



Textual demands of passages in three english/ language arts common core assessments and one core literacy program for intermediate elementary grades

Timothy G. Morrison, Brad Wilcox, Erica Murdoch & Lauren Bird

To cite this article: Timothy G. Morrison, Brad Wilcox, Erica Murdoch & Lauren Bird (2018) Textual demands of passages in three english/language arts common core assessments and one core literacy program for intermediate elementary grades, Reading Psychology, 39:4, 362-383, DOI: [10.1080/02702711.2018.1442969](https://doi.org/10.1080/02702711.2018.1442969)

To link to this article: <https://doi.org/10.1080/02702711.2018.1442969>



Published online: 15 Mar 2018.



Submit your article to this journal [↗](#)



Article views: 3



View related articles [↗](#)



View Crossmark data [↗](#)



TEXTUAL DEMANDS OF PASSAGES IN THREE ENGLISH/LANGUAGE ARTS COMMON CORE ASSESSMENTS AND ONE CORE LITERACY PROGRAM FOR INTERMEDIATE ELEMENTARY GRADES

**TIMOTHY G. MORRISON, BRAD WILCOX, ERICA MURDOCH,
and LAUREN BIRD**

Department of Teacher Education, Brigham Young University, Provo, Utah

The Common Core has emphasized reading for comprehension, including making inferences. However, little is known about the textual demands found within assessment and instructional passages that are promoted as being in line with Common Core expectations. The purpose of this content analysis was to identify the readability levels, passage length, and types of inferences found within 72 fiction and nonfiction passages from the PARCC, Smarter Balanced, and SAGE Common Core assessments and one core literacy program, Journeys, for students in grades three, four, and five. Results show that the readability levels of passages in this study were written at higher levels than their associated grade levels. Anaphoric relationships were by far the most common inference type found in the texts, followed by Prior Knowledge and Retrospective inferences.

Reading is a complex activity that involves the coordinated use of many processes—both lower-level (e.g., phonics, phonemic awareness, fluency) and higher-level (e.g., vocabulary knowledge and inferential abilities). Orchestrated use of these and other processes allow individuals to understand the texts they read.

As they defined reading comprehension, the RAND Reading Study Group (2002) focused on three influential factors: the reader, the text, and the contextual activity. Readers vary in their knowledge of and ability to use phonological, decoding, word reading, and vocabulary skills, as well as the ability to complete these processes rapidly, accurately, and with expression (i.e., fluency). Texts also vary in the demands they present to readers. Length, reading level, quality of the writing, text structure, and

Address correspondence to Timothy G. Morrison, Department of Teacher Education, Brigham Young University, 210-P MCKB, Provo, Utah 84602, USA. E-mail: tim_morrison@byu.edu

vocabulary load are prominent text factors. Common activity factors include reasons for reading (assigned or recreational), careful reading or skimming of the text, and the time constraints to complete the reading.

The Common Core State Standards (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) provide expectations for readers that address many issues included in the RAND Reading Study Group definition. Beginning in second grade students are expected to read texts of increasing difficulty so that by the end of 12th grade, students will be ready to deal with the level of texts they will encounter as they begin college. The CCSS include a figure that shows a staircase of text complexity, indicating that students will need to deal with increasingly difficult texts that are written several grade levels higher than what they have been expected to read in the past (CCSS, Appendix A, p. 9).

The Reading anchor standards in the CCSS deal with issues of comprehension at various levels. The standards require students to both “determine what the text says explicitly” and to “make logical inferences.” Students are instructed to “cite specific textual evidence” and to make “conclusions drawn from the text.” They are expected to “analyze how and why individuals, events, and ideas develop and interact over the course of a text.” Readers are also expected to “determine central ideas or themes” in texts. These are examples of standards that expect a great deal from readers in terms of reading comprehension.

With the adoption of the CCSS, publishers have produced core literacy programs that are said to meet the high expectations for reading comprehension in the CCSS. The three most-commonly used programs are *Wonders* (McGraw Hill Education, 2017), *Reading Street* (Scott Foresman-Pearson, 2017), and *Journeys* (Houghton Mifflin/Harcourt, 2009). All three include anthologies for each grade level, instructional materials for teachers, assessment measures, and additional online resources. Since these three programs are very similar, this study examined only one, *Journeys*.

We do not know if the passages in core literacy programs students read are of sufficient length and difficulty to match CCSS expectations. Similarly, little is known about how the instructional guidelines affect student comprehension.

In addition to these instructional efforts, a great deal of work has gone into developing instruments to measure students' abilities to achieve CCSS expectations. Two national Common Core assessments, PARCC (Partnership for Assessment of Readiness for College and Careers, 2017) and *Smarter Balanced Assessment Consortium* (2017), are used in many states to determine whether students are meeting Common Core goals. These assessment measures were developed through involvement by consortia of states and with considerable federal investment. Additionally, some states have created Common Core assessments similar to those two larger endeavors. One of these is the *SAGE Core Assessment* (Student Assessment of Growth and Excellence, 2017), that is used in Utah. Despite the effort to create these assessments, one issue that is not known is the length and reading levels of passages in these tests. Also unknown are the types of cognitive processing that the passages and tests require of students when teaching and measuring comprehension.

Kintsch's Construction-Integration (C-I) model (Kintsch, 1998, 2004) attempts to explain the coordination of the many processes identified by the RAND Reading Study Group. In the C-I model, readers begin the construction phase of the comprehension process at the surface level by decoding text through application of letter-sound relationship skills and use of semantic and syntactic knowledge. Readers create basic meaning through a textbase process in which textual information triggers activation of a limited amount of background knowledge. During this level of reading, activation usually happens automatically, but does not result in full understanding. The next phase is integration, during which readers use information in the text combined with deeper use of background knowledge to make inferences that allow readers to create a more complete mental representation of the text. This is termed the *situation model*. During this phase, readers use prior knowledge to support connections between information presented in text and their own personal knowledge, as well as other connections within text.

The Kintsch model is unique among comprehension theories because it highlights inferences as vital in reading comprehension. The model claims that readers' success in comprehending a text depends in large part on how they resolve the inference demands that are present in passages they read. As readers create situation

models during reading, they coordinate their own abilities to make inferences, draw on prior knowledge and experiences, and understand text features. They also approach the reading context with their dispositions toward both the topic and the reading process (RAND Reading Study Group, 2002).

We should expect passages within instructional materials from core literacy programs to provide opportunities for students to make these inferences. We should also expect assessment passages to invite students to demonstrate ability to make inferences that lead to deep comprehension.

Literature Review

Of the three aspects of reading comprehension outlined by the RAND Reading Study Group, this study focused on two: (1) text length and readability; and (2) processes used by the reader. Both of these factors influence how well individuals can comprehend what they read.

Readability Levels and Length of Passages

Demands of texts contribute to the comprehension process. The CCSS deal with text demands with its focus on text complexity. To examine how challenging texts are, the Common Core recommends using three factors: quantitative dimensions, qualitative aspects, and reader and task considerations.

Quantitative factors are commonly addressed using readability formulas. These formulas have been in use for more than 75 years. Traditionally, these formulas use word length, sentence length, and word familiarity to determine how challenging a text will be for readers. Common readability formulas include Flesch-Kincaid, Dale-Chall, and the Lexile Framework. The CCSS recommend using one or more readability formulas to make an initial estimate of text difficulty.

When using the qualitative dimension, teachers are advised to examine aspects of text that are not as quantifiable. Teachers should examine the structure of the text, language conventionality and clarity, the knowledge demands the text presents, including life experience, cultural/literary, and content/discipline knowledge.

Reader and task considerations constitute the third factor to use when determining text difficulty. Among many factors to consider are readers' purposes for reading, their interest in the content, and their prior experiences with the content of the text.

Some texts may have a single, simple theme and others may have multiple or complex themes. Some texts are structured in straightforward ways, while authors of other texts employ less common structural devices (e.g., flashbacks or documents that tell the story). Some texts require the reader to use a great deal of cultural understanding or technical knowledge. However, two factors that are constant across multiple contexts and genre differences (e.g., fiction, nonfiction, fantasy, and historical fiction) are the difficulty and length of texts.

Stenner, Koons, and Swartz (2010) documented that the demands that college, careers, and citizenship place on readers have held steady or increased over the last 50 years. Yet, the CCSS points out that students in high school are generally unprepared for post-secondary reading (CCSS, 2010) and suggests that students be expected to read texts above grade level. For example, the high-end expectation for the 2nd–3rd grade band is that students read texts a fifth-grade level, 5.15 (ATOS, 2018) and 5.34 (Flesch-Kincaid, 2018). Similarly, the high-end expectation for 6th–8th graders is 9.98 (ATOS) and 10.334 (Flesch-Kincaid). The instructional materials and assessments that have been designed with Common Core expectations in mind have not been examined extensively to determine the reading level of passages.

Another text factor that affects comprehension is the length of passages. Hiebert (2014) has pointed out that readers experience fatigue when they read long passages. She identified six stamina patterns among students: (1) nonstarters, (2) quitters after two passages, (3) quitters after three passages, (4) students who failed to comprehend text; (5) persisters who obtained a minimal level of comprehension, and (6) comprehenders, students who comprehended well throughout the all passages. Clearly, length of text is connected to what Boushey and Moser (2006) labeled reading stamina. Little is known about the length of passages in CCSS instructional and assessment materials.

Reader Processes

In addition to interactions between text difficulty and length, comprehension is affected by reader processes. The Kintsch model acknowledges readers' motivations and purposes to read (e.g., assigned reading, pleasure reading, and reading to solve a problem) affect how they approach the task and how successful they will be in creating meaning. If readers can choose texts about topics of high interest, their increased motivation can compensate to a degree for lower ability. According to the Kintsch model, ability to make inferences is key to completing all these processes.

An inference is defined as "any piece of information that is not explicitly stated in a text" (McKoon & Ratcliff, 1992, p. 440). Authors of texts imply relationships among ideas in text and make reference in text to readers' prior knowledge, expecting that readers will be able to connect the ideas to create meaning. Readers need to draw inferences to completely understand authors' messages. Sometimes inferences are made during reading and sometimes after, but either way "the ability to draw inferences remains the cornerstone of reading comprehension" (Kendeou, 2015, p. 160). It is important to understand various types of inferences, as well as the context in which they are found. In this case, in Common Core assessments and one core literacy program.

Some inferences are simple and automatic, while others require more effort from readers. Researchers have created various taxonomies of inferences (Graesser, Li, & Feng, 2015; Van den Broek, Beker, & Oudega, 2015), but "the most relevant distinction is between local cohesion inferences and global coherence inferences, also referred to as text-connecting (or bridging) and gap-filling inferences" (Oakhill, Cain, & McCarthy, 2015, p. 143). Inferences require readers to make connections within text locally (anaphoric inferences), use a combination of prior knowledge and text information (background knowledge inferences), and make within-text connections across large sections of text (retrospective inferences). An additional type of inference requires readers to anticipate what will happen next in text (predictive inferences).

ANAPHORIC INFERENCE

Anaphoric relationships are instances in text when one word replaces another word in close proximity. For example, consider the following sentences: “John went to the store. He bought some milk.” The word *he* in the second sentence replaces the antecedent *John* in the first sentence. Many anaphoric relationships are easy for most readers to resolve automatically. However, some are more difficult for readers and require closer consideration. For example, this excerpt about efforts to reintroduce wolves in the Yellowstone National Park (Bishop, 2008) contains several anaphoric relationships: “But are the animals truly safe? Scientists are doing all they can to make sure they are.” (p. 2). The first *they* in the second sentence refers to scientists, the second *they* refers to animals, and the word *are* refers to being *truly safe*. Incidentally, the word *animals* in the first sentence refers to *wolves*, which are described earlier in the passage.

BACKGROUND KNOWLEDGE INFERENCE

In addition to anaphoric inferences, readers make some inferences by combining prior knowledge they already possess and text information to better understand the author’s message. The author may give information about the setting of a story without specifically naming the place. For example, in the book, *The Sweetest Fig* (Van Allsburg, 1993), the author provides the names of main characters (*Bibot* and *Marcel*), a holiday (*Bastille Day*), and a landmark (*Eiffel Tower*) to let the reader know that the story is set in Paris, France, which is never directly stated. By doing so, the author assumes that readers will use text clues to infer the setting of the story based on background knowledge.

RETROSPECTIVE INFERENCE

Along with anaphoric and prior knowledge inferences, readers also make retrospective inferences, requiring them to combine text information in at least two places in text separated by multiple pages. In the children’s book, *The Three Princes* (Kimmel, 1984), a princess sends three princes out on a year-long quest to find the rarest thing they can find. They travel together for many days before going in different directions. After one year, they reunite where they first separated, far from the palace. Much has happened in the interim, so many readers may not remember

that the princes are still far away from the princess who is ill and will probably die before the next morning. One prince cries out, “Even if I rode all night, I could never arrive by morning” (np). In order to infer why he cannot arrive in time, the reader needs to make a retrospective inference, recalling parts of text encountered earlier in the book. Retrospective inferences are vital when drawing conclusions or examining cause-effect relationships.

PREDICTIVE INFERENCES

In addition to these three types of inferences, readers are also asked at times to make predictions during reading. Authors provide information to invite readers to think about what may happen next in the text. For example, in the short story, “Two Were Left” (Cave, 1980), a young boy, Noni, and his beloved husky dog, Nimuk, were stranded on an ice floe in the Arctic Ocean. They both became desperate because of their isolation on the floating island. When readers read, “One of us will soon be eating the other, Noni thought. So . . .” (np), they are invited to predict what will happen. Authors frequently provide a context that allows and expects readers to predict what will next occur. Such predictions are considered to be a type of inference that requires readers to use text information and prior knowledge to anticipate coming events.

In this content analysis, prediction inferences were not formally examined, because of the difficulty of labeling them without the involvement of students. An inference might or might not be labeled as predictive, depending on how a reader may choose to resolve it. Qualitative methods, including interviews and observations, would be needed to explore how individual readers go about resolving predictive inferences.

Purpose of the Study

The purpose of this study was to identify passage readability and length as well as types of inferences that are found within passages from three Common Core assessments and one core literacy program for students in grades three, four, and five. The following research questions were addressed in this study:

1. What are the readability levels and lengths of passages, fiction and nonfiction, used for testing and instructional purposes?
2. What inference types are present in passages, fiction and nonfiction, used for testing and instructional purposes?

Methodology

This study analyzed three assessment instruments used to measure students' mastery of the Common Core State Standards and one major core literacy program used to teach the Standards. We examined passages from the *PARCC Core Assessment* (Partnership for Assessment of Readiness for College and Careers, 2017), the *Smarter Balanced Assessment Consortium* (2017), the *SAGE Core Assessment* (Student Assessment of Growth and Excellence, 2017), and the *Journeys* core literacy program (Houghton Mifflin Harcourt, 2009). Specifically, we examined passages in these materials to measure word counts and readability levels of passages and to identify the types and frequencies of inferences found in them.

We randomly selected three narrative and three informational passages for grade levels three, four, and five from all four sources ($n = 72$ passages) for analysis. Our rationale for choosing from each genre was to be sure that both major categories of text were represented. However, the decision to select only six passages from each grade level was to keep the data manageable, since we were drawing from four different sources. The selection process was random, but was not meant to be proportional. Our limited sample was not meant to provide a comprehensive review, but it was appropriate to identify the types of inferences included for students to read. This decision is in line with direction given by Hoffman, Wilson, Martinez, and Sailors (2011). Intermediate grade levels were chosen because students at this point are typically reading with greater fluency, allowing for increased attention to comprehension. During grades K through two, students focus greater attention on word identification tasks, leaving fewer mental resources to create meaning (Stanovich, 1980).

Readability and Passage Length

For each passage, we identified the word count and readability level, using an online source (readable.io, 2017) that averaged reading

levels using multiple formulas, as suggested by Hiebert (2010/2011). This website employs eight commonly-used readability formulas to estimate text difficulty, including the highly utilized Flesch-Kincaid, New Dale-Chall, and Spache measures. This website's use of eight formulas is superior to other popular websites drawing on only four to six readability formulas. We averaged word counts and readability levels by grade level, genre, and passage source.

Identification of Inferences

We also identified all instances of three inference types: anaphoric, background knowledge, and retrospective inferences. Two researchers independently read the 72 selected passages and identified all instances of the three inference types. When disagreements arose, the researchers discussed their judgements and reached agreement on inference type. After those researchers who made these initial decisions came to agreement, they presented their results to the lead investigator and additional discussions ensued. Consensus was reached among the three investigators who used the following rules to identify anaphoric, background, and retrospective inference types.

ANAPHORIC INFERENCE

In order to make anaphoric inferences, the reader is expected to understand when a word or phrase in the text is replaced with another word or phrase. In a 4th-grade fiction passage called, "Just Like Home," in the *PARCC* assessment, the following sentence was classified an anaphoric inference: "The only thing Priya liked about her new school was art." *Priya* is replaced by *her*. Similarly, in a 3rd-grade nonfiction text in the *Smarter Balanced* assessment ("Life in Space"), the following was also classified as an anaphoric inference: "Astronauts' other muscles and their bones can also get weaker. This is because they do not have to work as hard to move the astronaut's body." *They* in the second sentence refers to *muscles* and *bones* rather than *astronauts*.

BACKGROUND INFERENCE

To make background knowledge inferences the reader is expected to have knowledge of a concept or event not

explained or not yet explained in the text and feel triggered to use that knowledge to understand the text. Background knowledge inferences are more than just understanding vocabulary words. They refer to the reader's prior knowledge of specific time periods, historical circumstances, places, people, languages, animals, events, concepts, cultural knowledge, and so on. They also can connect the reader to his or her personal experiences or understanding of the human condition. In this study, instances of figurative language, including idioms, similes, and metaphors, were not included in category of background knowledge; rather, they were considered to be an issue of vocabulary. The link between figurative language and vocabulary development (rather than inference) has been well established (Beck, McKeown, & Kucan, 2002; Blachowicz & Fisher, 1996; Lazar, 1996).

An example of a background knowledge inference is from a 5th-grade fiction passage ("Teachers' Lounge") in the *Smarter Balanced* assessment: "Near the side gate, they saw Eric O'Flanagan, the worst bully in the school." Readers must infer characteristics of Eric based on their background knowledge of how school bullies behave. Another background knowledge inference is from a 4th grade nonfiction passage from *Journeys* ("My Brother Martin") about Martin Luther King, Jr.: "And when he was much older, my brother M.L. dreamed a dream." Readers must use background knowledge to infer that Dr. King gave a speech about his dream, or vision of the future, and was not just talking about a dream at night.

RETROSPECTIVE INFERENCES

To make retrospective inferences, the reader is expected to integrate current input from the text with text information that has occurred previously within the same text. Almost always, this takes place through active, strategic processes, but those processes can vary from simply remembering a previous part of the text to actively searching through pages of text to find a specific reference. Retrospective inferences fall into one of three categories: (1) Returning to one place in one text to make a connection; (2) Returning to two or more places in one text; and (3) Returning to one or more places in multiple texts.

An example of a retrospective inference is found in “Saving Buster,” a 3rd-grade fiction text from *Journeys*: “‘Okay!’ Mrs. Parker laughed. ‘But this is the last time! I’m exhausted!’” In order to understand this text, readers must go back to an earlier part of the text, in which Mrs. Parker yelled, “Go get it, boy!” and started a game of fetch. The reader must infer that Mrs. Parker and Buster had been playing this game for a while, and that the dog is not as tired of the game as Mrs. Parker. Another example of a retrospective inference comes from a 4th-grade fiction text called, “Me and Uncle Romie,” in *Journeys*: “But I wasn’t worried anymore. Aunt Nanette would make my birthday special.” To understand this sentence, readers must recall that earlier in the text they learned that James is on a train traveling from North Carolina to New York City to visit his Aunt Nanette and Uncle Romie. He had never met them before, and was a little concerned, because he had left behind his friend, his dad, and his mom, who would soon have twin babies. Readers can infer that James was worried that he might not have fun on this summer vacation, especially since his birthday was coming up. Readers need to connect the previous information in the text with the fact that he is not worried anymore because things have occurred in the story that have allowed him to trust his aunt.

The above definitions and parameters guided all decisions in classifying inference types. After adjustments were made to assignment of inference types, frequency counts and percentages were calculated, and comparisons were made across sources of passages, genres, and grade levels.

Results

Results are based on the 72 examined passages included in three Common Core assessments and one core literacy program. We first show results related to the reading levels and word counts of passages. Second, we present results about the types of inferences present in selected assessment and instructional materials.

Readability Levels and Length of Passages

Table 1 presents mean word counts and readability levels for passages, both fiction and nonfiction, from all four sources.

Generally, the average word counts in the assessment passages across the three grade levels were longer for the fiction than for nonfiction passages. These ranged from 342 words (4th-grade *SAGE* fiction passages) to 863 words (3rd-grade *PARCC* fiction passages). Nonfiction passages were shorter than fiction passages (292 for 4th-grade *SAGE* to 798 words for 5th-grade *PARCC*). In the core literacy program, *Journeys*, the results were the opposite, with average fiction word counts of 724 (4th grade), 929 (3rd grade) and 1379 (5th grade) and nonfiction word counts of 1143 (3rd grade), 1167 (4th grade), and 1170 (5th grade). As would be expected, the mean word counts for instructional materials were larger than for assessment materials.

In almost all cases mean readability levels of both assessment and instructional passages were higher than the identified three grade levels for both fiction and nonfiction text (see Table 1). In the 24 total cases (not passages), there were only three exceptions to this pattern: all fifth-grade fiction readability levels were approximately at grade level (*Journeys*, 5.57; *PARCC*, 5.67; and *Smarter Balanced*, 5.97). Similar to word count, readability levels increased sequentially by grade level for only three of the eight

TABLE 1. Word Count and Readability of Fiction and Non-Fiction Passages in One Core Literacy Program and Three Common Core Assessments.

	Word Count		Readability Level	
	Fiction	Non-Fiction	Fiction	Non-Fiction
Smarter Balanced				
Grade 3	482	319	5.33	6.03
Grade 4	513	474	5.70	7.43
Grade 5	586	645	5.97	8.13
PARCC				
Grade 3	863	730	4.70	7.17
Grade 4	781	755	5.70	7.37
Grade 5	851	798	5.67	7.27
SAGE				
Grade 3	783	471	6.43	4.93
Grade 4	342	292	5.83	6.87
Grade 5	417	572	7.70	8.10
Journeys				
Grade 3	929	1143	4.97	6.33
Grade 4	724	1167	6.23	8.10
Grade 5	1379	1170	5.57	6.93

sources (fiction and nonfiction passages in *Smarter Balanced*, *PARCC*, *SAGE*, and *Journeys*). Two of the three were nonfiction. Except in one case (*SAGE* grade 3), average readability levels of nonfiction texts were higher than fiction texts.

Types of Inferences

All three inference types were found in all passages we examined. Table 2 presents the mean instances and percentages of inferences in fictional passages in all four sources. As expected, anaphoric relationships were by far the most common inference type found in the texts with percentages ranging from a low of 60.66% (4th-grade *SAGE*) to a high of 80.23% (4th-grade *Smarter Balanced*). Prior Knowledge inferences ranged from 5.43% (3rd-grade *Smarter Balanced*) to 25.63% (4th-grade *SAGE*), and Retrospective inferences ranged from 8.40% (4th-grade *Journeys*) to 14.75% (5th-grade *Smarter Balanced*).

TABLE 2. Mean Instances and Percentages of Inferences in Fictional Passages in Three Common Core Tests and One Core Literacy Program.

	Anaphora		Prior Knowledge		Retrospective	
	Mean Instances	Mean %	Mean Instances	Mean %	Mean Instances	Mean %
Smarter						
Balanced						
Grade 3	53.67	79.97	4.67	5.43	10.67	14.59
Grade 4	61.67	80.23	5.67	7.35	9.67	12.40
Grade 5	79.00	74.35	12.00	10.90	15.67	14.75
PARCC						
Grade 3	97.33	78.55	10.00	8.36	17.00	13.43
Grade 4	111.67	75.93	19.67	14.12	14.67	9.92
Grade 5	134.67	78.67	19.33	11.12	17.33	10.20
SAGE						
Grade 3	98.33	71.88	23.33	16.58	16.00	11.53
Grade 4	39.67	60.66	12.67	25.63	8.33	13.70
Grade 5	54.00	75.60	7.33	10.84	10.00	14.27
Journeys						
Grade 3	121.00	73.32	24.67	15.07	19.33	11.61
Grade 4	229.33	79.83	31.33	11.77	23.00	8.40
Grade 5	178.67	71.39	35.33	14.90	33.00	13.72

Table 3 refers to mean instances and percentages of inferences in nonfiction passages in all four sources. As with fiction, anaphoric inferences occurred much more frequently than the other two inference types, but the percentages of occurrence were not as large as in fictional texts, ranging from 47.63% (4th-grade *Smarter Balanced*) to 66.52% (4th-grade *Journeys*). Background Knowledge inferences were present in nonfiction texts much more frequently than fiction (21.90% of all inferences in 5th-grade *Smarter Balanced* passages to 35.51% in 4th-grade *SAGE* passages). Although nonfiction texts included more retrospective inferences than fiction passages, the range was not substantially different (10.68% of all inferences in 4th-grade *Journeys* passages to 19.59% in 4th-grade *Smarter Balanced* passages).

Discussion

The results for length of passages showed inconsistency in that non-fiction passages were longer in instructional materials and

TABLE 3. Mean Instances and Percentages of Inferences in Non-Fictional Passages in Three Common Core Tests and One Core Literacy Program.

	Anaphora		Prior Knowledge		Retrospective	
	Mean Instances	Mean %	Mean Instances	Mean %	Mean Instances	Mean %
Smarter						
Balanced						
Grade 3	24.66	56.05	13.67	30.32	6.00	13.60
Grade 4	29.00	47.63	21.33	32.78	12.33	19.59
Grade 5	44.00	62.61	19.00	21.90	12.00	15.50
PARCC						
Grade 3	62.00	55.03	37.33	30.97	15.00	14.00
Grade 4	55.67	51.41	35.33	34.87	13.67	13.72
Grade 5	93.67	61.14	33.33	22.50	23.33	16.35
SAGE						
Grade 3	37.00	54.07	21.33	32.46	9.67	14.25
Grade 4	20.00	48.55	12.67	35.51	5.67	15.93
Grade 5	53.33	62.10	22.33	27.04	9.67	10.86
Journeys						
Grade 3	107.33	62.56	44.67	23.99	23.00	13.45
Grade 4	109.00	66.52	33.33	22.81	16.00	10.68
Grade 5	92.67	54.29	51.67	31.52	25.00	14.19

the opposite was true for assessment materials. However, two findings deserve more discussion. First, the passages we evaluated, both instructional and assessment, were written at higher readability levels than the grade levels for which they were intended. Second, the passages we evaluated included all three inference types. Practical implications are apparent and are discussed.

Difficulty Levels of Passages

In almost all cases, the readability levels of passages in this study were judged to be written at higher levels than their associated grade levels. This finding is in alignment with the stance of the Common Core State Standards in regard to text complexity.

Since 2010, the Common Core has expected educators from second grade through high school to use increasingly challenging texts for instruction. This expectation has been set so that students at the end of the high school years may be able to read texts that are as difficult and challenging as those they are expected to read in college (CCSS, 2010; Council of Chief State School Officers and National Governors Association, n.d.).

Some (Hiebert, 2013; Shanahan, Fisher, & Frey, 2012) have advocated the use of challenging text with students in instructional settings. They recognize that these challenging texts should be used only with adequate teacher support. This scaffolding can take the form of close reading (Fisher & Frey, 2012), dyad reading (Young, Rasinski, & Mohr, 2015), and small group reading instruction techniques. These supports are examples of different ways that teachers can organize instruction to provide assistance when students encounter texts above their instructional levels.

By contrast, others (Allington, McCuiston, & Billen, 2015; Gamson, Lu, & Eckert, 2013) have cautioned against the use of frustration-level text for instruction. They argue that text used for instruction should be written at students' instructional reading levels. They reason that when children read texts that are too difficult for them, they struggle to read the words and comprehend what they are reading. They also point out that when children read instruction-level text they can develop self-regulating

skills and increase their reading vocabulary to a much greater degree than when confronting difficult text, even with teacher support.

Regardless of one's view on this issue, results from this study show that publishers of instructional materials and standards-based assessments seem to be following the recommendations of the Common Core by providing challenging texts that are beyond the reading levels previously used. Although teachers may be able to offer supports to students in instructional settings, the assistance readers may need to navigate text that is beyond their instructional levels in testing settings is not permitted. In assessment situations, students may not be able to demonstrate their proficiency with text that is written above their instructional reading levels. Thus, results of such tests may not provide a true measure of their reading ability.

Types of Inferences

Despite the widespread understanding in education that inference is a specific term with one generally agreed-upon definition, this study demonstrated that three distinct inference types were present in both instructional and assessment passages. Anaphoric, prior knowledge, and retrospective inferences were all present in the passages we examined that were written for children.

Anaphoric relationships were the most prevalent type of inference in the texts we examined in this study. This finding is not surprising given the constant use of anaphora in everyday oral communication and by authors in written communication. Writers frequently use anaphora for simple linguistic purposes, as well as for stylistic reasons. Because of the common usage of anaphoric inferences by authors, many teachers may not consciously or systematically teach this type of inference. Teachers may not consciously consider anaphoric relationships as examples of inferences. They may assume children may make anaphoric inferences automatically. Another reason teachers may not teach children explicitly about anaphoric inferences is because core literacy programs may not provide instructional guidelines to do so. Whatever the reason, the prevalence of

instances requiring students to make this type of inference while reading testing and instructional materials seems to merit some explicit instruction, especially since some anaphoric inferences can be quite difficult for children to resolve.

As expected, prior knowledge and retrospective inferences did not occur as frequently in inspected passages as anaphoric inferences. Learning to resolve these less-common inferences often requires instruction that may not be occurring systematically in school. Nonfiction writing often deals with topics with which readers are less familiar. Instruction in the use of prior knowledge to resolve some inferences is especially important despite the lack of attention given to them in the Common Core.

Practical Implications

All inferences are not created equal. As teachers and teacher educators become aware of differences among inferences, they can focus their instruction on specific inference types. Instruction in anaphoric inferences may be most effective in the context of fictional text, since they appear more frequently there. For prior knowledge inferences, this instruction may be most effective in the context of nonfiction text, since this inference occurred most often in those passages. For retrospective inferences, it appears that either fiction or nonfiction text could be used appropriately, since no major differences in occurrence surfaced between the two genres in this study.

All types of inferences can and should be taught explicitly. “A significant group of children who are proficient in lower-level reading skills but have difficulties in reading comprehension struggle because of failures in inference generation” (Kendeou, 2015, p. 170). The ability to draw inferences is stressed in the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, [CCSSO] 2010), but only generally. The first anchor standard for reading calls for readers to be able to “draw logical inferences” and to “cite specific textual evidence” when supporting conclusions drawn from text (p. 10). Fifth-graders are expected to be able to “compare and contrast two or more characters, settings, or events in a story” (p. 12). Second-graders

should be able to “compare and contrast the most important points presented by two texts on the same topic” (p. 13). These are just a few of many standards that require students to become proficient in using inferences as they read.

Some data from this study highlight the need for research to identify which inferences students can make easily and which ones cause them to stumble. Such research could include ways for teachers to identify student needs and tailor instruction to meet those specific needs.

Beyond the implications for teachers and teacher educators, this research may be helpful to publishers of core literacy programs and assessments. Publishers would do well to recognize different types of inferences and provide teaching suggestions and strategies for each of them using both fiction and nonfiction texts.

Limitations and Future Research

One limitation of this study is its small sample size. Only one instructional program was examined and a limited number of passages from that program were included. Similarly, only six passages from each of the three grade levels in the three assessments were examined ($n = 72$ passages). A larger number of passages from a range of sources and a broader representation of grade levels should be studied to ensure that the results of this study are representative.

A second limitation is the number of inferences studied. Although it is clear that the three inference types we highlighted were found in the examined passages, there may be other types of inferences that should also be examined. For example, we did not include the prediction inference in this study because it was difficult to anticipate if readers would be prompted to make the predictions implied by the author. In fact, it is not clear if readers would make the inferences that were present in the texts of this study. Future work in studying inferences should involve examining how students process the inferences they encounter as they read text. It is one thing to identify what inference types are present in text, but we really need to know how readers resolve a variety of inferences as they read. These differences could be a

function of developmental progress, reading ability of individual readers, language differences in readers, or a host of other possible factors. It is possible that some inference types may be more difficult to resolve than others. It is also likely that some inferences within a specific type are more difficult to understand than others. A future study could examine how readers resolve inferences as they read and could result in the creation of a taxonomy of inferences.

Future work could also involve creating instructional materials designed to help teachers learn how to assist children as they encounter inference demands in texts. Specific suggestions could be created considering grade levels and cultural backgrounds of readers. Much more research is needed as we learn more about the various roles that inferences play in the improvement of children's reading comprehension.

References

- Allington, R. L., McCuiston, K., & Billen, M. (2015). What research has to say about text complexity and learning to read. *The Reading Teacher*, 68(7), 491–501. doi:10.1002/trtr.1280
- ATOS. (2018). *Text complexity, ATOS, and Lexile measures*. Retrieved from www.renaissance.com/products/practice/accelerated-reader-360/atos-and-text-complexity/
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford.
- Bishop, G. (2008). A howling success! *Ranger Rick*, 48(11), 2.
- Blachowicz, C., & Fisher, P. (1996). *Teaching vocabulary in all classrooms*. Upper Saddle River, NJ: Merrill/Prentice Hall.
- Boushey, G., & Moser, J. (2006). *The daily five*. Portland, ME: Stenhouse.
- Cave, H. (1980). Two were left. In J. Hancock & C. Steben (Eds.), *Building reading skills, Level 2*. Evanston, IL: McDougal, Littell.
- Council of Chief State School Officers & National Governors Association. (n. d.). *Supplemental information for Appendix A of the Common Core State Standards for English language arts and literacy: New research on text complexity*. Retrieved from www.corestandards.org/assets/E0813_Appendix_A_New_Research_on_Text_Complexity.pdf
- Fisher, D., & Frey, N. (2012). Close reading in elementary schools. *The Reading Teacher*, 66(3), 179–188. doi:10.1002/TRTR.01117
- Flesch-Kincaid. (2018). *Flesch grade level readability formula*. Retrieved from <http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php>

- Gamson, D. A., Lu, X., & Eckert, S. A. (2013). Challenging the research base of the Common Core State Standards: A historical re-analysis of text complexity. *Educational Researcher*, 42(7), 381–391. doi:10.3102/0013189X13505684
- Graesser, A. C., Li, H., & Feng, S. (2015). Constructing inferences in naturalistic reading contexts. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 290–320). Cambridge, UK: Cambridge University Press.
- Hiebert, E. H. (2010/2011). Beyond single readability measures: Using multiple sources of information in establishing text complexity. *Journal of Education*, 191(2), 33–42. doi:10.1177/002205741119100206
- Hiebert, E. H. (2013). Supporting students' movement up the staircase of text complexity. *The Reading Teacher*, 66(6), 459–468. doi:10.1002/TRTR.1149
- Hiebert, E. H. (2014). *The forgotten reading proficiency: Stamina in silent reading*. Santa Cruz, CA: Text Project.
- Hoffman, J. V., Wilson, M. B., Martinez, R. A., & Sailors, M. (2011). Content analysis: The past, present, and future. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 28–49). New York, NY: Guilford.
- Houghton Mifflin Harcourt. (2009). *Journeys*. Boston, MA: Houghton Mifflin Harcourt.
- Kendeou, P. (2015). A general inference skill. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 160–181). Cambridge, UK: Cambridge University Press.
- Kimmel, E. (1984). *The three princes: A tale from the Middle East*. New York, NY: Holiday House.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Kintsch, W. (2004). The constriction-integration model of text comprehension and its implications for instruction. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1270–1328). Newark, DE: International Reading Association.
- Lazar, G. (1996). Using figurative language to expand students' vocabulary. *ELT Journal*, 50(1), 43–51. doi:10.1093/elt/50.1.43
- Education, McGraw-Hill. (2017). *Wonders*. New York, NY: McGraw-Hill Education.
- McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, 99, 440–466. doi:10.1037/0033-295X.99.3.440
- National Governors Association Center for Best Practices & Council of Chief State School Officers (CCSS). (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.
- Oakhill, J., Cain, K., & McCarthy, D. (2015). Inference processing in children: The contribution of depth and breadth of vocabulary knowledge. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 140–159). Cambridge, UK: Cambridge University Press.
- Partnership for Assessment of Readiness for College and Careers. (2017). *PARCC English language arts/literacy assessment*. Retrieved from <https://parcc.pearson.com/practice-tests/english/>

- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R & D program in reading comprehension*. Santa Monica, CA: RAND.
- Readable.io. (2017). *Measure text readability*. Retrieved from readable.io/text/
- Scott Foresman-Pearson. (2017). *Reading Street*. Upper Saddle River, NJ: Scott Foresman-Pearson.
- Shanahan, T., Fisher, D., & Frey, N. (2012). The challenge of challenging text. *Educational Leadership*, 69(6), 58–62.
- Student Assessment of Growth and Excellence. (2017). *SAGE assessment*. Retrieved from <http://sageportal.org/>
- Smarter Balanced Assessment Consortium. (2017). *Smarter Balanced English language arts/literacy summative assessment*. Retrieved from <https://www.smarterbalanced.org/assessments/>
- Stanovich, K. E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32–71. doi:10.2307/747348
- Stenner, A. J., Koons, H., & Swartz, C. W. (2010). *Text complexity and developing expertise in reading*. Chapel Hill, NC: MetaMetrics.
- Van Allsburg, C. (1993). *The sweetest fig*. Boston, MA: Houghton Mifflin.
- Van den Broek, P., Beker, K., & Oudega, M. (2015). Inference generation in text comprehension: Automatic and strategic processes in the construction of a mental model. In E. J. O'Brien, A. E. Cook, & R. F. Lorch (Eds.), *Inferences during reading* (pp. 94–121). Cambridge, UK: Cambridge University Press.
- Young, C., Rasinski, T., & Mohr, K. A. J. (2015). Read two impress. *The Reading Teacher*, 69(6), 633–636. doi:10.1002/trtr.1391