional goals for beginners (especially those who are at risk) such as those related to letter-sound features and meaning features. Our theoretical framework also considers features present across multiple texts (the text level), since it is the "diet" of text experiences, such as the repetition of words or the introduction of new words, that has potential to influence self-teaching. Finally, the framework takes into account the program-level feature of lesson-to-text match (LTTM), which refers to the match between the phonics instruction in a program's teacher guides and the words in the texts students are given to practice reading. In this study, the rationale for this framework includes both theoretical and empirical investigations of beginning texts at each level of analysis—word, text, and program.

Prominent Perspectives on Beginning Reading Texts

In 1984, Aukerman categorized 165 approaches to beginning reading instruction as using one of three types of texts: phonics, whole word, and whole text. A review similar to Aukerman's has not been conducted in the past 25 years, but Mesmer (2008) identified two types as dominant in today's classrooms: decodable (the current term for phonics texts) and leveled (the current term for whole-text texts). Before describing these text types below, we note that phonics/decodable texts are typically aligned with a code-emphasis approach to teaching reading, where the emphasis is on teaching students to use the alphabetic code to identify words. Whole-word and whole-text/leveled texts are generally aligned with a meaning-emphasis approach to teaching reading, which emphasizes reading for meaning (recently referred to as the whole language approach, the term used throughout this article). Our focus is on present-day forms of these two text types: decodable and leveled. In the upcoming descriptions of the philosophical origins of decodable and leveled texts, we first give a brief history of whole-word texts to help in understanding the impact of state mandates on current beginning reading texts. Excerpts from the three prototypical texts appear in Table 1.

Whole-Word Texts

From approximately 1930 to 1990, when the whole-word approach dominated texts of beginning reading instruction (Hoffman, Sailors, & Patterson, 2002), stories were mainly constructed around high-frequency words (e.g., *where, yes, can, out*), using algorithms to repeat and introduce new words. It was believed that repetition of whole words enabled beginning readers to memorize useful and frequent words so that early texts could be read successfully (Aukerman, 1984). A national commission (Anderson, Hiebert, Scott, & Wilkinson, 1985) called for an end to such algorithmic control of word repetition in text, and subsequent guidelines for textbook adoption cycles of California (California English/Language Arts Committee, 1987) and Texas (Texas Education Agency, 1990) mandated that beginning reading texts be authentic,

Table 1. Prototypical Text Types and Sample Excerpts

Text Type	Sample Excerpts
Whole word	Lucky wakes up. Something is new today. Something is different. "Hurry up, Lucky," his boy tells him. "Today is the first day of school." (Herman, 2001)
Decodable (phonics)	See the fun in the tent. They skip and flip. They bump and they slip. Brent and Pam flip on a mat. Dad can flip, flip, flip. (Roberts, 2008)
Leveled (whole text)	Ant met a bird. "Can you fly?" said the bird. "No, I can't," said the ant. Ant met a bee. "Can you buzz?" said the bee. "No, I can't," said the ant. (Dobeck, 2009)

not controlled. The whole language approach—as these mandates came to be known—lasted only a single adoption cycle, until it was overturned to mandate decodable text (described next). Interestingly, the backlash against controlling whole-word repetition remained, as up to 70% of the words in beginning reading programs published after the decodable text mandate appeared only a single time (Foorman et al., 2004). The only texts in the current marketplace that retain a trace of an emphasis on whole-word repetition of high-frequency words are those sold by trade book publishers (e.g., the *I Can Read* series published by HarperCollins), an excerpt of which appears in Table 1. High-frequency words in texts such as *Lucky*, however, are neither presented sequentially nor repeated according to the algorithms of the earlier period.

Decodable Texts

Decodable texts represent the latest in a long history of phonics texts emphasizing the alphabetic code. The emphasis in such texts is the use of letter and sound knowledge to decode unknown words (Mesmer, 2008). The labels for these texts have been many: phonics, association of sound and spelling, phonological linguistic, phonic-word, and the modified linguistic approach (Bond & Dykstra, 1967). The term "decodable text" was not used until the late 1970s (Google ngram viewer; Michel et al., 2011), and in the early 1990s, it emerged in the educational literature in reference to texts having a high proportion of phonetically regular words matched to common letter-sound relationships previously taught in phonics lessons within accompanying teacher guides (Mesmer, 2008). Today the term "decodable text" is also used for texts with high percentages of phonetically regular words that do not match to a set of instructional lessons (see, e.g., Decodable Books Written by Teachers, 2013).

Leveled Texts

Most beginning reading texts are arranged or leveled by one or more criteria of difficulty, including some publishers' sets of decodable texts. Thus, the word "lev-

eled" may be used in reference to multiple kinds of texts, but the term "leveled text" is known in the marketplace (Afflerbach et al., 2011; Bear et al., 2009; Beck, Farr, & Strickland, 2008) as texts leveled according to a gradient of difficulty based on multiple supportive features of the *whole text*, which allow for an emphasis on meaning, such as text structure, themes and ideas, or language and literary features (Fountas & Pinnell, 2012). Leveled texts typically feature naturalness of language, close picture-text matches, and predictable text structures (Cunningham et al., 2005).

The theory behind the whole-text perspective of leveled texts is that when students read the repetitive and predictable sentences, words, and phrases in a text such as "This is the House that Jack Built," they experience "wholebooksuccess" (Martin & Brogan, 1971), which in turn motivates them to reread the texts, and when they do, they attend to the meaning of the words, how words go together, and their letter/sound patterns. According to Goodman (1969), children rely on the three-cueing system to solve unknown words—semantics (meaning), their knowledge of syntax, and the visual features of words.

Leveled texts were originally part of a one-on-one tutoring intervention called Reading Recovery implemented in the United States in the mid-1980s (Pinnell, Huck, & DeFord, 1986). These texts were from collections having numerous texts, which did not need to be used in a prescribed order. As these texts became increasingly popular and new texts were developed, a need for assigning them to levels was recognized. A leveling system developed for use in Reading Recovery's contexts (Peterson, 1988) was adapted for general classroom use by Fountas and Pinnell (1996). Leveled texts are now a dominant and central part of American reading instruction, in both regular and special education classrooms (Cunningham et al., 2005).

Research on Beginning Reading Texts

Decodable and leveled texts have been included in two types of studies: descriptive studies that have compared and analyzed the relationship between words in texts and corresponding phonics lessons, and intervention studies examining the influence of text types on students' reading. Results from both groups of studies helped to inform our research questions as well as the interpretation of our findings.

Descriptive Studies Comparing Beginning Reading Texts and Phonics Lessons

Descriptive studies have typically focused on lesson-to-text match (LTTM): the match between the instruction of phonics elements in teacher guides and the words in student texts (Stein et al., 1999). Such a focus began with Chall's (1967/1983) analyses of four first-grade reading programs: two code emphasis and two meaning emphasis. Chall observed that the teacher guides of the meaning-emphasis basal programs included phonics instruction; however, the phonics elements taught did not systematically match the words in students' texts as they did in the code-emphasis programs.

For each of the four decades following Chall's (1967/1983) work, researchers have analyzed and compared LTTM in meaning- and code-emphasis first-grade reading programs, and, as a result, shifts in various copyrights are evident. In reading programs copyrighted in the 1970s, Beck and McCaslin (1978) reported that patterns of LTTM had not changed from those reported by Chall (1967/1983) and noted that the

code-emphasis programs provided a higher "potential for accuracy" when decoding words, whereas the LTTM of meaning-emphasis programs did not. Four copyrighted programs of the 1980s were analyzed by Meyer et al. (1987), who noted that meaning-emphasis programs continued to have low LTTM. Three out of four of the programs analyzed were meaning-emphasis, and their LTTM was less than 10%. Stein et al. (1999) found that decodable texts and lessons mandated for adoption in California and Texas in the 1990s featured LTTMs similar to the meaning-based programs analyzed by Beck and McCaslin (1978). Finally, in an analysis of six programs with copyrights from 1995 to 2000, Foorman et al. (2004) reported much variability in LTTMs, with only two programs approaching the state mandates of 75% (California) and 80% (Texas) decodability.

Intervention Studies Using Beginning Reading Texts

There is extensive intervention literature in which either decodable or leveled texts have been used in intervention contexts (Blachman et al., 2004; Ehri, Dreyer, Flugman, & Gross, 2007; Hiebert, Colt, Catto, & Gury, 1992; Menon & Hiebert, 2005; Torgesen et al., 2001); however, studies in which the focus was on isolating and measuring the influence of specific text types on beginning reading are few. Cheatham and Allor (2012) sought to establish effects of decodability on beginning reading by reviewing studies that attempted to isolate the effects of decodable text. Results of these studies were inconsistent, with positive effects reported for word reading accuracy or reading fluency in some studies but not in others. Cheatham and Allor's observations are useful when considering text features for at-risk readers during first grade. They identified the need to consider text features in relation to particular stages of students' reading development. Referring to Ehri's (2005) four phases of sight-word learning, they noted that decodable texts may be more critical in the second stage—partial alphabetic—when students begin to use some letter names and sounds to read words, and in the third stage—full alphabetic—when more complete knowledge of letter-sound correspondences is used. Such an observation suggests that the influences of text types may also lie in the particular developmental phases of beginning readers.

The Current Study

Absent from the literature altogether is research specifically analyzing and comparing words in texts and their accompanying phonics instruction from commercially published intervention programs. Because intervention programs loom large due to RTI—and promise to loom even larger as predictions of student performance on the assessments of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) are made (e.g., Loveless, 2012)—a focus on the texts of intervention programs is critical.

In the current study, two commercially available first-grade intervention programs were compared. *Leveled Literacy Intervention* (LLI; Fountas & Pinnell, 2008) features leveled texts and is currently used in all 50 states in the United States (Lori Lampert, personal communication, June 28, 2012). *My Sidewalks on Reading Street* (MS; Juel, Paratore, Simmons, & Vaughn, 2008) features decodable texts and is used in over 1,500 districts in all 50 states (National Center on Response to Intervention,

2013). These programs are similar in that both offer instruction in the major literacy components such as phonics, along with practice reading and rereading an assigned text. The framework of word-level, text-level, and program-level analyses emerged from the features outlined below.

Word-Level Features

Word-level features are those pertaining to the nature of the words themselves. Features of words in texts of beginning reading are critical, since independent, accurate, and automatic word recognition is essential to proficient reading (Adams, 1990; Cunningham, Nathan, & Raher, 2011; Tunmer & Nicholson, 2011). The importance of these features has been supported by research using computerized simulations representing the development of word recognition. Connectionist model researchers who create and test these representations have found that words' phonological, orthographic, and semantic characteristics, along with frequency, familiarity, and imageability, are all involved in determining readability (Plaut, 2005; Plaut & Shallice, 1991). In this study we measured four types of word features from the framework of Mesmer et al. (2012): word decodability (words that adhere to common spelling-sound patterns), highly frequent words (words occurring frequently in print), word concreteness (words that can elicit a mental picture), and multisyllable words.

Because English is an alphabetic language, extensive research validates that phonological awareness, alphabetic knowledge, and word reading each strongly predict later reading ability (National Reading Panel, 2000). Word decodability (phonetic regularity) serves as a scaffold for beginning readers in that students can apply knowledge of letter-sound relationships to words by either attending to individual letters or by identifying known word parts (Ehri, 2005).

English words do not always have regular correspondences between letters and sounds (i.e., many speech sounds can be spelled in different ways). These irregularly spelled words tend to occur with great frequency in English (e.g., *have, of, they*), and in texts read by beginning readers, they account for more than 50% of words (Adams, 1990). A reader's word-reading accuracy and automaticity are positively influenced by the frequency with which words appear in print (Gernsbacher, 1984; Zinna, Liberman, & Shankweiler, 1986). Texts should feature enough highly frequent words so that they can be learned; however, research has not established guidelines indicating how many of these words should be included or how often they should be repeated within texts (Mesmer et al., 2012).

Another word-level feature that can influence the ease with which children learn to recognize words is imageability. Imageability is the degree to which a word's meaning, if known to the reader, elicits a mental picture (Paivio, Yuille, & Madigan, 1968). This makes a word such as "butterfly" easier to learn and remember than "hopeful" because it conjures a mental picture, which in turn activates response retrieval (Paivio, 1969). To operationalize imageability of words in texts in this study, we measured word concreteness, which is related to but not synonymous with imageability. For example, the word *fiddlehead* is concrete because it has the potential to elicit a mental picture, but it may not be imageable if the reader does not know what a fiddlehead is. Descriptive studies of word-level features in texts can only measure word concreteness, since there is no way of knowing what words readers know.

The last word-level feature in this study is multisyllable words. Multisyllable words in texts are often not a focus in the word learning of beginning or at-risk readers because an implicit assumption in word recognition acquisition is that, at least in English, single-syllable words are learned before multisyllable words. One exception is the role of multisyllable words in the "prealphabetic" or first phase of Ehri's (2005) model of sight-word learning, where children learn words on the basis of visual characteristics. The word *dinosaur*, for example, may be remembered because of the height of the first letter ("a dinosaur has a long neck") and the length of the word ("a dinosaur has a long body"); however, this strategy becomes ineffective as children are introduced to more words. At this point, they need to apply their knowledge of letters and sounds to read words. Although scholarship on multisyllable words in beginning reading texts is scant, patterns in Table 2 depict changes coinciding with textbook mandates in the last several decades.

Text-Level Features

While word-level features are important to the texts at-risk children read, text-level features need to be considered as well. Text-level features relate to words within the context of other words in a program, such as how often words repeat within or across particular levels (e.g., type-token ratio) or appear only a single time (i.e., singletons). In learning theory (e.g., Bruner, 1966; Ericcson, Krampe, & Tesch-Römer, 1993), the deliberate pacing and repetition of content can maximize learning or, if not adequately controlled, may explain a lack of it. We chose to measure text-level features from this theoretical basis and from limited research on the benefits of word repetition in texts (Juel & Roper/Schneider, 1985; Reitsma, 1983).

From a cognitive perspective, both word features and children's existing proficiency determine appropriate pacing and repetition of new content. More proficient children may be equipped to handle a faster pace at which new content appears and can therefore engage in more "self-teaching" (Share, 1995) compared to struggling readers. High repetition may benefit struggling readers the most, whereas a quicker introduction of novel words may be suitable for more proficient readers. Research on self-teaching typically focuses on children in their second or third year of reading

Table 2. Percentages of Repetition and Multisyllable Words in Reading Programs from Five Decades[a]

Copyright	Grade/Level	Total Words per Text	Type-Token Ratio[b]	Singletons[b]	Multisyllable Words[b]
1962	Grade 1/Entry	18	10	0	6
	Grade 1/Exit	378	8	7	23
1983	Grade 1/Entry	144	5	5	8
	Grade 1/Exit	481	10	16	27
1993	Grade 1/Entry	79	29	46	26
	Grade 1/Exit	385	20	41	41
2000	Grade 1/Entry	83	21	40	9
	Grade 1/Exit	334	19	40	34
2007	Grade 1/Entry	131	25	38	9
	Grade 1/Exit	352	26	50	34

[a] Information for 1962, 1983, 1993, and 2000 copyrights from Hiebert (2005).
[b] Type-token ratio, singletons, and multisyllable words are calculated on levels of 10 texts.

instruction—not on novices, since there needs to be a modicum of acquired information and knowledge before students can self-teach (Cunningham et al., 2011; Ehri, 1998). With this in mind, it may be safe to assume that intervention programs may need to be more repetitive and more conservative in the introduction of new words compared to nonintervention programs, since students who struggle are less likely to be able to self-teach.

The use of learning theory provides a rationale for why educators and developers of reading programs should consider the pacing and the introduction of new content, since empirical evidence is limited on the effects of these factors. Pacing and repetition of content are usually not addressed explicitly in frameworks and theories of text design or selection today. An example of this, applied to the code-emphasis approach to reading, pertains to California's most recent mandates (California State Board of Education, 2006) requiring teachers to provide two texts per each taught phonics element in first grade; however, no prescription is provided regarding repetition of spelling-sound correspondences or the pace at which texts are presented. Within the whole-language approach, the theory is that repeated reading will lead to proficiency in word recognition, although the research base supporting this view is scant and even contradictory (see, e.g., Johnston, 2000). Even though optimal levels of word repetition are not known, a lack of knowledge about pacing and repetition does not mean that these variables are not theoretically important, but without empirical evidence backing up the theories, there is little to guide program designers or publishers.

Because of the potential importance of word repetition in beginning reading texts, we analyzed two measures of the pace at which new content is presented: type-token ratio and the percentage of singletons within text levels. We also compared current levels of repetition with baseline data from analyses of previous copyrights of first-grade core programs (see Table 2).

Program-Level Feature

Program-level features pertain to the relation between the instruction featured in a program's teacher guides and texts students read. In particular, the match between the content of the phonics instruction in teacher guides and the words in student texts (LTTM) has been a focus since Chall's (1967/1983) reviews. With respect to LTTM, we examined the programs' adherence to guidelines such as those proposed by Beck (1997), which were based on her earlier reviews of first-grade texts and programs (Beck & Block, 1979; Beck & McCaslin, 1978). She recommended 70% to 80% decodability, since only 30% to 50% may not provide beginning readers with enough opportunity to practice what they had learned, and 100% would result in a stilted, artificial-sounding text. It should be noted that Beck's work was descriptive rather than empirical—student performance was not connected to the match of lessons and student texts.

Research Question and Hypotheses

The overarching research question guiding our analyses was, how do LLI and MS compare at the word level, text level, and program level? A review of the literature regarding philosophical foundations of reading and previous empirical analyses

helped us to hypothesize what differences and similarities would be expected between LLI and MS. A secondary question was, how do some of these features correspond to programs of previous decades?

In terms of word-level and program-level features, we hypothesized that MS texts would feature a greater percentage of phonetically regular words and a greater LTTM compared to LLI texts, since texts designed to promote the development of decoding skills by definition have greater regularity in letter-sound relationships (Mesmer, 2005). We also hypothesized that LLI texts would feature a greater percentage of highly frequent words, concrete words, and multisyllable words. A greater percentage of highly frequent words was predicted to occur in LLI leveled texts, because they are founded in the history of "wholebooksuccess" (Martin & Brogan, 1971), and a greater percentage of concrete words was predicted to occur in leveled texts in which the theory of the three cueing systems underlies their design. LLI texts were also expected to feature a greater percentage of multisyllable words compared to MS texts, since texts featuring more concrete words deliberately matched to pictures (e.g., *pancakes, rooster, blueberries*) emphasize semantic cues over orthographic cues.

With respect to text-level features (e.g., type-token ratio and singletons), we did not generate specific hypotheses as to whether either LLI or MS texts would have a higher or lower percentage of repetition. As evident in Table 2, type-token ratios and percentages of singletons in text have increased in the anthologies of core (formerly called basal) reading programs in the past two decades. There is neither prior research nor theory to suggest that either text type would have more or less repetition than texts in core reading programs.

Method

Program Selection

In the current study, we compared first-grade versions of *Leveled Literacy Intervention* (LLI) and *My Sidewalks* (MS), which are used with small groups of at-risk readers. The programs include instruction and text reading practice in 30 minutes of daily lessons designed to supplement (i.e., be taught in addition to) classroom English language arts instruction. LLI is designed to be delivered for approximately 18 weeks, whereas MS is delivered for up to 30 weeks. LLI includes 100 required lessons (there are 10 optional "getting started" lessons that were not included in this analysis), along with 100 corresponding leveled texts, whereas MS includes 150 required lessons, along with 90 corresponding decodable texts. Both programs include instruction in phonemic awareness, phonics, fluency, vocabulary, comprehension, and writing. Our analysis of word-level (within words), text-level (between texts), and program-level (LTTM) features included only the texts read by students and the phonics and word work instruction from the teacher guides, which is similar to procedures used by Stein et al. (1999) in their comparison of the components of first-grade reading programs.

Measures

Word-level features. For word-level features, all words from the student texts were entered into a word-processing program and then grouped into 10 levels for each program, resulting in 10 word lists for LLI and 10 word lists for MS. From these

lists, the total number of words at each level was established (i.e., all words), along with total unique words (i.e., words that are counted the first time they occur in a level only—any subsequent occurrences are not counted), total multisyllable words, and total singletons at each level (i.e., words that appear only a single time). Word-level features were obtained, including concrete words, highly frequent words, multisyllable words, and phonetically regular words, by dividing words with a given feature (e.g., phonetic regularity) by the total words at that level and multiplying by 100. For example, if a level included 183 phonetically regular words in a level with 359 total words, then the percentage of phonetically regular words was $183/359 \times 100 = 51\%$. Percentages of words that are phonetically regular, highly frequent, and concrete were obtained using the Text Analysis: Beginning Books (TABB; see Hiebert, 2011) digital program, since words with these features were not able to be counted directly from our lists. TABB contains reference lists of words based on previous research that specify the degree of phonetic regularity, frequency, and concreteness of words based on sources listed in Table 3. This study analyzed the concreteness of words, consistent with available databases (e.g., Graesser, McNamara, & Kulikowich, 2011; Paivio et al., 1968).

Text-level features. Next, text-level percentages for singletons and type-token ratios were established. The percentage of singletons was obtained by dividing the total number of singletons by the total of unique words (not total words, as with word-level features) and multiplying by 100. Type-token percentages were obtained by dividing unique words by the total words in each level and multiplying by 100.

Program-level feature. For the program-level feature of lesson-to-text match (LTTM), phonics and word work lessons from LLI and MS teacher guides were reviewed, and corresponding phonics elements were inventoried. Cumulative lists of phonics elements for LLI and MS were created so that words in texts could be matched to them (e.g., short "a," silent "e" patterns). Then, beginning with the first text in each program, two raters examined every word in each text and marked whether the words are decodable based on the phonics element(s) taught in that lesson or in previous lessons. Two raters worked independently, circling each word in the texts that matched phonics elements taught. Raters compared their matches, and in cases of disagreement, word match was discussed and consensus was reached before matches were totaled. Because of the laborious and error-prone nature of the

Table 3. Word-Level Features and Text Analysis: Beginning Books (TABB) Procedures

Word-Level Feature	TABB Procedure
Phonetically regular	TABB's database consists of 2,650 single-syllable words with the most common vowel patterns, based on Bloomfield and Barnhart (1961), Fries (1963), and Fry (1990) lists, which contain a short or long vowel and occur in word families where at least two of the word family members are frequently found in print (e.g., *cat, sat*).
Highly frequent	TABB's database consists of 2,046 words that begin with the 930 words identified by Zeno, Ivens, Millard, and Duvvuri (1995) as occurring 100 times or more per million word sample, along with derivatives identified by Thorndike (1921; e.g., plurals, past tense, comparatives, superlatives).
Concrete	TABB's database consists of 1,200 words derived from 10 sources of familiar/concrete words for young children (e.g., Biemiller, 2009; Johnson, Moe, & Baumann, 1983) or that identify key concepts for young children, such as farms, automobiles, and oceans (e.g., DK Publishing, 2008; Scarry, 1985).

counting task, a third person checked final counts for accuracy. LTTM was then calculated for each text by dividing the number of decodable words by the total words in each text. Finally, the average LTTM of all texts within a level was calculated (see Table 4).

To determine the consistency with which the two raters matched words in texts to phonics lessons taught (LTTM), each rater's number of matches was compared across entire programs as well as for each lesson level. Both raters were highly consistent in matching words and phonics elements for both programs: LTTM for LLI and MS was highly similar across both raters (LLI: Rater 1 = 31.2%, Rater 2 = 30.8%; MS: Rater 1 = 68.3%, Rater 2 = 68.2%). In addition, agreement across the 10 levels for each program was also highly consistent: LTTM differed at only two of the 10 levels for LLI, and at only one level for MS, and in each case differences were equal to or less than 1%.

It should be noted that we needed to be more liberal in deciding whether a phonics element was taught in LLI compared to MS because the nature of the phonics lessons in the programs is so different. When a phonics element is introduced and practiced in MS, letter names and their associated sounds are explicitly taught. This made matching the phonics elements with the words used in MS texts relatively straightforward. In LLI phonics lessons, the sounds associated with vowel letters are not always explicitly taught. For example, although letter "o" is introduced in Lesson 19 in the word *got*, the sound that letter "o" represents is not explicitly taught. In this lesson, students are instructed to watch the teacher write letters "g," "o," and "t" in a sequence of three boxes. The suggested instructional language includes the teacher only saying the sound made by letter "g"—not the sounds of letters "o" and "t." In these instances, we still considered the vowel to be "taught," and later words featuring the short "o" sound were counted in LTTMs, provided the other letters and sounds in the words were also taught. This approach is similar to Beck and McCaslin's (1978) comparisons of code-emphasis and basal programs. They defined a phonics element as being taught if it received "explicit attention," which included either explicit or indirect instruction. We took a similarly liberal approach to determine the LTTM for LLI texts in order to maximize potential matches.

Table 4. Percentages of Lesson-to-Text Match for *Leveled Literacy Intervention* and *My Sidewalks*

Leveled Literacy Intervention			*My Sidewalks*		
Lesson Range	LTTM Range	Mean LTTM	Lesson Range	LTTM Range	Mean LTTM
Lessons 11–20	0–30	4	Lessons 1–15	0–76	28
Lessons 21–30	0–41	18	Lessons 16–30	41–84	68
Lessons 31–40	9–40	22	Lessons 31–45	64–77	72
Lessons 41–50	19–50	30	Lessons 46–60	68–80	75
Lessons 51–60	24–52	33	Lessons 61–75	70–79	74
Lessons 61–70	23–47	39	Lessons 76–90	67–80	73
Lessons 71–80	26–51	35	Lessons 91–105	68–81	74
Lessons 81–90	28–46	37	Lessons 106–120	69–80	74
Lessons 91–100	34–53	43	Lessons 121–135	65–84	70
Lessons 100–110	42–62	50	Lessons 136–150	67–81	76
Total mean		31			68

Note.—LTTM = lesson-to-text match.

Results[1]

Word-Level Results

Similarities and differences in word-level features for LLI and MS texts are reported in Table 5. With respect to average percentages of concrete words, LLI and MS texts are similar (i.e., a difference of less than 5%) at each level (LLI = 30%, MS = 28%). On average, slightly more than one-quarter of the words in LLI and MS levels are "easily imagined." MS texts also have more variability across levels compared to LLI (LLI = 25% to 34%, MS = 19% to 39%).

With respect to average percentages of highly frequent words, LLI and MS texts differ slightly (i.e., a difference greater than 5% but less than 10%) at each level (LLI = 66%, MS = 59%). The practical significance of these differences is not known, but it is notable that both programs feature texts in which the majority of words (above 50%) are highly frequent. LLI texts include a similar percentage of highly frequent words (LLI = 64% to 70%) at each level, compared to MS texts, which were more variable (MS = 51% to 64%).

With respect to the average percentage of phonetically regular words, LLI and MS texts differ more extremely (i.e., a difference greater than 10%) at each level (LLI = 42%, MS = 62%). MS texts appear to consistently make greater use of phonetically regular words compared to LLI texts across all levels (LLI = 25% to 51%, MS = 51% to 71%). More substantial differences were also found with respect to the average percentages of multisyllable words found at each level (LLI = 23%, MS = 11%). Additional analyses at each program level revealed that the first three levels of LLI texts feature an average of 23% multisyllable words in running text, whereas the first three levels of MS texts feature only 6%. Conversely, the last three levels of LLI texts feature an average of 12% multisyllable words, whereas the last three levels of MS texts feature 17%. This indicates that early LLI texts contain many multisyllable words, which decrease over the course of the intervention by 48%[2], whereas MS texts begin with a relatively low percentage of multisyllable words and increase by 180%.

Table 5. Percentages of Word-Level and Text-Level Features in *Leveled Literacy Intervention* (LLI) and *My Sidewalks* (MS) Texts

	Word-Level Features								Text-Level Features			
	Highly Concrete		Highly Frequent		Multisyllable		Phonetically Regular		Singletons		Type-Token Ratio	
Level	LLI	MS	LLI	MS	LLI	MS	LLI	MS	LLI	MS	LLI	MS
1	31	19	65	59	33	7	51	51	68	45	26	27
2	34	24	64	54	20	9	39	66	58	38	22	24
3	25	26	70	54	16	2	46	69	52	37	23	27
4	29	39	68	51	11	6	47	71	43	39	20	25
5	31	31	67	62	18	10	44	67	33	33	17	24
6	31	33	66	59	20	12	44	67	39	44	20	29
7	29	25	64	64	19	13	42	63	37	41	19	25
8	31	29	66	64	17	15	39	54	38	49	17	32
9	28	28	67	63	8	16	40	57	32	38	17	24
10	26	26	64	62	10	20	25	56	46	47	21	29
Total	30	28	66	59	23	11	42	62	45	41	20	26

Note.—In *My Sidewalks*, each level contains nine texts. In *Leveled Literacy Intervention*, each level contains 10 texts.

Text-Level Results

Similarities and differences in text-level features for LLI and MS are also reported in Table 5. With respect to the overall mean percentages of singletons contained at leach level, text-level features are similar (LLI = 45%; MS = 41%), although further comparisons of LLI's first three and last three levels reveal a marked decrease in number singletons (LLI = 59% to 39%). With respect to type-token ratio, text-level features for LLI and MS are slightly different (LLI = 20%; MS = 26%). Both programs demonstrate relative stability across levels (LLI = 17% to 26%; MS = 24% to 32%); however, MS texts do contain slightly less word repetition.

Program-Level Results

Lesson-to-text match (LTTM) percentages between phonics lessons in each program and corresponding texts are reported in Table 4. For the first level of each program, LTTM is low (LLI = 4%, MS = 28%) when compared to program averages (LLI = 31%, MS = 68%); however, at later levels, more distinct differences emerge. By the beginning of the second lesson group in MS (Lessons 16–30), the majority of texts contain words matching phonics patterns taught. By this point in the program, students have been taught a total of 13 consonant sounds and five short vowel sounds. When these phonics elements are paired with texts, LTTM approaches 70% and exceeds it thereafter. Conversely, it is not until Lesson 18 (Lesson 8 in our analysis) that LLI incorporates explicit use of a vowel sound, beyond the first 10 "getting started" lessons that are not taught to all students. Theoretically, then, in Lessons 1–17, phonetically regular words may not be decodable for students who do not know their vowel sounds, since all words must contain at least one vowel sound. When the short "a" sound is introduced in Lesson 18, the associated LLI text becomes the first with an LTTM exceeding 0% (*Getting Dressed* = 30%).

Results showed marked differences between the programs' LTTM averages, with MS texts ($M = 68\%$) more aligned with phonics lessons compared to LLI texts ($M = 31\%$). In addition, when the first level of each program was eliminated (due to discrepancies from other levels), standard deviations for the remaining levels were 9.93% for LLI and 2.52% for MS. This means that after the first set of lessons, MS phonics lessons and texts remain aligned (i.e., as the phonics lessons become more complex, the words in the texts adjust accordingly), whereas LTTM percentages for LLI start low and trend upward to a maximum of 50%.

Discussion

This study provided much-needed systematic analyses of word-, text-, and program-level features of two widely used intervention programs in the United States. The discussion of the current findings is exclusive to these programs and should not be generalized to other programs. This study is in line with a history of analyses comparing meaning- and code-emphasis attributes in first-grade reading programs and texts (Beck & McCaslin, 1978; Chall, 1967/1983; Foorman et al., 2004; Meyer et al., 1987; Stein et al., 1999); however, no analysis of first-grade commercially published intervention programs has even been completed. These programs are used to teach beginning readers who are at risk or struggling, and gaining knowledge of the differences and similarities in these programs can launch additional research to clarify

their impact on at-risk readers, as well as research on how to better design effective interventions. Differences or similarities in interventions not obvious to the untrained eye will also assist educators in recognizing which aspects align with their instructional goals.

The results of this comparison of intervention programs correspond with some notable traditions and viewpoints in beginning reading instruction. The LLI program appears to be in alignment with a meaning-emphasis philosophy. Comparisons at the word and text levels revealed that LLI features more highly frequent words and more multisyllable words, thus confirming our initial hypotheses. This makes sense in a program emphasizing meaning, semantic cues, natural language patterns, predictable syntactic patterns and word repetition, and a philosophical foundation underlying texts promoting "wholebooksuccess" (Martin & Brogan, 1971). We also hypothesized that LLI leveled texts would have a higher percentage of concrete words than MS decodable texts, but this did not turn out to be the case. Although we did not have hypotheses related to the text-level features of repetition, we found a slightly lower type-token ratio in LLI compared to MS. This is likely due to the repetition of syntax and words in LLI, which matches the characteristic of texts with a foundation in whole language traditions. Also related to repetition, the average percentage of singletons in LLI texts was similar to the percentage in MS texts; however, LLI texts were distinctly more variable.

The MS program appears to be in alignment with a code-emphasis philosophy. MS texts feature more phonetically regular words, fewer multisyllable words, and a high level of LTTM, thus confirming our hypotheses related to these features. The alignment of letters and sounds in both the texts and the instruction appears to be consistent with the perspective of a code-emphasis program. By definition, decodable texts include a high percentage of phonetically regular words matched to texts and fewer multisyllable words because multisyllable words are more difficult to decode.

A particular contribution of the present study is the attention to multisyllable words, a feature of beginning texts that is infrequently addressed in theory and research on intervention for at-risk readers. LLI texts had a relatively high percentage of multisyllable words at beginning levels and lower percentages at the final levels (33% to 10%). This shift is explained by early LLI texts' inclusion of many multisyllable words that are matched to pictures such as *pancakes, rooster*, and *blueberries*. The high percentage of multisyllable words in LLI texts may serve to assist students in the earliest stages of reading (e.g., Ehri's prealphabetic phase) when they do not yet know alphabet letters or sounds (Ehri, 1998, 2005). This is because visual aspects of the words, such as their shapes, may provide a beneficial scaffold. An alternative explanation is that the pictures serve as the scaffold in supporting beginning readers, rather than the length or appearance of the word.

The pattern of multisyllable words was reversed with the MS texts, with percentages increasing throughout the program (7% to 20%). This pattern likely reflects the influence of LTTM in the design of MS texts. That is, multisyllable words are not included in MS texts until phonics elements to read them have been taught and single-syllable words have been read successfully—for example, instruction on compound words, *-es* endings, multisyllabic words with two consonants in the middle, or consonant + *le* endings followed by the reading of words such as *pancake, gashes, napkin*, and *apple*. Each program's philosophy may serve to justify inclusion or exclusion of multisyllable words. Further research is required to establish how these patterns influence the independent word-recognition proficiencies of at-risk readers.

Although MS had a higher LTTM than LLI, it bears mentioning that the method of measuring LTTM in this analysis for both programs was derived using an assumption of "once taught, then learned." At the present time, no research validates that a phonics element that is taught will actually be learned. Mesmer (2010) emphasized this point, arguing that a true LTTM requires coordination between the reader's knowledge and the words in texts and not necessarily just the lessons and the words in texts. LTTM also assumes that each letter-sound pattern requires the same level of instructional attention and student practice (California Department of Education, 2006), even though letter-sound patterns vary enormously in the number of times they appear in written English (Venezky, 1967). Analyses that treat each letter sound equally, such as ours, may mean that some infrequently occurring patterns that are introduced too early may be to the detriment of students' acquiring letter-sound patterns that have higher usefulness in decoding words (e.g., learning the short sound of letter "a" that occurs in a lot of words will be more useful than learning the sound of letter "x" that occurs less frequently).

When LLI and MS texts are compared to texts from reading programs used since the 1960s (see Table 2), LLI texts most closely match those used during the 1990s, when the pace of content (e.g., introduction of new words) doubled over previous eras and proportions of singletons and multisyllable words increased two- to three-fold. MS texts appear most closely matched to entry-level grade 1 texts of 2007, when proportions of these same features were also high compared to earlier eras, but the proportion of multisyllable words was lower.

It is interesting to note that similar to previous descriptive studies comparing the LTTM of meaning- and code-emphasis first-grade classroom reading programs (Beck & McCaslin, 1978; Chall, 1967/1983; Foorman et al., 2004; Meyer et al., 1987; Stein et al., 1999), our findings show the meaning-emphasis program that included phonics instruction (LLI) featured much less of a systematic match to the words in students' texts than the code-emphasis program (MS) did. It appears that philosophical foundations continue to exert influence on interventions, much like they influenced the construction and content of first-grade whole-class reading programs in previous decades.

Findings related to type-token ratio and other text features need to be viewed from the perspective of our unit of analysis—a set of texts rather than a single text. Our choice of reporting results for levels of texts rather than individual texts reflects the nature of instruction and learning. Numerous texts are part of the learning-to-read process, particularly with at-risk students. When a program consists of 90–100 texts (as was the case with both programs), a report of the features of individual texts can be unwieldy. In interpreting data, however, it is critical to remember that readers do experience texts one at a time. If, for example, the mean type-token ratio for a level is in the vicinity of 10%, the ratio will be substantially higher for some of the individual texts. In other words, individual texts may present even more of a challenge for beginning readers than the view garnered from the findings of levels of texts.

Additional Instructional Implications

What do differences mean for learning to read—and for the teaching of reading? These are critical questions, especially for beginners who specifically are at risk. These are the children for whom quality of schooling may make a big difference, especially if they have teachers who do not have experience recognizing the ramifications of

differences in texts and programs (e.g., too much word variability). Although it may be argued that differences between programs such as those found in this study are not surprising, it is critical to know the values of these differences so that logical and practical intervention decisions can be made. Educators and publishers must recognize that there may be advantages in some of these differences.

With respect to findings at the word, text, and program levels of our framework, a "diet" of LLI texts may provide numerous benefits. The majority of words in the texts are highly frequent, which may enable the development of a sight-word vocabulary. LLI texts also have a large proportion of multisyllable words at the beginning levels, allowing students to successfully read words at the prealphabetic stage. Finally, LLI texts feature a relatively high repetition of words, which may lead to increased sight-word vocabulary. A diet of LLI may be inadequate in reference to the following: percentages of multisyllable words remain high even after students have presumably moved into the alphabetic stages of reading, which may encourage guessing at words or overreliance on picture cues. A low percentage of phonetically regular words (and LTTM) may discourage the use of decoding strategies, and the high percentage of singletons may cause difficulty for at-risk beginners who cannot "rely on either semantic supports or repetition to identify or retain meanings of words in texts" (Mesmer et al., 2012, p. 241).

A diet of MS texts may provide benefits as follows: the majority of words in the texts are highly frequent, which, like LLI, may help enable the development of a sight-word vocabulary, although the percentage is lower than LLI and we still do not know the implications of these differences in practice. Percentages of multisyllable words that are low, along with high percentages of phonetically regular words and a high LTTM, may especially encourage the use of what is learned in phonics lessons (Juel & Roper/Schneider, 1985). A diet of MS may be inadequate in reference to the following: a high percentage of singletons and relatively low percentage of word repetition may cause difficulty for at-risk beginners because the introduction of too many new words (and insufficient word repetition) may inhibit development of a sight-word vocabulary (all things being equal).

Because intervention texts, like those included in LLI and MS, are used to supplement classroom instruction, educators working with at-risk students must also consider the types of texts students read in classrooms. If leveled texts are used during classroom instruction and students continue to struggle, it may help to add decodable texts matched to phonics lessons. Since LLI texts are not highly decodable, at-risk students may not gain sufficient experience in the application of the letter-sound knowledge they are taught, particularly if decodable texts are also not used in their classrooms. Similarly, since MS does not include leveled texts, at-risk students may not gain sufficient experience in reading texts with language patterns not constrained by the use of phonetically regular words.

Comparisons of intervention programs, such as those reported in this study, can help teachers identify potential deficiencies and in the selection of texts that may help lead to a resolution of reading problems. Whereas systems for the selection of qualitatively leveled texts are popular and standard in many commercially available reading programs (e.g., Rigby, Scholastic), supplementing with decodable texts requires different expertise on the part of teachers if they need to select and sequence texts that feature a strong LTTM.

Future Research Directions

We believe analyses such as those presented in this study are essential. Although previous descriptive studies of first-grade reading programs exist, this study is the first comparison of two intervention programs using a framework of word-, text-, and program-level features. Additional research on other commercially available intervention programs, at various grade levels, would add to this initial exploration. Research on features not included here (e.g., the nature of the phonics lessons in the program, vocabulary instruction), as well as other ways of measuring various features (such as repetition and the introduction of new words), would also be valuable. For example, our analyses used percentages to compare features—perhaps the proper diet of word characteristics has little to do with percentages but the frequency of words within a program.

Another valuable approach would be to determine the effects of intervention programs and text types on at-risk students with differing reading levels. Would readers who are extremely delayed versus only slightly delayed respond similarly to these programs or to the influences of leveled and decodable texts? As Cheatham and Allor (2012) intimated, the interaction of text type with a particular stage of students' reading development may be important. Using a variety of approaches to conceptualize and operationalize text and program characteristics, along with a variety of participant samples, would provide a knowledge base for researchers, publishers, and educators who need to know more about intervention programs they use or widely distribute.

Given the current preferential peaks of LLI and MS programs, intervention research specifically comparing their effectiveness is urgently needed regarding their immediate and long-term impact on the prevention and treatment of reading difficulties. To our knowledge, there is no peer-reviewed research on either program to support their effectiveness, let alone the influence of their texts.

Research isolating the effects of decodable and leveled text types is needed but, alas, is rare and is complicated by numerous threats to internal validity. Most difficulties encountered in this type of research pertain to not being able to employ randomized control group designs in school settings, not adequately controlling classroom factors when delivering supplemental intervention, and not having a sufficiently strong manipulation of the "text type" independent variable (see Jenkins, Peyton, Sanders, & Vadasy, 2004). To better isolate the effects of text types and LTTM, future research could randomize students at different stages of reading into groups who receive the same phonics instruction but read decodable or leveled texts. In addition, in nonrandomized designs, classroom instruction must be tightly controlled so that texts read in the classroom or elsewhere do not confound results.

Finally, as future research determines which program characteristics best support at-risk readers, researchers and publishers should partake in the reporting of LTTM and other features such as word repetition and percentages of concrete words or highly frequent words. Publishers are already expected to present readability and leveling information about texts; it would also be helpful for teachers to have information regarding these characteristics. Consistent with a point made by one of our reviewers, such reporting has the potential to raise awareness of important dimensions regarding words and texts included in intervention programs.

Notes

Address all correspondence to Maria Murray, Wilber Hall Room 113, Oswego, NY 13126; e-mail: maria.murray@oswego.edu.

1. Word-level, text-level, and program-level analyses were limited to descriptive interpretations. This is because we were not trying to make inferences or apply findings beyond the two programs we examined. It is the descriptions of similarities and differences that are most beneficial to educators who need to know more about the programs and texts they use.

2. To determine the percent increase or decrease in multisyllable words across programs, the mean of the first three levels was subtracted from the mean of the last three levels and then divided by the mean of the first three levels.

References

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

Afflerbach, P., Beers, J., Blachowicz, C., Boyd, D., Diffily, D., Gaunty-Porter, D., . . . Wixson, K. K. (2011). *Reading street.* Glenview, IL: Scott Foresman.

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The report of the Commission on Reading.* Washington, DC: National Institute of Education.

Aukerman, R. C. (1984). *Approaches to reading.* New York: Wiley.

Bear, D. R., Dole, J. A., Echevarria, J., Hasbrouck, J., Paris, S. G., Shanahan, T., & Tinajero, J. V. (2009). *Treasures.* New York: Macmillan/McGraw-Hill.

Beck, I. L. (1997, October/November). Response to "Overselling Phonics" (Letter to the Editor). *Reading Today*, 17.

Beck, I. L., & Block, K. K. (1979). An analysis of two beginning reading programs: Some facts and some opinions. *Theory and Practice of Early Reading*, **1**, 279–318.

Beck, I. L., Farr, R. C., & Strickland, D. S. (2008). *Storytown.* Orlando, FL: Harcourt School Publishers.

Beck, I. L., & McCaslin, E. S. (1978). *An analysis of the dimensions that affect the development of code-breaking ability in eight beginning reading programs* (Technical Report No. 6). Pittsburgh: Learning Research and Development Center.

Biemiller, A. (2009). *Words worth teaching.* Columbus, OH: SRA/McGraw-Hill.

Blachman, B. A., Schatschneider, C., Fletcher, J. M., Francis, D. J., Clonan, S. M., Shaywitz, B. A., & Shaywitz, S. E. (2004). Effects of intensive reading remediation for second and third graders and a 1-year follow-up. *Journal of Educational Psychology*, **96**, 444–461. doi:10.1037/0022-0663.96.3.444

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

Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, **2**, 5–142. doi:10.2307/746948

Bruner, J. S. (1966). *Toward a theory of instruction.* New York: Norton.

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

California State Board of Education. (2006). *Reading/language arts framework for California's public schools: Kindergarten through grade twelve.* Sacramento, CA: California Department of Education.

Chall, J. S. (1967/1983). *Learning to read: The great debate.* Fort Worth, TX: Harcourt Brace.

Cheatham, J. P., & Allor, J. H. (2012). The influence of decodability in early reading text on reading achievement: A review of the evidence. *Reading and Writing*, **24**, 1–24. doi:10.1007/s11145-011-93552

Cunningham, A. E., Nathan, R. G., & Raher, K. S. (2011). Orthographic processing in models of word recognition. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research* (Vol. **4**, pp. 259–285). New York: Routledge.

Cunningham, J. W., Spadorcia, S. A., Erickson, K. A., Koppenhaver, D. A., Sturm, J. M., & Yoder, D. E. (2005). Investigating the instructional supportiveness of leveled texts. *Reading Research Quarterly*, **40**, 410–427. doi:10.1598/RRQ.40.4.2

Decodable Books Written by Teachers. (2013). Retrieved from http://www.auburn.edu/academic/education/reading_genie/teacherbooks.html

DK Publishing. (2008). *My first words: Let's get talking.* New York: DK Preschool.

Dobeck, M. (2009). *Ant can't.* Portsmouth, NH: Heinemann.

Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme & R. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 87–111). Mahwah, NJ: Erlbaum.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, **9**, 167–188. doi:10.1207/x1532799xssr0902_4

Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading Rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, **44**, 414–448. doi:10.3102/0002831207302175

Ericcson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, **100**, 363–406. doi:10.1037%2F%2F0033–295X.100.3.363

Foorman, B. R., Francis, D. G., Davidson, K. C., Harm, M. W., & Griffin, J. (2004). Variability in text features in six grade 1 basal reading programs. *Scientific Studies of Reading*, **8**, 167–197. doi:10.1207/s1532799xssr0802_4

Foorman, B. R., Francis, D. G., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read. *Journal of Educational Psychology*, **90**, 37–55. doi:10.1037//0022-0663.90.1.37

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

Fountas, I. C., & Pinnell, G. S. (2008). *Leveled literacy intervention.* Portsmouth, NH: Heinemann.

Fountas, I. C., & Pinnell, G. S. (2012). *The Fountas and Pinnell text level gradient: Revision to recommended grade-level goals*. Portsmouth, NH: Heinemann.

Fries, C. C. (1963). *Linguistics and reading*. New York: Holt, Rinehart & Winston.

Fry, E. (1990). A readability formula for short passages. *Journal of Reading*, **33**, 594–97.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**(2), 256–281.

Goodman, K. (1969). Analysis of oral reading miscues: Applied psycholinguistics. In F. Gollasch (Ed.), *Language and literacy: The selected writings of Kenneth Goodman* (pp. 123–134). Boston: Routledge & Kegan Paul.

Graesser, A., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-metrix: Providing multilevel analyses of text characteristics. *Educational Researcher*, **40**, 223–234. doi:10.3102/0013189X11413260

Herman, G. (2001). *Lucky goes to school*. New York: Penguin Young Readers.

Hiebert, E. H. (2005). State reform policies and the task textbooks pose for first-grade readers. *Elementary School Journal*, **105**, 245–266.

Hiebert, E. H. (2011). Curious George and Rosetta Stone: The role of texts in supporting automaticity in beginning reading. In T. Rasinski, C. L. Z. Blachowicz, & K. Lems (Eds.), *Teaching reading fluency: Meeting the needs of all readers* (pp. 289–309). New York: Guilford.

Hiebert, E. H., Colt, J., 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. doi:10.3102/00028312029003545

Hoffman, J. V., Sailors, M., & Patterson, E. (2002). Decodable texts for beginning reading instruction: The Year 2000 basals. *Journal of Literacy Research*, **34**, 269–298. doi:10/1207/s15548430jlr3403_2

Jenkins, J. R., Peyton, J. A., Sanders, E. A., & Vadasy, P. F. (2004). Effects of reading decodable texts in supplemental first-grade tutoring. *Scientific Studies of Reading*, **8**, 53–85. doi:10.1207/s1532799xssr0801_4

Johnson, D., Moe, A., & Baumann, J. F. (1983). *The Ginn word book for teachers: A basic lexicon.* Boston: Ginn.

Johnston, F. R. (2000). Word learning in predictable text. *Journal of Educational Psychology*, **92**, 248–255. doi:10.1037/0022–0663.92.2.248

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grades. *Journal of Educational Psychology*, **88**, 3–17. doi:10.1007/978-1-4612-4282-6

Juel, C., Paratore, J. R., Simmons, D., & Vaughn, S. (2008). *My sidewalks on reading street*. Glenview, IL: Scott Foresman.

Juel, C., & Roper/Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly*, **20**, 134–152. doi:10.2307/747751

Loveless, T. (2012). *How well are American students learning?* Washington, DC: Brookings Institution.

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

Menon, S., & Hiebert, E. H. (2005). A comparison of first graders' reading with little books or literature-based basal anthologies. *Reading Research Quarterly*, **40**, 12–38. doi:10.1598/RRQ.40.1.2

Mesmer, H. A. (2005). Text decodability and the first-grade reader. *Reading and Writing Quarterly*, **21**, 61–86. doi:10.1080/10573560590523667

Mesmer, H. A. (2008). *Tools for matching readers to texts*. New York: Guilford.

Mesmer, H. A. (2010). Textual scaffolds for developing fluency in beginning readers: Accuracy and reading rate in qualitatively leveled and decodable text. *Reading Research and Instruction*, **49**, 1–20. doi:10.1080/19388070802613450

Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of text complexity for the early grades: Learning from the past, anticipating the future. *Reading Research Quarterly*, **47**, 235–258.

Meyer, L. A., Greer, E. A., & Crummey, L. (1987). An analysis of decoding, comprehension, and story text comprehensibility in four first-grade reading programs. *Journal of Reading Behavior*, **19**, 69–98. doi:10.1080/10862968709547588

Michel, J., Shen, Y., Aiden, A. P., Veres, A., Gray, M., The Google Books Team . . . Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, **331**, 176–182. doi:10.1126/science.1199644

National Center on Response to Intervention. (2013, May 9). Retrieved from http://www.rti4success.org/tools_charts/popups_instruction/my_sidewalks_intensive_reading_intervention_info.php

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects.* Washington, DC: Author. Retrieved from http://www.corestandards.org/the-standards

National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: Reports of the subgroups.* Washington, DC: National Institute of Child Health and Human Development, National Institutes of Health.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, **76**(3), 241–263. doi:10.1037/h0027272

Paivio, A., Yuille, J. C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values of 925 nouns. *Journal of Experimental Psychology*, **76**, 1–25. doi:10.1037/h0025327

Pearson, P. D. (2004). The reading wars. *Educational Policy*, **18**, 216–252.

Pearson leads the pack. (2010). Retrieved from http://www.publishersweekly.com/pw/by-topic/international/international-book-news/article/43564-global-publishing-rankings-2009.html

Peterson, B. (1988). *Characteristics of texts that support beginning readers* (Unpublished dissertation). Ohio State University, Columbus.

Pinnell, G. S., Huck, C. S., & DeFord, D. E. (1986). *The Reading Recovery project in Columbus, Ohio studies: Vol. 3. Year 1: 1985–1986*. Columbus: Ohio State University.

Plaut, D. C. (2005). Connectionist approaches to reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 24–38). Oxford: Blackwell.

Plaut, D. C., & Shallice, T. (1991). Effects of word abstractness in a connectionist model of deep dyslexia. *Proceedings of the 13th Annual Conference of the Cognitive Science Society* (pp. 73–78). Hillsdale, NJ: Erlbaum.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, **36**, 321–339. doi:10.1016/0022–0965(83)90036-X

Roberts, F. (2008). *Circus families*. Glenview, IL: Pearson Scott Foresman.

Scarry, R. (1985). *Richard Scarry's biggest word book ever!* New York: Random House.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, **55**, 151–218. doi:10.1016/00100277(94)00645-2

Stein, M., Johnson, B., & Gutlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Remedial and Special Education*, **20**, 275–287. doi: 10.1177/074193259902000503

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

Thorndike, E. L. (1921). *The teacher's word book*. New York: Columbia University Press.

Torgesen, J. K., Alexander, A. W., Wagner, R. K., Rashotte, C. A., Voeller, K., Conway, T., & Rose, E. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities*, **34**, 33–58. doi:10.1177/002221940103400104

Tunmer, W. E., & Nicholson, T. (2011). The development and teaching of word recognition skill. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research* (Vol. **4**, pp. 405–431). New York: Routledge.

Venezky, R. L. (1967). English orthography: Its graphical structure and its relation to sound. *Reading Research Quarterly*, **2**, 75–105.

Zeno, S. M., Ivens, S. H., Millard, R., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science.

Zinna, D. R., Liberman, I. Y., & Shankweiler, D. (1986). Children's sensitivity to factors influencing vowel reading. *Reading Research Quarterly*, **21**, 465–479. doi:10.2307/747617