Research Paper

A Comparison of Genre, Lexile, and Word Count of Read-Aloud Texts in Core Reading Series

Catherine Rand* [ID]

Rutgers University, United States

Abstract

This quantitative content analysis compares the Lexile level, estimated volume, and genre of Year Five and Six read-aloud texts used in four popular US core reading series to analyze the similarities between programs focusing on word count, Lexile level, and genre. The four series include Houghton Mifflin Harcourt's Into Reading, Lucy Calkins' Units of Study for Teaching Reading, Fountas & Pinnell's Classroom Collection, and McGraw-Hill's Wonders. Descriptive statistics showed considerable differences between programs, including low mean Lexile levels, considerable differences in volume of text, and too little informational text across all programs. These findings align with other scholars' critiques over the last 20 years. This data is important in understanding the effects of curriculum on comprehension and reading growth, as well as the equity of reading opportunities given to students.

Keywords: *Comprehension, Curriculum, Text Analysis, Read-Aloud, Elementary School*

INTRODUCTION

Core reading series are used or required in many Language Arts classrooms in the United States (edReports.org, 2022). Most of these programs are created and sold by private companies and offer a school year's worth of texts and instructional activities to teach reading. Most reading series are designed to be used as the primary source of curriculum for Language Arts classrooms, offering texts with accompanying lessons that fit into themed units with a pacing guide for the year. The texts from core reading programs can make up a large part of what students read in school (Duke, 2000). Yet, the effectiveness of the programs and how well they follow best practices in reading is questionable (Braker-Walters, 2014; Brenner & Hiebert, 2010; Dewitz, Jones, & Leahy, 2009; Foorman et al., 2004; McGill-Franzen, Zmach, Solic & Zeig, 2006; Al Otaiba et al., 2005; Torgesen; Pilonieta, 2010; Neuman & Wright, 2015). The features of the texts in these programs that affect comprehension, such as complexity, genre, or length (Amendum et al., 2018), are not readily apparent nor well-researched. Since they are sold by private companies, there is not much incentive for programs to be transparent about how their programs differ from other programs.

Student reading scores in the United States hover around two-thirds of students not passing the national literacy test (National Center for Educational Statistics, 2022). The Science of Reading has reignited the focus on the value of systematic phonics practice (DiMarco, 2022; Durán & Hikida, 2022; Sanden et al., 2022), yet the lack of comprehensive comprehension instruction and stagnation of students' reading progress in the US, beginning in upper elementary grades, is just as pressing an issue (Duke, 2000; National Center for Educational Statistics, 2022). Upper elementary grades, between Year Four and Year Six or grades three and five in the US, are important for preparing students for middle school and to get them on track to being prepared for college and careers (Valentine et al., 2002). Yet, there is much less reading research on the curricular texts and

activities used with that age group than there is for older or younger students. Knowing the quantity and volume of the texts that children learn with is crucial. Without that knowledge, understanding why our students are not making progress in reading is impossible. Studying the curricula in upper elementary level classrooms can offer insight into that pivotal time when so many students begin to fall behind in reading.

There is very little research that tracks the comprehension practice students receive in the classroom every day or the materials that teachers use to teach comprehension. In upper elementary grades, students are predominantly asked to read for meaning and comprehension rather than decoding for the first time. Comprehension becomes important across multiple subject areas, and students who learned to decode in early grades but did not master comprehension of multiple kinds of texts can fall behind. Duke and Block (2012), in a review of reading practices in elementary schools, found that comprehension, content knowledge, and vocabulary are frequently neglected instructionally in classrooms. EdReports has reviewed most of the current core reading series on the market and found that many of them do not offer comprehension instruction that is comprehensive or standards aligned (edReports.org, 2022). One of the primary activities that most reading programs use to teach and practice reading comprehension is the read-aloud or shared reading, where a teacher reads a text aloud to students while modeling a comprehension strategy and then having them practice that strategy with the same text.

Research is needed to analyze differences between comprehension instruction across different core reading series so that stakeholders can make informed decisions and we can further understand the issue of plateauing reading performance. One way to compare them is by analyzing the read-aloud texts they offer. This study aimed to measure and compare the genre, estimated volume of reading, and Lexile level of the whole-group, comprehension-focused, read-aloud texts across the most recent version of four popular core reading series in grades four and five, Year Five and Six in most international schools, when students are approximately nine to eleven years old. The reading programs were chosen based on popularity (Tuma et al., 2020). This study answers the question:

- What are the genres, volume of reading, and Lexile levels of fourth-and fifth- grade whole group shared-reading texts in four popular core reading programs, and how do they compare between programs?

LITERATURE REVIEW
Active View of Reading

The Active View of Reading asserts that reading requires several skills that all work together in order for one to successfully read (Duke & Cartwright, 2021). This view of reading builds on Gough and Tunmer's earlier Simple View of Reading, which argues that word recognition and language comprehension are both needed for reading to happen effectively (1986). Duke and Cartwright add two major components: bridging processes, including vocabulary understanding, fluency, and print concepts, and active self-regulation skills. This study uses these four foundational skill sets—word recognition, language comprehension, bridging processes, and self-regulation—as a guideline for what skills curricula should be offering practice with for their reading instruction. This study looks at read-aloud texts as models of those reading skills. Read-aloud texts are important for students to see and hear sophisticated texts, offer diverse structures and background knowledge, and are at a higher level than they can read independently (Alme, 2020; International Literacy Association, 2018; Greene et al., 2002; Layne, 2023). Analyzing the Lexile level of texts measures the relative challenge of word recognition and language comprehension (MetaMetrics, Inc., 2022). Measuring the volume of words used in a text or program suggests the amount of

stamina, working memory, and focus needed to complete the text and therefore measures the self-regulation practice offered (Calfee & Hiebert, 2011; Dixon, LeFavre, & Twilley, 1988; Hatcher, 2000; Mesmer, Cunningham, Hiebert, 2012; Nevo & Breznitz, 2011). Finally, analyzing the genre of the texts helps measure the diversity of texts students are exposed to and, therefore, the measure of print concepts and fluency practice students get by reading those texts. Students comprehend informational text in a different way than they comprehend literary texts and need cognitive flexibility to understand a text's structure (Duke, 2000; Hiebert, 2014; Marr & Gormley, 1982).

Research on Curricula

The research that does exist on reading curricula highlights some of the shortfalls of the programs. Many studies that examine curricula have focused on the instructional techniques used, and found that (Al Otaiba, et al., 2005; Dewitz et al., 2009; edReports.org, 2022; McGill-Franzen, et, 2006; Pilonieta, 2010). Others have focused just on vocabulary instruction, texts in younger grades, or reading assessment complexity (Brenner & Hiebert, 2010; Foorman et al., 2004; Toyama et al., 2017). One of the most recent reviews of how curricula programs affect reading comprehension argues that the Common Core recommendations led to instructional practices that do not align with best practices (Dewitz & Graves, 2021). This finding aligns with Duke and Block's work from 2012, which found that short-term decoding goals were emphasized more than long-term reading goals, namely comprehension.

Brenner and Hiebert (2010) and Foorman et al. (2004) focused on the text complexity of core reading program texts, which were looked at in younger grades. or the complexity of popular reading assessments (Toyama et al., 2017). Some studies have focused on the vocabulary instruction provided in core reading series or on the availability and use of informational texts within the series (Graves et al., 2019; Wright & Neuman, 2013). This study reflects just how necessary it is to delve further and examine what is being used in the classroom to understand why our students are not making more progress.

In 2012, Duke and Block compared the research done in classrooms to the recommendations given in the 1998 National Research Council report, Preventing Reading Difficulties in Young Children. They found that short-term decoding goals were emphasized more than long-term reading goals, like comprehension, that students need to succeed later in school. EdReports, a company that reviews curriculum programs for how well they meet the Common Core State Standards, found that of the 105 elementary-level programs reviewed, only 34.3% of them, or 36 programs, were aligned enough to the Common Core to meet their expectations. Over the last five years, the study reviewed curricula across all subjects and measured whether the program contained instruction related to all the standards, as well as how much time and attention each standard received. EdReports did not, however, evaluate the specific texts used in any of the programs and what they are used for instructionally (EdReports, 2022).

Vocabulary

One of the most important features of a text that affects its complexity is vocabulary. The accepted threshold for the percentage of words within a text that need to be familiar for a reader to comprehend a text fully is 98% (Schmitt, Jiang, and Grabe, 2011). According to Schmitt et al.'s data, the participants who were able to reach at least 70% comprehension needed to score between 98% and 99% on their vocabulary pre-test. Most vocabulary growth is not through direct instruction but rather through exposure to new vocabulary in texts (Anderson & Nagy, 1993). If a student were able to successfully learn ten words a week for all the years they were in school, it would only amount to 4,320 words. Yet, Anderson and Nagy (1993) estimate that students are exposed to

roughly 20,000 to 40,000 different words through texts each year.

A study by Foorman et al. (2004) found very little repetition of vocabulary words within a curricular unit and did not provide enough practice with vocabulary (Foorman et al., 2004). Neuman and Wright measured the target vocabulary students were expected to learn in four weeks' worth of text from four popular programs at the kindergarten level (Neuman & Wright, 2015). The programs varied from about two words to more than 20 words learned each week. Over a full year, this difference amounts to 81 compared to 720 words (Neuman & Wright, 2015).

Word Count and Syntax

The length can affect the complexity of a text in several ways. Sentence length and the number of clauses in each sentence are predictors for text complexity because of their impact on working memory and cognitive processing. Hatcher (2000) and Hiebert (2014) found that 83-88% of the variance in book level could be explained by sentence complexity. The overall length of a text can also increase the need for working memory and cognitive load. This is necessary for understanding the discourse of a text and can even predict reading comprehension and reading speed (Dixon et al., 1988; Nevo & Breznitz). Stamina is required to be able to continually use reading strategies or actively comprehend all of a text (Calfee & Hiebert, 2011). Calfee and Hiebert found that texts between 350 and 450 words were easily read, but passages as long as 800-1000 words proved more challenging (Calfee & Hiebert, 2011). Different genres also vary in their sentence lengths and variety of sentences (Hiebert, 2014), so practice with multiple genres is important to practice flexible thinking and stamina. Without these skills, overall comprehension of the text is difficult. Multiple studies have found that the disparity between the balance of text types disproportionately affects students in low-SES schools (Braker-Walters, 2014; Jeong et al., 2010).

Brenner and Hiebert (2010) reviewed a sample of text from six 3rd-grade core reading programs for the number of words per text and the volume of reading given to students. The research they cite recommends that students need between 45 and 60 minutes of reading practice each day. The programs only offered between 21.9 to 13.7 minutes a day of reading (Brenner & Hiebert, 2010). Hatcher (2000) and Hiebert (2014) found that 83-88% of the variance in book level could be explained by sentence complexity. One of the few recent studies to measure aspects of text in core reading programs was an analysis of vocabulary rarity and repetition (Graves et al., 2019). They reviewed the whole group texts from the student anthology from the four best-selling curriculum programs of 2014 for grades one to six. They found that some programs offered half the total number of words and pages read as other programs (Graves et al., 2019).

Genre & Volume of Reading

Reading volume is also an important factor to consider in text complexity. While there is little evidence as to what specific linguistic factors are needed for students to improve the most, there is evidence that students who read more text have been shown to make more gains in reading than students who read less (Foorman et al., 2006). Student access to reading longer texts and many texts is also an important factor because of the Matthew Effect, in which the gap between the successful students and those who aren't continues to grow over time (Mol & Bus, 2011; Stanovich, 1986). Anderson, Wilson, & Fielding (1988) and Locher and Pfost (2020) have shown that students' exposure to text, or time spent reading, correlates with the differences between more and less proficient readers.

Another critical aspect of text complexity is text genre. There are many ways to define genre and it is not consistent across the literature. This study will discuss genre based on structure rather than content, using the terms literary and informational. Exposure to equal amounts of literary and

informational text is important for students because many informational types of texts are more likely to be what students will see and interact with outside of school, and they require different skills than reading literary texts (Gee, 2018; National Reading Panel, 2000; The New London Group, 2000). The term literary in this study is used to include narrative texts, dramatic texts, and poetry. Informational text, on the other hand, is defined as conveying information accurately to increase a reader's knowledge. This includes content-rich nonfiction in the sciences and history, how-to and technical writing, persuasive texts, etc.

Hiebert (2014) found that literary text had a wider variation in sentence length, with a mean variation of 9.5 words between sentences, than the informational text, with a mean variation of 6.9 words between sentences. She argues that texts of different genres have different factors of complexity and therefore require different skills to comprehend them (Hiebert, 2014). Background knowledge is also critical for readers to construct knowledge from a text. Without a schema about a topic, even if a reader knows individual words, they cannot make meaning from them (Rumelhart & Schemata, 1980). Marr and Gormley (1982) showed that reading level did not account for comprehension differences between students, but prior knowledge about the subject accounted for 23% of the variance in scores. Multiple studies have found disparity between the balance of text types in the classroom, disproportionately affecting students in low-SES schools (Braker-Walters, 2014; Jeong et al., 2010).

Duke (2000) found that in first-grade classrooms of low-SES schools, literary texts comprised 68.34% of the texts that students were exposed to, while only 6.28% were informational. In the high-SES schools, informational text went up to 11.00%. Seven of the twenty classrooms studied did not use any informational texts over the course of the study (Duke, 2000). More recent studies have found similar disparities (Braker-Walters, 2014; Jeong et al., 2010). This difference affects student achievement. McGill-Franzen, Zmach, Solic, and Zeig (2006) found that student state test scores correlated with the third-grade core reading programs they experienced. One of the major differences between programs is the large discrepancy in the number of longer texts used. Both were lacking in the number of informational texts in comparison to literary texts. Braker-Walters (2014) reviewed three different core reading programs at the 4th grade level and found that none of them had more than a third informational text, accounting for 14%-18% of overall reading, despite longstanding recommendations of 50% informational text (Common Core State Standards Initiative, 2022). Jeong, Gaffney, and Choi (2010) had similar findings, looking at the proportion of informational texts in 2nd, 3rd, and 4th-grade classroom libraries. They found that across 15 classrooms in three school districts, the average percentage of informational text in the class libraries was 14.6% in 2nd grade %, 13.1% in 3rd grade, and 13.8% in 4th grade (Jeong et al., 2010).

Other Curricular Concerns

Pilonieta (2010) looked at five core reading programs at the 1st, 3rd, and 5th grade levels to determine if the programs offered instruction and application of fourteen research-recommended comprehension strategies. She found, like Al Otaiba et al. (2005), that none of the programs taught and practiced all of the strategies, averaging about two-thirds of the strategies while neglecting others. Programs also did not use the gradual release of responsibility, and some had little to no form of guided or independent reading practice (DeWitz, Jones, and Leahy, 2009; McGill et al., 2009; Pilonieta, 2010). They also found that the programs offered less instruction and intensity of strategy practice than the research suggests is appropriate (Dewitz, Jones, & Leahy, 2009).

These studies taken together suggest that core reading programs may offer the same volume of reading as one another, and do not necessarily align with best practices or the standards. However, there is little clear comparison between curriculum programs. Additionally, many of

these studies are no longer relevant, as publishers have introduced newer, updated versions of their programs. There is also too little research on the complexity of the texts used in these programs, which matters for student learning (Pearson, 2013).

Knowing about the texts that children learn with is crucial. Without that knowledge, understanding why our students are not making progress in reading is impossible. This study examines these curricular read-aloud texts in the most popular core reading programs in grades four and five for word count, genre, and Lexile level.

RESEARCH METHOD
Data Collection
The data for this study were collected from the fourth- and fifth-grade, or Year Five and Year Six, versions of the following core reading programs using the most recent version at the time of study:

- Houghton Mifflin Harcourt's Into Reading (Ada et al., 2020)
- Lucy Calkins' Units of Study for Teaching Reading (Calkins, 2015)
- Fountas & Pinnell's Classroom Collection (Fountas & Pinnell, 2020)
- McGraw-Hill's Wonders (August et al., 2020)

These programs were chosen because they were listed in the top ten of the most popularly used in a national survey of teachers and administrators in both 2021 and 2022 (Tuma et al., 2020). Online review copies of the programs, available to researchers, academics, and school professionals, were used to gain access to a list of all texts, a scope and sequence that showed how each text was used, and many copies of the texts. Any texts not available through the reading series website were obtained using public libraries or available online copies. Titles, authors, genre, Lexile level, and word count were collected from all whole-group, reading comprehension instructional texts and organized in a spreadsheet. Texts used for small group instruction, guided reading, decoding practice, or individual reading were excluded. The total texts assessed included 11 texts in the Lucy Calkins fourth-grade program, 8 texts in Lucy Calkin's fifth-grade program, 12 texts in McGraw-Hill's fourth-grade program, 12 texts in McGraw-Hill's fifth-grade program, 43 texts in Houghton Mifflin Harcourt's fourth-grade program, 42 in Houghton Mifflin Harcourt's fifth-grade program, 120 texts in Fountas and Pinnell's fourth-grade program, and 120 texts in Fountas and Pinnell's fifth-grade program.

Analysis
This study used a quantitative content analysis to examine the texts from five fourth- and fifth-grade reading programs (Krippendorf, 2018). Content analysis is a method for evaluating texts in which different variables, or features of a text, are defined, and then each occurrence of that variable is coded so that variables can be analyzed quantitatively (Neuendorf, 2017). The variables in this study included the text genre, the total number of words, and the Lexile level of a 500-word sample from the middle of the text. Each occurrence of all three variables was coded using the following process.

In order to measure genre in a way that can be compared across texts and across core reading programs, it was coded as informational or literary. The structure of the text was considered, rather than the content (NGACBP & CCSSO, 2023). Texts organized by sequence and events, including narratives, poetry, graphic novels, and drama, were labeled as literary. Texts or videos organized by topic rather than sequential events were labeled as informational. For the sake of brevity, the term "text" will generally include any media, as well as written text. When texts fell into both

categories, such as narratives meant to teach about history or a biography, text features were used to determine which was the dominant genre. Features such as subheadings, text boxes, and informational graphics pushed texts into the informational category. Dialogue, sensory description, and narrative elements like alliteration and rhythmic phrasing or strong use of metaphor pushed texts into the narrative category.

Table 1. Informational vs. Literary Categorization of Text Topics

Literary	Informational
Fiction novels	Biography/Autobiography
Memoir	Science
Short fiction	History
Poetry	Reference (Encyclopedias, Dictionaries, etc.)
Humor	Infographics
Graphic Novels	Expository
	Persuasive
	Technical writing
	Narrative Nonfiction

Text volume was measured with the estimated word count of a text and the number of texts used. This was calculated by collecting a page count and an estimated average number of words per page. Page count was measured by counting tokens, using the text version used by the core reading series. Titles and footnotes were not included in the word count, as they may affect stamina differently. For texts five pages or shorter, word count was measured for each page and then averaged based on the number of pages. For texts six pages or longer, a set of five pages, taken from the middle of the text to adjust for possible anomalies in word count at the beginning or end of a book, was counted, averaged, and multiplied by the number of pages in the text to calculate an estimated word count for the text. All texts used over the course of the whole school year were measured.

The Lexile level of each text was recorded or analyzed using the Lexile® Analyzer tool (MetaMetrics, Inc., 2022). The Lexile Framework for Reading was used because of its popularity with publishers, such as Scholastic and Houghton Mifflin Harcourt, and for its reliance on transparent complexity measures (Cunningham, Hiebert, & Mesmer, 2018; Gunning, 2003; Scholastic, Inc., 2021; Smith, Burdick, Sanford, Burdick, 2006). The Lexile Text Analyzer, an online tool designed to give an estimated Lexile level of any text, uses mean log word frequency – or the rareness of the vocabulary in the passage – and the log of the mean sentence length as the factors to determine complexity (MetaMetrics, Inc., 2022). The Lexile Analyzer was used to assess the Lexile level – specifically the mean log sentence length and mean log word frequency – of any texts that don't already have a publicly-released score.

If the text was more than 10 pages or 2000 words, a 1000-word passage was taken from the middle of the text and run through the Analyzer. The author used MetaMetrics' Lexile Text Analyzer and is not a trained MetaMetrics employee, so the Lexile levels are estimates only. The Analyzer produces an estimated Lexile level that was calculated based on the mean sentence length of the text and the mean log word frequency. Texts that are poems, dramas, or other similar forms are not designed to be measured by Lexile and, therefore, do not have a Lexile level collected.

Since the reader and task aspects of text complexity cannot be measured in large quantities, this study controlled for them instead. Only whole-group, teacher-led read-aloud texts were used

as data. In this way, the study controlled for reading tasks, eliminating such tasks as guided reading, fluency practice, or independent reading. The reader was controlled for by assuming that the students using these programs are in the 50th percentile in terms of reading level, measured by Lexile (MetaMetrics, 2023). To ensure validity and reliability, two other researchers recounted 10% of the texts and determined the genre of the texts using the established criteria with a 98% agreement rate.

After coding all variables, descriptive statistics were run, including the mean, median, and range for the estimated total number of words per program and Lexile level for each program. The percentage of literary and informational texts in each grade's program from each publisher was calculated as well. The data from each program were graphed to show the range of Lexile levels of the texts within and between programs, the difference in Lexile levels between fourth and fifth grade for each program, and the difference in Lexile levels by word count.

MetaMetrics conducted a national study with more than 3 million students that measured the level of texts they were able to read compared to other students. From that study, they released national student norms of the Lexile level that a student in the 50th percentile and a student in the 90th percentile are typically able to read. A Lexile measure for the beginning of the year, middle of the year, and end of the year is given for each grade for both percentile ranks. The range of Lexile levels in each reading program in this study was compared to the range of student norms for fourth and fifth grade, using the beginning of the year 50th percentile as the bottom of the range and the end of the year 90th percentile as the top of the range to approximate what texts could be read independently by most students.

FINDINGS AND DISCUSSION

Analysis of the fourth- and fifth-grade core reading series' Lucy Calkins' Units of Study for Teaching Reading (LC), McGraw-Hill's Wonders (MH), Houghton Mifflin Harcourt's Into Reading (HMH), and Fountas & Pinnell's Classroom Collection (F&P) resulted in a total N of 368 texts, 186 fourth-grade texts, and 182 fifth-grade texts. Within the fourth-grade programs, LC included 11 texts, MH included 12 texts, HMH included 43 texts, and F&P included 120 texts. Within the fifth-grade programs, LC included 8 texts, MH included 12 texts, HMH included 42 texts, and F&P included 120 texts. Of the 368 texts, 1 text in the MH fifth-grade curriculum, 7 texts in the F&P fourth-grade curriculum, and 6 texts in the F&P fifth-grade curriculum were unable to be obtained for a word count but were included in the Lexile and genre analysis.

Lexile levels, estimated word counts, and the genre were recorded for each text in each program and compared.

Table 2. Overview of Number of Texts, Estimated Total Word Count, and Genre by Program

	# Texts (N)	Mean Lexile	Estimated Total # of Words	Genre	
				# Texts Narrative (%)	# Texts Informational (%)
LC 4th	11	866L	214,662	8 (73%)	3 (27%)
MH 4th	12	758L	273,014	9 (75%)	3 (25%)
HMH 4th	43	821L	62,188	26 (60%)	17 (40%)
F&P 4th	120	720L	325,671	115 (96%)	5 (4%)
LC 5th	8	793L	128,984	6 (75%)	2 (25%)
MH 5th	12	803L	279,864	10 (83%)	2 (17%)
HMH 5th	42	914L	107,901	24 (57%)	18 (43%)

|	# Texts (N)	Mean Lexile	Estimated Total # of Words	Genre	
				# Texts Narrative (%)	# Texts Informational (%)
F&P 5th	120	789L	476,623	106 (88%)	14 (12%)

Lexile Levels

The programs varied in their mean Lexile levels, as well as their ranges. While LC had the highest mean and median of the fourth-grade programs, 866L and 880L, respectively, FP's fourth-grade program had the lowest mean and median Lexile levels, with 720L and 690L, respectively. LC's fourth-grade program had a mean of 866L and a median of 880L. MH's fourth-grade program had a mean of 758L and a median of 820L. HMH had the highest mean and median Lexile level of the fifth-grade programs, while FP had the lowest mean and median, 789L and 770L, respectively. LC's fifth-grade program had a mean of 793L and a median of 790L. MH's fifth-grade program had a mean of 803L and a median of 800L.

The program with the widest range of Lexile levels was FP's fourth-grade program, a range of 800. The narrowest range belonged to MH's fifth-grade program, a range of 320. MH's fourth-grade program had the greatest discrepancy between its mean and median Lexile, a value of 62. Two of the four programs, HMH and FP, had mean and median Lexile levels that increased from their fourth-grade program to their fifth-grade program. The LC programs decreased in their mean and median Lexile levels from fourth to fifth grade. The MH programs increased by an average of 45 points from fourth to fifth grade, but the median Lexile level decreased by 20 points.

Table 3. Mean and Median Lexile Levels of Each Core Reading Program

Lexile	4th Mean	4th Median	4th Range	5th Mean	5th Median	5th Range
LC	866L	880L	590-1220L	793L	790L	550-1120L
MH	758L	820L	520-930L	803L	800L	650-970L
HMH	821L	810L	520-1140L	914L	940L	590-1090L
F&P	720L	690L	310-1110L	789L	770L	460-1140L

The change in Lexile between grades in each program and a comparison to the national student norms can be seen in Figure 1.
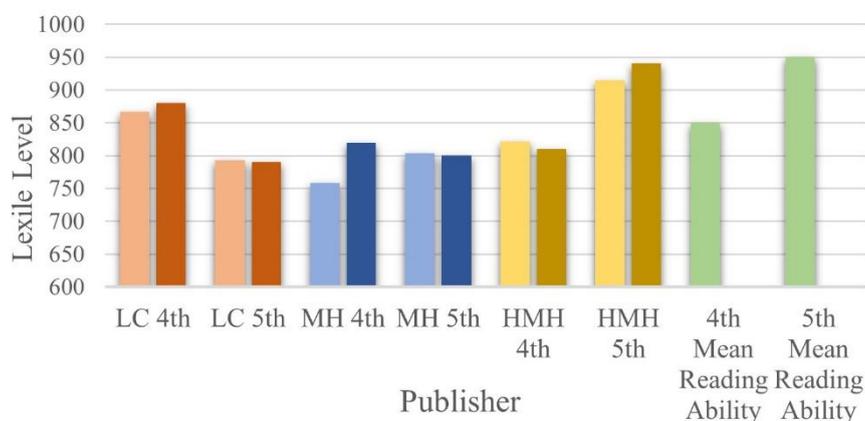
Figure 1. Mean and Median Lexile Level of Texts by Publisher

The Lexile ranges of the programs studied spanned about four grade levels, giving students a range of texts to read. However, the range in six of the eight programs in this study has mean Lexile levels well below the Lexile level that most end-of-year fourth graders can read (MetaMetrics, 2023). Houghton Mifflin Harcourt's fifth-grade Lexile mean approached the fifth-grade national student end-of-year 50% norm of 950L, and Lucy Calkins fourth-grade Lexile mean was the only program to exceed the fourth-grade national student end-of-year 50% norm of 850L (MetaMetrics, 2023).

For students' reading abilities to grow, they need to be exposed to texts that challenge their current skills and are within their Zone of Proximal Development. These core reading programs predominantly offered read-aloud texts that are below where 50% of the nation's fifth graders are reading independently (MetaMetrics, 2023). Even the Fountas and Pinnell and Houghton Mifflin Harcourt series, which have many more texts than the other two programs, have fewer than half of the texts at or above where most students can read independently. In Lucy Calkins' fifth-grade series, only 6% of the words were from texts on or above 950L, the national norm at which 50% of fifth graders can read independently (MetaMetrics, 2023).

Read-aloud texts are designed to be read out loud by the teacher to the whole group of students to model reading comprehension strategies and skills (Fisher et al., 2004). Because the teacher is the person who is reading, students can focus on comprehension because they do not have to focus on decoding difficult words or fluency of reading, and therefore, the text can be more difficult than what students read independently. Read-alouds are important because they introduce more sophisticated literary techniques, text structures, genres, and high-level vocabulary than students would otherwise see in their own reading (Fisher et al., 2004).

Low Lexile levels suggest a lack of high-level vocabulary and syntax (MetaMetrics, 2022). These low levels call into question the effectiveness of these programs in helping students improve their reading ability. If these are the only models students are receiving, they are not being introduced to many more sophisticated texts than what they are already able to read independently, which has the potential to limit student reading growth. Other scholars have raised the alarm about how much vocabulary students are learning in school (Anderson & Nagy, 1993; Foorman et al., 2004; Neuman & Wright, 2015). Vocabulary is one of the dividers between high- and low-SES students, and a lack of vocabulary in a core reading series only worsens that divide, as low-SES students are less likely to increase their vocabulary from experiences outside of school (MaGuire et al., 2018).

Volume of Reading

The volume of reading was measured both by the number of texts a program included and by the estimated number of words in each text. Videos were not included in the word count. Once the estimated total number of words for each text was calculated, the estimated number of words in each program was calculated. The programs varied greatly in their estimated number of words. In the fourth-grade programs, LC had 214,662 words across 11 texts, MH had 273,014 words across 12 texts, HMH had 62,188 words across 43 texts, and FP had 325,671 words across 120 texts. The mean number of words per text in each fourth-grade program was 19,514 (LC), 22,751 (MH), 1,446 (HMH), and 2,714 (F&P).

In the fifth-grade programs, LC had 128,984 words across 8 texts, MH had 279,864 words across 13 texts, HMH had 107,901 words across 42 texts, and FP had 476,623 words across 120 texts. The mean number of words per text in the fifth-grade programs was 16,123 (LC), 21,528 (MH), 2,569 (HMH), 3,971 (F&P). HMH's read-aloud texts were mostly excerpts, while LC, MH, and FP used full texts.

I also calculated the amount of instructional time each text accounts for. The amount of time each text would take to read aloud by the teacher was calculated using a 180-day school year, assuming an oral reading rate of 183 words per minute – a rate determined by Brysbaert's meta-analysis of the average number of words adults read aloud per minute (Brysbaert, 2019). It also assumes that the teacher is reading at a pace typical of most adults. Given these parameters, within the fourth-grade programs, LC provided 6.5 minutes of reading a day from the read-aloud texts, MH provided 8.3 minutes, HMH provided 1.9 minutes, and FP provided 9.9 minutes. Within the fifth-grade programs, LC provided 3.9 minutes, MH provided 8.5 minutes, HMH provided 3.3 minutes, and FP provided 14.5 minutes.

Table 4. Number of Texts, Estimated Word Count, and Reading Time by Program

	# Texts (N)	Estimated Total # of Words	Estimated Mean # of Words per Text	# Minutes per Day of Reading
LC 4th	8	214,662	19,514	6.5
MH 4th	9	273,014	22,751	8.3
HMH 4th	26	62,188	1,446	1.9
F&P 4th	115	325,671	2,832	9.9
LC 5th	6	128,984	16,123	3.9
MH 5th	10	279,864	21,528	8.5
HMH 5th	24	107,901	2,569	3.3
F&P 5th	106	476,623	3,971	14.5

The distribution of word counts for each text by Lexile level also differed across programs. The majority of texts in the HMH fourth- and fifth-grade programs and FP fourth- and fifth-grade programs, as well as the LC fifth-grade program, were under 10,000 words. There was a wider range of word count in the other three programs, the largest range being a difference of 414,435 words. In the LC's fourth-grade program, there was a fairly even distribution of high Lexile levels and high word count. In the MH fourth-grade program, the higher Lexile texts were the shortest, and the lowest Lexile levels accounted for the most words. In contrast, in the publisher's fifth-grade program, the text with the highest Lexile level was also the longest. The fourth- and fifth-grade HMH and FP programs had few longer texts at all.

The wide variation in the volume of reading across these programs is a concern. Studies have shown that students who read more perform better in reading (Cummingham & Stanovich, 2003). Students who struggle to progress in their reading typically read fewer texts independently or at home than strong readers do (Allington & McGill-Franzen, 2021). Furthermore, students from low-SES backgrounds, more likely to be students of color, are generally exposed less to literacy activities outside of school (Allington & McGill-Franzen, 2021). In a recent study, 25% of caregivers reported never reading with their children (Logan et al., 2019). The texts that students are exposed to in school become even more important to ensure they are given an equitable, high-quality reading education that will allow them the opportunity to succeed as adults. Higher levels of education can equate to fewer health issues, higher economic gains, and an improved sense of happiness (Assari, 2019; Berkman et al., 2004; Rocha & Ponczek, 2011; Wang, Liu, & Cai, 2022). Higher literacy also correlates with lower dropout rates (Hernandez, 2011).

The volume of text is also important for building skills like working memory and stamina, and for offering practice with strategies like text structure. Calfee and Hiebert (2011) suggested that texts need to be at least 500 words for a reader to need stamina. The need for working memory

and the presentation of text structure, especially narrative structure, is much stronger in longer texts, such as those in Lucy Calkins' programs, than in excerpts and short texts, like those in Houghton Mifflin Harcourt's or Fountas and Pinnell's programs. When reading programs offer predominantly short texts, it means students do not get much practice with narrative development. On the other hand, when programs only offer a handful of longer texts, students get less practice with a variety of genres, background knowledge, vocabulary, and breadth of strategy practice.

The difference in word count between the fourth-grade programs is more than 500%. Houghton Mifflin Harcourt's read-aloud texts for fourth grade offer just 19% of the words that Fountas and Pinnell offer in fourth grade, a difference of 263,483 words. Assuming 15 minutes for a read-aloud each day – the most amount of time any program offered in terms of text volume – the difference between those two programs equates to 96 extra days of read-aloud instruction. In the fifth-grade programs, the biggest difference is between the same two publishers. Because these fifth-grade programs have more volume of text than in fourth grade, this discrepancy amounts to 33.6 hours, or nearly 134 days of read-aloud instruction.

If these programs have similar differences in the lower grades and are used from kindergarten through sixth grade, that could amount to a discrepancy of between three and a half and five years' worth of read-aloud time students are not receiving, a difference of between 1.8 and 2.6 million words. Young children who are read to more have shown to have more brain activity in the areas that deal with mental imagery and narrative comprehension (Hutton et al., 2015). Elementary students were also found to have improved vocabulary learning from more texts read aloud to them (Elley, 1989). Yet, students in low-SES households are far less likely to have family read to them outside of school (Allington & McGill-Franzen, 2021). For students who already get less exposure to literacy activities and texts outside of school, receiving much less exposure to complex texts and much less modeling of comprehension only compounds the inequality they already face in their educational experiences. The differences in the volume of read-aloud text students are exposed to in these programs raise questions about how student reading growth in later grades is influenced by the reading series students experienced in elementary grades.

Duke found that students were given an average of 3.6 minutes' worth of informational reading each day, but only 1.9 minutes in low-SES school districts (Duke, 2000). Racial and socio-economic segregation of schools in the U.S. has ensured that schools or districts in low-SES communities have fewer resources than schools in higher-SES communities, including money for curricula (Mathewson, 2020; U.S. Government Accountability Office, 2022).
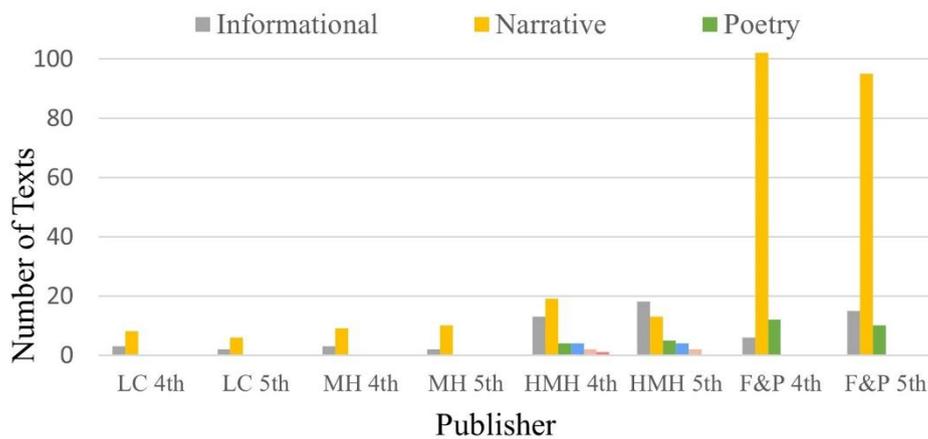
Genre

The LC program had 27% informational texts, MH had 25% informational texts, HMH had 40% informational texts, and FP had 3% informational texts. In the fifth-grade programs, LC had 25% informational texts, MH had 17% informational texts, HMH had 43% informational texts, and FP had 12% informational texts.

When the volume of reading is considered, the percentage of informational texts mostly decreases. The number of words was categorized by the genre of each text and totaled for each program. In the fourth-grade programs, the percentage of words in the program that belonged to informational texts was 17% in LC, 4% in MH, 41% in HMH, and 4% in FP. In the fifth-grade programs, the percentage of words in informational texts was 12% in LC, 6% in MH, 37% in HMH, and 13% in FP.

See Table 2 for the breakdown of narrative and informational texts in each program. Two programs, Houghton Mifflin Harcourt and Fountas and Pinnell, included poetry in their texts, which were included under literary texts. Only Houghton Mifflin Harcourt included other forms of text,

including dramas, videos, and graphic novels.

Figure 2. Number of Texts by Genre by Program

The discrepancy between literary and informational texts in these reading series follows the findings of other studies that show there continues to be a deficit of informational text in classrooms, despite decades-old recommendations for equal literary and informational text (Braker-Walters, 2014; Duke, 2000; Jeong et al., 2010). Houghton Mifflin Harcourt's programs had the most balance between types of text, with a percentage higher than other scholars found in older series, but it is still not the fifty percent informational text that is recommended by the Common Core Standards for decades and judged by edReports to be a high-quality program (edReports.org, 2022; NGACBP & CCSSO, 2023).

An important discovery in this study was the lack of a clear definition when it comes to genre. For example, McGraw-Hill states on their website that "Wonders also delivers an equal blend of literary and informational text with selections like 'Wild Weather' by Seymour Simon and a variety of articles that cover a wide range of topics…" (McGraw-Hill, 2023). Their read-aloud texts have just 25% informational texts in fourth grade and 17% in fifth grade. If considered by the number of words rather than the number of texts, the amount of informational text drops to just 4% in fourth grade and 6% in fifth grade. To have an "equal blend" of both genres would mean that the vast majority of informational texts in the series are in small group or independent reading activities, not during the read-aloud, where the texts and corresponding comprehension strategies are being modeled for students, an unlikely scenario.

The National Reading Panel (2000) and Duke and Block (2012) found that reading strategies focused on comprehension are undertaught in schools. Braker-Walters (2014) and Jeong et al. (2010) also found that informational texts in many classrooms, libraries, and reading programs are lacking, with similar percentages to this study. The terms informational and nonfiction sometimes make it difficult to tell what types of text publishers are using (Gill, 2009; Moss, 2008). Given McGraw-Hill's statement about informational text in their core reading series that mentions the topic of texts, it seems clear they are using a definition based on text content, not structure. It is important to define texts by structure because defining a genre by content does not consider the comprehension skills and strategies needed to understand the text. For example, analyzing different text structures, finding evidence or quotes, or evaluating the reasoning of a speaker's argument are not strategies that can be used with most literary texts, even if they are about true concepts. Instead, students require an informational structure to learn those skills. Complexity research that shows that syntactic and lexical complexity, as well as qualitative factors like text structure and author's voice, affect students' ability to understand a text (Frantz, et al., 2015;

Fountas & Pinnell, 2017; Frey, et al., 2016; Fry, 2002; Graesser, et al., 2014; Hiebert, et al., 2019).
Defining a text by content and whether that content is true or not also disregards the issue of author bias and unreliable sources, and does not leave space for interpretation of what is true.

The disparity between genres of text is concerning, as well, because students are expected to be able to comprehend informational text not only in their reading classes, but in science, social studies, and other content subjects. The demands of reading informational text in content subjects continue to increase in later grades, yet national student reading performance continues to decrease as students get older (National Center for Educational Statistics, 2022). Some students also prefer informational texts and may not be building the positive associations with reading that they could because of an over-prevalence of literary texts. The percentage of informational text in these programs is more than found in first-grade classrooms twenty years ago (Duke, 2000), but there is still much progress to be made.

Similarly, media, such as images, videos, and games, have been neglected in these programs, despite having been identified as important and authentic sources of text and literacy in the modern world that students need to be able to interact with (Gee, 2018; Leu et al., 2013). Media is included in the Common Core State Standards and supported by scholars as a necessary part of contemporary reading instruction (NGACBP & CCSSO, 2023). New Literacy scholars argue that students need to learn how to read and comprehend in the 21st century, using the technology and forms of text they are more likely to encounter outside of school (Gee, 2018; Leu et al., 2013). Only one publisher, Houghton Mifflin Harcourt, included any media as read-aloud material, with four informational videos, two dramas, and one graphic novel in their fourth-grade program, and four informational videos and two dramas in their fifth-grade program. 33% of surveyed adults in the U.S. read at least one e-book, rather than physical text (Faverio & Perrin, 2022), 74% of polled Americans played some form of video game (The NPD Group, Inc., 2020), 73% of adults report using YouTube to watch videos, and more than 60% use social media platforms (Perrin & Anderson, 2019). Students need practice with those forms of communication in order to use them effectively and critically, both in a personal and professional capacity. Based on the texts from the four programs in this study, students are not seeing texts that reflect the literacy types that they will encounter as adults. The texts in these programs are not preparing students to be literate in the current world, where 92.1% of businesses with over 100 employees use the internet for communication, marketing, hiring, or to find information (Statista, 2023). Preparing students for how to be literate with new technology and media is especially important for equity, as students from the lowest SES households and Black and Hispanic students are the most likely to not have access to technology at home and rely on schools for computers and the internet (National Center for Educational Sciences, 2018).

CONCLUSIONS

This study analyzed the Lexile level, volume of reading, and discrepancy between literary and informational read-aloud texts in four popular fourth- and fifth-grade commercial core reading series. These included Lucy Calkins' Units of Study for Teaching Reading, Houghton Mifflin Harcourt's Into Reading, Fountas & Pinnell's Classroom Collection, and McGraw-Hill's Wonders. There were several differences between programs that led to inequities for the students who were given instruction through each program.

The mean Lexile level for texts in these eight programs ranged from nearly 200L, and the median ranged from 250L, a difference of almost two full grade levels (MetaMetrics, 2023). The mean Lexile level for all programs fell below the mean reading level of students in each corresponding grade, suggesting that the read-aloud texts that are supposed to offer challenging

models of text to support reading growth are not offering the challenge they need to.

The estimated volume of text for each program ranged from 62,188 words (HMH 4th) to 476,623 words (F&P 5th), with Fountas and Pinnell's 5th grade program offering more than seven times the volume of read-aloud practice as Houghton Mifflin Harcourt's fourth grade program. The discrepancy in text volume means that some students get practice with vastly more text than others, which can, in turn, mean a large discrepancy in practice with working memory, stamina, analyzing text structures, genre, or background knowledge.

The ratio of literary to informational texts was not equal in any program, with informational texts being underrepresented in all programs. Media was also only present as a read-aloud in one of the four publishers' programs. The lack of informational texts means that students are not getting the practice at how to read texts that they will encounter in content areas, and are needed to gain fluency in all literacy strategies. The findings in this study align with other studies that have found curriculum programs do not provide enough vocabulary exposure or informational text (Braker-Walters, 2014; Duke, 2000; Jeong et al., 2010).

This study highlights important differences between four core reading series that aren't immediately apparent and aligns with some of the critiques other scholars have raised about core reading series in the past. Stakeholders need this type of comparison and information to be able to make more informed decisions about curricula, and in order to understand further the reasons why so many readers stop progressing in reading starting in fourth grade. The findings in this study show that there is much progress needed for reading series to fully present evidence-based instructional curriculum, and legislation surrounding transparency and adherence to research-based practices in curriculum or the creation of public curricula, could help with that issue. In summary, the four programs had Lexile levels of text too low to be challenging to most students, too little informational text, very little media, and a big disparity in the volume of reading offered. No curriculum series will meet the needs of all students and cannot be the only resource used in a classroom. No program will fully solve the problem of reading performance, but the resources that students see make a difference in their reading development. More transparency about the differences between programs is necessary for schools to make informed decisions that best help their students.

LIMITATION & FURTHER RESEARCH

All of the programs in this study are commercial reading series that are bought by a school or district. The next step in analyzing the effects of core reading programs on students is to evaluate whether the cost of these programs is correlated with the volume of reading, Lexile levels, or levels of informational text. Research is also needed to directly evaluate how the differences between the series are affecting student reading growth and motivation.

There are several limitations to this study. While I have categorized certain texts as informational, other researchers or practitioners may consider content more important than structure and categorize texts differently. I am not a trained MetaMetrics analyst, so all Lexile levels in this study are only estimated levels. This study evaluates quantitative aspects of texts' complexity and does not claim to have found any text's full complexity. Rather, this study is intended to give information as an effort of transparency in the materials used in schools and to understand how these texts may affect instruction. This study also analyzed only read-aloud texts. Small group reading texts, texts designed for struggling readers, English language learners, or special education students, or independent reading texts may have different characteristics. Finally, this study does not consider the individual context in which the texts would be used, which could affect the effectiveness and practice they offer to students.

This study is important for gaining a basic understanding of the texts used in these programs, but it is not a stopping point. The findings of this study raise several more questions that need to be answered. Qualitative analysis needs to be done with both literary and informational texts from each program to determine if the Lexile levels reflect the complexity of the texts in the context of how they are used and what values the texts hold beyond Lexile, such as particular literary elements or content focused on social-emotional learning or representation of minority populations. Additional studies could also include the fidelity with which the teachers or districts use the programs, and what supplemental materials are being used in classrooms. Finally, and most pressing, information about which core reading series districts use is needed to establish whether there are patterns of inequity in what resources students have access to. This information is needed so that school and district decision-makers can make more informed decisions about curriculum.

REFERENCES

Ada, A. F., Beers, K., Campoy, F. I., Carroll, J. A., Clemens, A., Cunningham, A., Hougen, M. C., Howard, T. C., Izquierdo, E., Jago, C., Palmer, E., Probst, R., Templeton, S., & Washington, J. (2020). *HMH Into Reading*. Houghton Mifflin Harcourt.

Al Otaiba, S., Kosanovich-Grek, M. L., Torgesen, J. K., Hassler, L., & Wahl, M. (2005). Reviewing core kindergarten and first-grade reading programs in light of No Child Left Behind: An exploratory study. *Reading & Writing Quarterly, 21*(4), 377–400.

Allington, R. L., & McGill-Franzen, A. M. (2021). Reading volume and reading achievement: A review of recent research. *Reading Research Quarterly, 56*(S1), S231–S238.

Allyn, P., Hiebert, E. F., Pearson, P. D., & Vaughn, S. (2016). *ReadyGen*. Savvas Learning.

Alme, S. (2020). *What is the significance of reading aloud to middle school aged students in the language arts classroom as measured by the STAR Reading Assessment?* (Master's thesis, Hamline University). DigitalCommons@Hamline.

Amendum, S. J., Conradi, K., & Hiebert, E. H. (2018). Text complexity in elementary classrooms. *The Reading Teacher, 72*(1), 43–52. https://doi.org/10.1002/trtr.1683

Anderson, R. C., & Nagy, W. E. (1993). *The vocabulary conundrum* (Technical Report No. 570). Center for the Study of Reading.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*(3), 285–303.

Assari, S. (2019). Race, education attainment, and happiness in the United States. *International Journal of Epidemiologic Research, 6*(2), 76–82. https://doi.org/10.15171/ijer.2019.14

August, D., Bear, D., Bumgardner, K., Cerna, P., Echevarria, J., Fisher, D., Francis, D., Gibson, V., Hasbrouck, J., Justice, L., Mwaler, T., McTighe, J., Sesame Workshop, Shanahan, T., Soto, G., Spinrad, T., Tinajero, J., Walker-Dalhouse, D., & Zike, D. (2020). *Wonders K–6*. McGraw Hill Education.

Berkman, N. D., DeWalt, D. A., Pignone, M. P., Sheridan, S. L., Lohr, K. N., Lux, L., Sutton, S. F., Swinson, T., & Bonito, A. J. (2004). *Literacy and health outcomes* (Evidence Report/Technology Assessment No. 87; AHRQ Publication No. 04-E007-2). Agency for Healthcare Research and Quality.

Braker-Walters, B. A. (2014). Informational text and the Common Core: A content analysis of three basal reading programs. *SAGE Open, 4*(4), 1–8. https://doi.org/10.1177/2158244014555119

Brenner, D., & Hiebert, E. H. (2010). If I follow the teachers' editions, isn't that enough? Analyzing reading volume in six core reading programs. *The Elementary School Journal, 110*(3), 347–363.

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language, 109*, Article 104047.

Calfee, R. C., & Hiebert, E. H. (2011, July). Using cohort analyses to examine long-term effects of reading initiatives in California. Paper presented at the annual conference of the Society for the Scientific Study of Reading, St. Petersburg, FL, United States.

Calkins, L. (2015). *Units of study for teaching reading: Grades K–5*. Heinemann.

Common Core State Standards Initiative. (2022). *Key shifts in English Language Arts*. https://www.corestandards.org/other-resources/key-shifts-in-english-language-arts

Cunningham, J. W., Hiebert, E. H., & Mesmer, H. A. (2018). Investigating the validity of two widely used quantitative text tools. *Reading and Writing, 31*(4), 813–833.

Cunningham, A. E., & Stanovich, K. E. (2003). Reading can make you smarter! Principal components analysis of the relationship between print exposure and cognitive ability. *Journal of Educational Psychology, 95*(1), 1–15. https://doi.org/10.1037/0022-0663.95.1.1

DiMarco, B. (2022, October 6). Sounding out a better way to teach reading. *The New York Times*. https://www.nytimes.com/2022/10/06/education/learning/schools-teaching-reading-phonics.html

Dewitz, P., & Graves, M. F. (2021). The science of reading: Four forces that modified, distorted, or ignored the research finding on reading comprehension. *Reading Research Quarterly, 56*(S1), S131–S144. https://doi.org/10.1002/rrq.392

Dewitz, P., Jones, J., & Leahy, S. (2009). Comprehension strategy instruction in core reading programs. *Reading Research Quarterly, 44*(2), 102–126. https://doi.org/10.1598/RRQ.44.2.1

Dixon, P., LeFevre, J. A., & Twilley, L. C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology, 80*(4), 465–472. https://doi.org/10.1037/0022-0663.80.4.465

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*(2), 202–224. https://doi.org/10.1598/RRQ.35.2.2

Duke, N. K., & Block, M. K. (2012). Improving reading in the primary grades. *The Future of Children, 22*(2), 55–72. https://doi.org/10.1353/foc.2012.0017

Duke, N. K., & Cartwright, K. B. (2021). The science of reading progresses: Communicating advances beyond the simple view of reading. *Reading Research Quarterly, 56*(S1), S25–S44. https://doi.org/10.1002/rrq.411

Durán, L., & Hikida, M. (2022). Making sense of reading's forever wars. *Phi Delta Kappan, 103*(8), 14–19. https://doi.org/10.1177/00317217221104633

edReports.org. (2022). The state of the instructional materials market: 2021 report. https://www.edreports.org/resources/article/state-of-the-instructional-materials-market-2021-the-availability-and-use-of-aligned-materials

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*(2), 174–187. https://doi.org/10.2307/747863

Faverio, M., & Perrin, A. (2022, January 6). Three-in-ten Americans now read e-books. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2022/01/06/three-in-ten-americans-now-read-e-books

Fisher, D., Flood, J., Lapp, D., & Frey, N. (2004). Interactive read-alouds: Is there a common set of implementation practices? *The Reading Teacher, 58*(1), 8–17. https://doi.org/10.1598/RT.58.1.1

Foorman, B. R., Francis, D. J., 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*(2), 167–197.

https://doi.org/10.1207/s1532799xssr0802_2

Foorman, B. R., Schatschneider, C., Eakin, M. N., Fletcher, J. M., Moats, L. C., & Francis, D. J. (2006). The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high-poverty schools. *Contemporary Educational Psychology, 31*(1), 1–29. https://doi.org/10.1016/j.cedpsych.2005.05.001

Fountas, I. C., & Pinnell, G. S. (2017). *Ten characteristics of texts for guided reading*. Heinemann.

Fountas, I. C., & Pinnell, G. S. (2020). *Fountas & Pinnell Classroom™ Collection*. Heinemann.

Frantz, R. S., Starr, L. E., & Bailey, A. L. (2015). Syntactic complexity as an aspect of text complexity. *Educational Researcher, 44*(7), 387–393. https://doi.org/10.3102/0013189X15584399

Frey, N., Lapp, D., & Fisher, D. (2016). *Text complexity: Stretching readers with texts and tasks*. SAGE Publications.

Fry, E. B. (2002). Readability versus leveling. *The Reading Teacher, 56*(3), 286–291. https://doi.org/10.1598/RT.56.3.7

Gee, J. P. (2018). Reading as situated language: A sociocognitive perspective. In D. L. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (6th ed., pp. 105–117). Routledge.

Gill, S. R. (2009). The complexity of nonfiction texts. *Journal of Adolescent & Adult Literacy, 53*(2), 97–100. https://doi.org/10.1598/JAAL.53.2.1

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*(1), 6–10. https://doi.org/10.1177/074193258600700104

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-Metrix measures text characteristics at multiple levels of language and discourse. *The Elementary School Journal, 115*(2), 210–229. https://doi.org/10.1086/674714

Graves, M. F., Elmore, J., & Fitzgerald, J. (2019). The vocabulary of core reading programs. *The Elementary School Journal, 119*(3), 386–416. https://doi.org/10.1086/701732

Greene Brabham, E., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94*(3), 465–473. https://doi.org/10.1037/0022-0663.94.3.465

Gunning, T. G. (2003). The role of readability in today's classrooms. *Topics in Language Disorders, 23*(3), 175–189. https://doi.org/10.1097/00011363-200307000-00002

Hatcher, P. (2000). Predictors of Reading Recovery book levels. *Journal of Research in Reading, 23*(1), 67–77. https://doi.org/10.1111/1467-9817.00104

Hernandez, D. J. (2011). *Double jeopardy: How third-grade reading skills and poverty influence high school graduation*. Annie E. Casey Foundation. https://www.aecf.org/resources/double-jeopardy/

Hiebert, E. H., Scott, J. A., Castaneda, R., & Spichtig, A. N. (2019). An analysis of the features of informational texts used in reading programs. *Elementary School Journal, 119*(3), 449–470. https://doi.org/10.1086/701731

Hiebert, E. H. (2014). The forgotten reading proficiency: Stamina in silent reading. In E. H. Hiebert (Ed.), *Stamina, silent reading, & the Common Core State Standards* (pp. 5–17). TextProject.

Houghton Mifflin Harcourt. (2022). *HMH Reading Inventory*. https://www.hmhco.com/programs/reading-inventory

Hutton, J. S., Horowitz-Kraus, T., Mendelsohn, A. L., DeWitt, T., Holland, S. K., & the C-MIND Authorship Consortium. (2015). Home reading environment and brain activation in preschool children listening to stories. *Pediatrics, 136*(3), 466–478. https://doi.org/10.1542/peds.2015-0359

International Literacy Association. (2018). *The power and promise of read-alouds and independent reading* [Literacy leadership brief]. Newark, DE: Author. https://literacyworldwide.org/resources

Jeong, J., Gaffney, J. S., & Choi, J. O. (2010). Availability and use of informational texts in second-, third-, and fourth-grade classrooms. *Research in the Teaching of English, 44*(4), 435–456. https://www.jstor.org/stable/40863242

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

Layne, S. (2023). *In defense of read-aloud: Sustaining best practice*. Routledge.

Leu, D. J., Kinzer, C. K., Coiro, J., Castek, J., & Henry, L. A. (2013). New literacies and the new literacies of online reading comprehension: A dual-level theory. In D. L. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 1150–1181). International Reading Association.

Locher, F. M., & Pfost, M. (2020). The relation between time spent reading and reading comprehension throughout the life course. *Journal of Research in Reading, 43*(1), 57–77. https://doi.org/10.1111/1467-9817.12289

Logan, J. A. R., Justice, L. M., Yumus, M., & Chaparro-Moreno, L. J. (2019). When children are not read to at home: The million word gap. *Journal of Developmental & Behavioral Pediatrics, 40*(5), 383–386. https://doi.org/10.1097/DBP.0000000000000657

Maguire, M. J., Schneider, J. M., Middleton, A. E., Ralph, Y., Lopez, M., Ackerman, R. A., & Abel, A. D. (2018). Vocabulary knowledge mediates the link between socioeconomic status and word learning in grade school. *Journal of Experimental Child Psychology, 166*, 679–695. https://doi.org/10.1016/j.jecp.2017.10.003

Marr, M. B., & Gormley, K. (1982). Children's recall of familiar and unfamiliar text. *Reading Research Quarterly, 18*(1), 89–104. https://doi.org/10.2307/747539

Mathewson, T. G. (2020, February 26). New data: Even within the same district some wealthy schools get millions more than poor ones. *The Hechinger Report.*

McGill-Franzen, A., Zmach, C., Solic, K., & Zeig, J. L. (2006). The confluence of two policy mandates: Core reading programs and third-grade retention in Florida. *The Elementary School Journal, 107*(1), 67–91. https://doi.org/10.1086/503740

McGraw Hill. (2023). *Authentic literature brochure: Wonders 2023*. McGraw Hill.

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*(3), 235–258. https://doi.org/10.1002/RRQ.030

MetaMetrics, Inc. (2022). *Lexile Text Analyzer*. Lexile & Quantile Tools. https://hub.lexile.com/analyzer

MetaMetrics, Inc. (2023). *What are Lexile student measures by grade?* Lexile & Quantile Tools. https://hub.lexile.com/lexile-grade-level-charts

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*(2), 267–296. https://doi.org/10.1037/a0021890

Moss, B. (2008). The information text gap: The mismatch between non-narrative text types in basal readers and 2009 NAEP recommended guidelines. *Journal of Literacy Research, 40*(2), 201–219. https://doi.org/10.1080/10862960802070411

National Center for Education Statistics. (2022). *The Nation's Report Card: 2022 reading assessment*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/nationsreportcard/reading/

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2023). *Common Core State Standards for English language arts*. Authors. https://www.corestandards.org/ELA-Literacy/

National Reading Panel (US), National Institute of Child Health and Human Development (US). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. U.S. Government Printing Office. https://www.nichd.nih.gov/publications/pubs/nrp

Nevo, E., & Breznitz, Z. (2011). Assessment of working memory components at 6 years of age as predictors of reading achievements a year later. *Journal of Experimental Child Psychology, 109*(1), 73–90. https://doi.org/10.1016/j.jecp.2010.10.009

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). SAGE Publications.

Neuman, S. B., & Wright, T. S. (2015). *All about words: Increasing vocabulary in the Common Core classroom, PreK–2*. Teachers College Press.

Pearson, P. D. (2013). Research foundations of the Common Core State Standards in English language arts. In S. Neuman & L. Gambrell (Eds.), *Quality reading instruction in the age of Common Core Standards* (pp. 237–262). International Reading Association.

Perrin, A., & Anderson, M. (2019, April 10). Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018

Pilonieta, P. (2010). Instruction of research-based comprehension strategies in basal reading programs. *Reading Psychology, 31*(2), 150–175. https://doi.org/10.1080/02702710903291269

Rocha, M. S. B., & Ponczek, V. (2011). The effects of adult literacy on earnings and employment. *Economics of Education Review, 30*(4), 755–764. https://doi.org/10.1016/j.econedurev.2011.03.001

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33–58). Lawrence Erlbaum.

Sanden, S., MacPhee, D. A., Hartle, L., Poggendorf, S., & Zuiderveen, C. (2022). The status of phonics instruction: Learning from the teachers. *Reading Horizons: A Journal of Literacy and Language Arts, 61*(1), Article 5. https://scholarworks.wmich.edu/reading_horizons/vol61/iss1/5

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01203.x

Scholastic, Inc. (2021). *Lexile collections*. Scholastic: The Teacher Store. https://shop.scholastic.com/teachers-ecommerce/teacher/shops/lexile-leveled-collections.html

Smith, D. R., Stenner, A. J., Morabin, I., & Smith, M. (1989). *The Lexile scale in theory and practice: Final report for NIH grant IID-19448*. U.S. Department of Education.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1

Statista. (2023). Social media marketing usage rate in the United States from 2013 to 2022. *Statista*. https://www.statista.com/statistics/203513/usage-trands-of-social-media-platforms-in-marketing/

Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2006). How accurate are Lexile text measures? *Journal of Applied Measurement, 7*(3), 307–322. https://www.researchgate.net/publication/233800006

The New London Group. (2000). A pedagogy of multiliteracies: Designing social futures. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Literacy learning and the design of social futures* (pp. 9–37). Routledge.

The NPD Group, Inc. (2020). *Gamer segmentation: 2020 syndicated report*. The NPD Group, Inc. https://www.npd.com/wp-content/uploads/2021/08/npd-gamer-segmentation-report-whitepaper.pdf

Toyama, Y., Hiebert, E. H., & Pearson, P. D. (2017). An analysis of the text complexity of leveled passages in four popular classroom reading assessments. *Educational Assessment, 22*(3), 139–170. https://doi.org/10.1080/10627197.2017.1346648

Tuma, A. P., Doan, S., Lawrence, R. A., Henry, D., Kaufman, J. H., Setodji, C. M., Grant, D., & Young, C. J. (2020). *American Instructional Resources Survey: 2019 technical documentation and survey results (RR-4402-BMGF/SFF/OFF)*. RAND Corporation. https://doi.org/10.7249/RR4402

U.S. Government Accountability Office. (2022). *K–12 education: Student population has significantly diversified, but many schools remain divided along racial, ethnic, and economic lines (GAO-22-104737)*. U.S. Government Accountability Office. https://www.gao.gov/products/gao-22-104737

Valentine, J. W., Clark, D. C., Hackmann, D. G., & Petzko, V. N. (2002). *A national study of leadership in middle level schools. Volume I: A national study of middle level leaders and school programs*. National Association of Secondary School Principals.

Wang, J., Liu, C., & Cai, Z. (2022). Digital literacy and subjective happiness of low-income groups: Evidence from rural China. *Frontiers in Psychology, 13*, Article 1045187. https://doi.org/10.3389/fpsyg.2022.1045187

Wright, T. S., & Neuman, S. B. (2013). Vocabulary instruction in commonly used kindergarten core reading curricula. *The Elementary School Journal, 113*(3), 386–408. https://doi.org/10.1086/669089