

# 1

## Description of the *Qualitative Reading Inventory-5*

### General Description of the *Qualitative Reading Inventory-5*

#### The Word Lists

#### The Passages

Pre-Primer, Primer, First Grade,  
and Second Grade Passages

Third Grade through Fifth Grade Passages

Sixth Grade Passages

Upper Middle School Passages

High School Passages

Use of the Passages

Measures of Comprehension

### GENERAL DESCRIPTION OF THE *QUALITATIVE READING INVENTORY-5*

The *Qualitative Reading Inventory-5* (*QRI-5*) is an individually administered informal reading inventory (IRI) designed to provide information about (1) conditions under which students can identify words and comprehend text successfully and (2) conditions that appear to result in unsuccessful word identification or comprehension. The *QRI-5* continues a long history of informal reading inventories, which for forty years have been used to identify subjects' reading levels—*independent*, *instructional*, and *frustration*—and to provide valuable diagnostic information. Like other informal reading inventories, the *QRI-5* provides graded word lists and numerous passages designed to assess the oral and silent reading and listening ability of students from the pre-primer 1 through the high school levels.

The *QRI-5* provides a number of assessment options. Results can be used to estimate students' reading levels, to group students for guided reading sessions, or to choose appropriate books for literacy circles, reading workshops, and independent reading. The *QRI-5* can provide information for designing and evaluating intervention instruction. Finally, the *QRI-5* can be used to document student growth.

The *QRI-5* is not a norm-referenced or standardized instrument. Norm-referenced tests provide comparative data; that is, an individual's score is evaluated in terms of the scores of the norm group. Standardized instruments are administered identically for all students. Users of an informal reading inventory, on the other hand, make their own decisions as to the number and type of passages to administer and which assessment options to use. While the *QRI-5*, like other IRIs, uses traditional percentages to determine *independent*, *instructional*, and *frustration* levels, student scores are interpreted only in regard to the individual and not to any norm group.

## THE WORD LISTS

---

The pre-primer 1 word list has 17 words; all of the other word lists contain 20 words that we have selected from *QRI* passages at the same level of readability. For example, the primer word list contains words from the primer passages. The word lists are designed

1. To assess accuracy of word identification
2. To assess speed and automaticity of word identification
3. To determine a starting point for reading the initial passage

The *QRI-5* includes a new feature. Some words on the pre-primer 1 through first grade lists contain phonetically regular vowel patterns. However, because these words are high-frequency words (as is appropriate at these levels), students may still pronounce them as memorized sight words and be relatively unaware of the sound patterns they contain. Therefore the *QRI-5* provides an additional list of low-frequency words that contain the same phonetically regular phonograms. This additional feature allows the examiner to determine two things. First, does the student actually recognize phonograms as phonetic units? Second, does she or he know the more common or frequent ones? This can suggest a beginning point for phonics instruction.

## THE PASSAGES

---

The passages to be read orally or silently assess the student's ability to read and comprehend different types of text. Passages can also be used to assess a student's listening level.

### Pre-Primer, Primer, First Grade, and Second Grade Passages

At the pre-primer levels, there are six passages of increasing difficulty. Five are narratives and one is expository. Narratives usually involve stories while expository text is written primarily to inform, as in textbooks, newspapers, and much magazine writing (Pearson & Fielding, 1991). Three of the narratives and the expository passage are presented with pictures. The primer, first, second, and third grade readability levels have six passages—four narrative and two expository. At the primer and first grade levels, three narrative passages and one expository passage are presented with pictures. At second grade there are two narratives with pictures. Research suggests that emergent readers depend on picture context for both word identification and passage comprehension. In addition, text with pictures more closely approximates the type of selections presented to beginning readers. However, good readers are not dependent on pictures for word identification (Stanovich, 1991b), and it may be important to ascertain whether word identification differs when pictures are present as opposed to when they are absent. The examiner can assess the effect of pictures on a student's word identification or comprehension by contrasting performance on passages with and without pictures.

Because children often have difficulty making the transition from narration to exposition, we felt that it was important to include expository material at all levels. The inclusion of expository material also makes the *QRI-5* more usable by teachers working with adult beginning readers who might be put off by children's narratives. Examiners who

give passages from pre-primer through second grade will be able to ascertain the reader's relative strengths in recalling and comprehending narrative versus expository material. In addition, we have included enough passages to assess differences between oral and silent reading, which, if they occur, should be more predominant at these early reading levels.

All passages contain concept questions that are designed to measure prior knowledge of three or four major concepts within each passage. While the topics of pre-primer 1 through third grade passages are generally familiar, scores on the concept task will help the examiner determine whether the student possesses knowledge of basic concepts necessary to comprehend the selection. All passages are also accompanied by a short prediction task where, given the title of the passage and the concepts included in the prior knowledge measure, the reader is asked to predict passage content.

### Third Grade through Fifth Grade Passages

The passages for third grade through fifth grade include three narrative and three expository passages at each level. The narratives for fourth through fifth grades are biographies of famous people who vary in their familiarity to students in these grades (e.g., Johnny Appleseed versus Amelia Earhart in fourth grade and Martin Luther King, Jr., versus Patricia McKissack in fifth grade). We chose biographies in order to provide a more controlled assessment of prior knowledge. For example, it is easier to assess prior knowledge of a real person than the content of a fictional narrative. Because both familiar and unfamiliar people are included, *QRI-5* results can be useful in suggesting why students are having trouble in comprehension.

The expository passages are descriptive science and social studies materials on various topics, modeled after or taken from representative textbooks. We include passages that, according to our pilot data, offer a range in familiarity because of research suggesting that familiarity, which is measured by students' prior knowledge, is an important determinant of reading comprehension.

### Sixth Grade Passages

At the sixth grade level, we have included three narrative passages, two social studies passages, and two science passages. The two expository passages are on the same topic—ancient Egypt, including *The Nile River* and *Building Pyramids*. The second of the two passages has been formatted to allow the examiner to engage in the think-aloud process with students, and the examiner copy provides a scoring grid to help the examiner keep track of the type of comments made by the reader. If the examiner does not wish to use think-alouds, a student sheet without think-aloud formatting is provided on the accompanying DVD. Because the second passage focuses on the same general topic as the first, it can be used as a pre–post assessment as long as the student's prior knowledge is considered. It can also be used for other purposes such as contrasting oral and silent reading and assessing note-taking ability.



### Upper Middle School Passages

At the upper middle school level, there are six passages. Two passages representative of middle school literature selections are biographical or autobiographical in nature, two passages represent science content, and two represent social studies content. Like the sixth grade level, the two social studies (*Immigration 1 and 2*) and science passages (*Life Cycle of Stars 1 and 2*) are on the same topic, with the second passage in each category formatted for the use of the think-aloud process. As in the sixth grade level, the second passage can also be used for pre–post assessment of progress or other purposes. All the

passages were taken from published literature, science, and social studies texts, so they are representative of classroom materials. Prior knowledge tasks allow the examiner to determine whether comprehension scores were due, in part, to the subject's level of prior knowledge.

## High School Passages

At the high school level, there are three passages taken from representative literature, social studies, and science texts used at that level. The passages within each content area are sections of a biography in the case of literature or a chapter for social studies and science. Readability formulas for determining text level are not particularly useful at the high school level and different formulas provided us with a wide range of readability levels for a single selection. We reasoned that readability levels mattered less than the content typically chosen for high school textbooks. We have therefore included selections that were especially representative of high school content across several publishers.

All high school passages represent relatively unfamiliar topics: the Vietnam War, World War I, and viruses, and each passage is divided into three sections. The second section can be used for post-testing, for assessing note-taking ability, or for other similar purposes. The third section of each selection is formatted so it can be used for engaging in the think-aloud process.

We do not provide guidelines for scoring word identification on the high school passages. Our rationale is that students should read passages at this level silently. However, if a student demonstrates lack of comprehension after silently reading a high school passage and the examiner wishes to evaluate word identification, we suggest two options. The examiner can administer the high school word list, or the examiner can drop back to an upper middle school passage and ask the student to read it orally.

## Use of the Passages

The passages are designed

1. To determine a student's independent, instructional, or frustration levels for word identification in context (pre-primer 1 through upper middle school)
2. To determine a student's independent, instructional, or frustration levels for comprehension
3. To assess a student's ability to read different types of text: narrative and expository text, text with and without pictures (pre-primer 1 through grade two), and text of varying familiarity
4. To assess a student's ability to comprehend in different modes: oral and silent (pre-primer 1 through upper middle school)
5. To assess a student's ability to use look-backs to locate missing or incorrect information (grade three through high school)
6. To assess the variety and quality of a student's think-alouds (grade six through high school)

## Measures of Comprehension

The *QRI-5* assesses comprehension of all passages in two ways: retelling and questions. In addition, for passages at the third grade through high school levels, the examiner may utilize look-backs to evaluate the quality of comprehension further. In sixth grade through high school levels, the examiner may employ think-alouds as a further assessment of comprehension quality.

The comprehension measures are designed

1. To assess the quality of the reader's unaided recall
2. To assess the reader's understanding of the text when prompted with questions
3. To examine the quality of a student's comprehension during reading

**Retelling.** After reading the selection, the student is asked to retell the passage as though telling it to someone who never heard it before. The student's retelling is scored from a map of important idea units contained in the passage. Retelling for narrative passages is coded according to setting or background, goal, events, and resolution. Coding for expository passages focuses on main ideas and details. What the student is able to retell and the organization of that retelling can provide information about the reader's ability to recall important ideas that are structured in some logical way.

**Questions.** Next, the examiner asks the student two types of questions. Questions with answers stated explicitly in the text are called text-explicit questions. Questions with answers that the subject must infer from information in the text are called text-implicit questions. Answers to text-implicit questions must be tied to information in the story and not simply derived from prior knowledge. Independent, instructional, and frustration levels for comprehension are derived from scores on the question measure.

**Look-Backs.** At the third grade through high school levels, the examiner may use look-backs to assess comprehension further. After scoring the questions, the examiner can ask the student to look back in the text to locate missing information or to correct erroneous answers. Look-backs allow the examiner to differentiate between comprehension during reading and memory after reading. Students may understand a concept while reading but then forget it. As a result, they are unable to answer an explicit or implicit question. Look-backs are particularly informative when the student has read unfamiliar or difficult, concept-dense text. While skilled readers can generally employ skimming or rereading to locate information, less skilled readers are unable to do this successfully.

**Think-Alouds.** At the sixth grade through high school levels, the examiner can ask the student to think aloud while reading. The examiner asks the student to pause at designated segments of text signaled by the word **STOP** and describe what he or she is thinking at that point. A think-aloud scoring grid is provided to help examiners keep track of the type of comments made by the reader. A passage that examiners can use to model the think-aloud procedure is also included.

In summary, the *QRI-5* consists of graded word lists and narrative and expository passages. The word lists and passages, which range from pre-primer to high school levels, vary in familiarity, and prior knowledge is assessed before the student reads each passage. Comprehension is assessed through retelling, questions, look-backs, and think-alouds.

# 2

## A Research Perspective

### Reading Acquisition: A Developmental Perspective

#### Factors Related to Comprehension

- Text Structure
- Prior Knowledge
- Oral and Silent Reading
- Questions and Look-Backs
- Retelling
- Think-Alouds

### Factors Related to Word Identification

- Speed and Automaticity
- Knowledge of Letter–Sound Matching
- Role of Dialect in Coding Miscues

### Factors Related to the Use of an Informal Reading Inventory

- Passage Leveling
- Response to Intervention

Since the first edition of the *QRI* was published, we have continued to read and conduct research on factors affecting word identification and comprehension as well as those related to usage of an IRI. Chapter 2 provides a brief summary of this research base.

## READING ACQUISITION: A DEVELOPMENTAL PERSPECTIVE

---

Developmental descriptions of the reading process have been provided by Adams (1990), Chall (1983), Ehri (1991), Ehri and McCormick (2004), Gough and Juel (1991), Spear-Swerling and Sternberg (1996), and Stahl (2006). These descriptions suggest that the knowledge sources children use to construct meaning from text and the aspects of constructing meaning that give them trouble vary depending on their stage or level of reading acquisition. Children tend to move through several stages as they learn to recognize words and construct meaning. The first level has been termed the logographic stage (Ehri, 1991; Ehri & McCormick, 2004), the stage of visual-cue reading (Spear-Swerling & Sternberg, 1996), or the awareness stage (Stahl, 2006). Children understand that print stands for language and for meaning and they attempt to identify words; however, they do so using visual cues as opposed to sound cues. They may use the shape of a word or letter or the presence of a logo as a cue. For example, they can recognize “McDonald’s” but only in the presence of the familiar golden arches. They might call any word that ends in *y* as “monkey” because the word has a tail at the end, just like a monkey.

Children then move to the next level, called the semialphabetic stage (Ehri, 1991; Ehri & McCormick, 2004), the stage of phonetic-cue word recognition (Spear-Swerling

& Sternberg, 1996), or the accuracy stage (Stahl, 2006). At this point, children recognize that sounds within words are represented by letters. Supported by developing phonemic awareness and emerging knowledge of grapheme–phoneme relationships, they begin to focus on letters and sounds usually in the initial and final positions. However, they have difficulty acquiring meaning directly from print without the aid of context, and they use context to facilitate word identification (Jett-Simpson & Leslie, 1997; Stanovich, 2000; Sulzby & Teale, 1991).

Gradually, children move to a phase of controlled word recognition (Spear-Swerling & Sternberg, 1996), where they make full use of letter–sound relationships but often do so slowly and laboriously. Aware that letters represent sounds in a consistent manner (Beck, 2006), they rely less on context. They become more economical in their word identification efforts and begin to use word chunks (i.e., *ap*, *ick*) to pronounce unfamiliar words (Chard, McDonagh, Lee, & Reece, 2007). They recognize that printed words contain onsets or beginning consonant sounds (*fl* as in “flip”) and rimes or the chunk beginning with the first vowel in a syllable (*un* as in “fun”), and the use of known word patterns to identify unknown words becomes an important development in reading acquisition (Adams, 1990; Stahl, Duffy-Hester, & Dougherty Stahl, 2006).

The next stage has been termed sight word reading (Ehri, 1991) or automatic word recognition (Spear-Swerling & Sternberg, 1996). Children recognize many words accurately and automatically without the need of “sounding out.” When they encounter unfamiliar words, their word recognition is rapid and automatic and context no longer plays a meaningful part. Instead, context becomes a “mechanism in the comprehension process” (Stanovich, 1993–1994, p. 282) and is often used to assign meanings to unknown words.

Spear-Swerling and Sternberg (1996) described the next level as the strategic reading stage. Students have mastered basic word-recognition skills. They can read familiar text with fluency and comprehend it fully and now focus on learning how to apply those same strategies to more difficult text and less familiar genres (Chall, 1983).

This developmental perspective does not imply rigid stages through which children pass at prescribed rates. Nor should it be inferred that at any point in learning to read, children are focused only on one aspect of reading. Stanovich (1980) described the interaction of different sources of information that students have available to them as they read. Beginning readers, having less knowledge of phoneme–grapheme relationships, use picture cues and other contextual elements to aid the word recognition process. A pronounced word may be a semantic fit but bear little resemblance to the actual word with regard to letters and sounds. When reading “Who do I see on the plant?” one child read, “Who do I see on the leaf?” As readers begin to develop knowledge of letter–sound patterns, pronounced words may resemble the actual words with regard to letters and sounds but not fit within the meaning of the text. When reading *Marva Finds a Friend*, one child read “looked” as “locked” and “cried” as “cared,” both mispronunciations that seriously changed the meaning of the text. When readers develop automatic word identification, they are able to focus attention on learning different strategies for comprehending and remembering information from text.

We believe that assessment should observe developmental differences in learning to read. The *QRI-5* includes passages for assessment of the stages of beginning word identification. Pre-primer through third grade passages contain high-frequency words as well as words that contain regular and common letter–sound patterns. The passages focus on familiar topics and some include pictures. Our pilot data indicated that at the

pre-primer and primer levels, texts with pictures were read more accurately by readers at these levels (Leslie & Caldwell, 1995). At the primer through second grade levels, pictures assisted retelling and comprehension, which suggests emerging use of context as an aid to comprehension as opposed to being a factor in word identification. Timed administration of the word lists and the determination of reading rate during passage reading allows for evaluation of a student's move into the stage of sight word reading. The gradual rise in reading rate on emergent through third grade passages suggests the students are taking less time to decode words and/or there are less words that need decoding. All levels in *QRI-5* include both narrative and expository text. Expository selections were taken from or modeled after representative classroom texts, which allows for assessment of a student's move to the stage of strategic reading.

## FACTORS RELATED TO COMPREHENSION

---

### Text Structure

Research has described the structures of narrative (Graesser, Golding, & Long, 1991; Johnson & Mandler, 1980; Stein & Glenn, 1979) and expository text (Meyer & Rice, 1984; Weaver & Kintsch, 1991). Narratives that follow the structure of fables (setting–character–goal/problem–events–resolution) have been found to be easier for children to recall than narratives without the major story components (Brennan, Bridge, & Winograd, 1986; Stein, 1979) and, throughout the elementary school years, narrative text is easier for children to recall and comprehend than expository text (Berkowitz & Taylor, 1981; Leslie & Caldwell, 1989; Leslie & Cooper, 1993; Pearson & Hamm, 2005). Paris, Carpenter, Paris, and Hamilton (2005) suggested that a “genuine predictor of reading comprehension is children's narrative reasoning, the ability to understand the elements and relations in goal-directed narratives” (p. 153).

It is highly likely that readers' familiarity with the structure of narratives is greater than their familiarity with the structures of expository text. Children have probably been read more narrative than expository texts. Primary grade instructional materials are predominantly narratives and the narrative texts with which children have the most experience tend to have a single common structure (Graesser, Golding, & Long, 1991; Graesser & Goodman, 1985; Mulcahy & Samuels, 1987). Another reason that narrative text is easier to comprehend may be related to readers' knowledge of content. Students tend to know more about the topics discussed in narrative writings compared to those usually presented in expository texts. Pearson and Hamm (2005) suggested that the difference between narrative and expository comprehension may actually be a function of assessment. The way in which understanding of the two genres is assessed may create “an artifactual difference,” that is, we may test more central ideas in narratives than in nonfiction (p. 61).

In contrast to children's rather stable knowledge of narrative structure, their knowledge of expository structures is more variable (Englert & Hiebert, 1984; Klingner & Vaughn, 2004). Children may be less familiar with any single structure of expository text because of the variety of expository structures—sequence or time order, listing or description, compare and contrast, cause and effect, and problem and solution (Caldwell, 2008b). Skilled readers recognize and use these patterns to facilitate comprehension and memory (Goldman & Rakestraw, 2000) even though they rarely appear in pure form. While authors do not always clearly signal the pattern and may combine two or more

patterns into a single text segment, skilled readers tend to structure their recall according to the dominant structure (Meyer, 2003).

Different languages and cultures have different structures for telling stories. Although all cultures have stories with episodic structures, the number of episodes may vary (Westby, 2004). Cultural groups also vary in the importance that they give to different components of the story. Several cultures emphasize settings and often omit the event sequence. Matsuyama (1983) reported that 80% of Japanese folktales did not have a goal for the main character. Although many cultures speak Spanish and the structures of their stories differ, one common feature is omission of pronouns when retelling important story segments, under the assumption that the character can be easily inferred from the story (Mahecha, 2003).

When an expected narrative structure is not present, the reader may reorganize the retelling to fit the anticipated structure. Students who encounter a story structure different from what is familiar to them may retell the story according to their cultural format. For example, Japanese children accustomed to folktales that do not include goals for the main character might not include goals in their retelling. Other readers may reduce the number of narrative episodes or combine two episodes into one (Kintsch & Green, 1978). Users of an IRI must be aware that a student's native language can influence recall of English stories and that such retelling is not necessarily inaccurate or poorly organized.

*QRI-5*, like previous editions, offers both narrative and expository texts at all levels. This allows the examiner to contrast comprehension following narrative text with comprehension following expository text. Our experience suggests that students may attain higher instructional levels in narrative text than in expository text.

## Prior Knowledge

It has been shown repeatedly that readers with greater prior knowledge, consistent with the content in the text, recall and comprehend better than those with less prior knowledge (Alvermann, Smith, & Readance, 1985; Lipson, 1983). This finding is true for adults (Anderson, Reynolds, Schallert, & Goetz, 1977) and children (Pearson, Hansen, & Gordon, 1979; Taft & Leslie, 1985). Furthermore, the results can be generalized to poor readers as well as good readers (Stevens, 1980; Taylor, 1979). In fact, research has found that good and poor readers' ability to recall and summarize did not differ significantly if the groups were similar in their levels of knowledge (Recht & Leslie, 1988). The implications of these findings for assessment seem obvious (Johnston, 1984). A student with high knowledge of particular content that is consistent with the information presented in the text will be able to recall more information and answer more questions correctly than on material of the same readability about which the student had less or inconsistent knowledge.

There are many ways to measure prior knowledge (Holmes & Roser, 1987): multiple-choice tests, interviews, oral free associations, written free associations, open-ended questions, and oral or written predictions. Valencia, Stallman, Commeyras, Pearson, and Hartman (1991) concluded that different prior knowledge assessments may measure different aspects of prior knowledge. Free-association tasks have been shown to be significantly correlated with comprehension of narrative (Langer, 1984; Leslie & Caldwell, 1989, 1990; Leslie & Cooper, 1993) and expository text (Hare, 1982; Langer, 1984; Leslie & Caldwell, 1989, 1990; Taft & Leslie, 1985). Research (Leslie & Cooper, 1993) also has suggested that instructions for a free-association task, which ask for precise responses rather than general associations, were more predictive of sixth graders' comprehension

and retelling of narrative text. Also, prediction tasks have been shown to correlate somewhat with comprehension, although the findings are less consistent than those using free-association concepts (Leslie & Cooper, 1993; Valencia & Stallman, 1989).

We have often been asked why we do not include questions that assess vocabulary after each *QRI* passage. While we acknowledge the importance of vocabulary knowledge as a critical component of comprehension (Paris, Carpenter, Paris, & Hamilton, 2005), what information is actually acquired by asking for the definition of a single word following the reading of a passage? If the word was explicitly defined in the passage, a correct answer suggests the reader can comprehend literal elements. If the word was not so defined, it suggests an ability to draw an inference. However, the correct definition of a single word does not offer any indication of the scope of a reader's vocabulary knowledge. We believe that our key concept questions are a more sensitive measure, not of a reader's total vocabulary base, but of vocabulary that is relevant to the comprehension of a specific selection.

The *QRI-5* measures prior knowledge by asking children to answer questions that tap their understanding of key concepts (Langer, 1984). In addition, a prediction task based on the concepts and the title of the selection is provided. Our pilot data suggested that at the first grade level and above, both conceptual knowledge and prediction were significantly correlated with some form of comprehension, be it retelling, answers to questions without look-backs, or answers to questions with look-backs.

## Oral and Silent Reading

Although overall differences in oral and silent reading comprehension may be minimal, some children may comprehend better in one mode than in the other. Results of research examining differences between oral and silent reading comprehension are mixed. The most consistent findings appear to be that poor readers at the fourth grade level and below tend to comprehend better when reading orally (Burge, 1983; Swalm, 1972). Rate differences may also suggest the reader's stage of reading development. Children for whom oral and silent rates are similar may not yet have moved into the stage of automatic word recognition. Generally, fluent readers read faster silently than orally.

In the *QRI-5*, multiple passages at the same level allow for a comparison of comprehension following oral and silent reading with regard to retelling and question answering. Multiple passages also allow for examination of possible rate differences between the two modes.

## Questions and Look-Backs

Question content is broadly categorized as literal or explicit questions and inferential or implicit questions (Applegate, Quinn, & Applegate, 2002; Bowyer-Crane & Snowling, 2005; Caldwell, 2008b; Ciardiello, 1998; Graesser & Person, 1994; Kintsch & Kintsch, 2005). Literal questions focus on what was explicitly stated in the text and usually begin with such words as "who," "what," "where," and "when." There are several variations or levels of inferential/implicit questions (Caldwell, 2008b; Leslie & Caldwell, 2009). Low-level inference questions have relatively obvious answers. These may be stated in the text but in different language than the question stem or they may require connecting text segments that are not signaled by grammatical markers such as "because." Higher-order inferential questions ask students to move beyond the text in order to predict, hypothesize, reconstruct, form opinions, or offer rationales. Applegate, Quinn, and Applegate (2002) also described response items that ask the reader to address the significance or meaning of the passage as a whole. They examined eight published informal reading inventories and differentiated inferential questions into low-level, high-level, and

response items. The percentages for low-level inference questions ranged from 6.6% through 36.6%; high-level inferences ranged from 0.8% to 17.5%. Percentages for inference questions on *QRI-4* were 23.7% for low-level inferences, 17.5% for high level inferences, and 18.4% for response-based inferences. Questions that ask students to engage in higher-order reasoning tend to be indicative of a deeper form of comprehension than that revealed by literal questioning (Graesser, Baggett, & Williams, 1996; Kintsch, 2005; McNamara, Kintsch, Songer, & Kintsch, 1996).

Little research has been done on the quality of questions. Few studies of technical adequacy exist (Hamilton & Shinn, 2003) and these generally take the form of correlations with standardized measures (Fuchs, Fuchs, & Maxwell, 1988). A serious issue with using questions is their passage dependency, that is, whether they can be answered correctly without reading the passage. This generally occurs when students already know much of the text content prior to reading and can infer answers based on extensive prior knowledge.

Another issue concerning question usage is whether students answer the questions from memory or have access to the text (Leslie & Caldwell, 2009). Looking back in the text to answer questions clearly changes the nature of the question-answering task (Johnston, 1984). However, it also differentiates between understanding during reading and memory for what was read and understood (Leslie & Caldwell, 2006). While research suggests that students younger than eighth grade rarely use spontaneous look-backs (Garner & Reis, 1981), Leslie and Caldwell (2001, 2006) found that students with levels above third grade were able to use look-backs effectively; that is, they were able to skim the text, find the location of the answer, and respond with a correct answer that was unavailable to them without looking back. Palincsar, Magnusson, Pesko, and Hamlin (2005) asked fourth graders to return to the text in order to answer questions and described the process as “one of constructing and revising coherent and sensible meaning” (p. 275). If looking back increases comprehension, this suggests that assessments that do not allow look-backs may actually underestimate a student’s level of comprehension.

The *QRI-5* includes the option of asking students to engage in look-backs after answering questions in order to resolve comprehension failures. We recognize that asking students to look back is not the same as spontaneous look-backs but we believe that examining whether students, when prompted, can look back and correct or add to answers provides valuable information for instruction. Our pilot research indicated that students with instructional levels at or above third grade were able to increase their explicit and implicit comprehension scores by looking back in the text. Prior to this level, a focus on word identification frequently resulted in the student rereading the entire text rather than looking back and locating the section that contained the relevant information.

## Retelling

Measuring retelling after reading has taken different forms. Kintsch and Van Dijk (1978) measured recall as the percentage of propositions retold by the reader with a proposition defined as a predicate and one or more objects such as agent, object, or goal. Others have broken sentences into clauses that consist of a main verb (Kendeou & van den Broek, 2005; Magliano, Trabasso, & Graesser, 1999; van den Broek, Lorch, Linderholm, & Gustafson, 2001). Retelling has also been defined as the total number of words retold (Fuchs, Fuchs, & Hamlett, 1989; Roberts, Good, & Corcoran, 2005) and the percentage of words that are exact matches or synonyms of text words (Fuchs, Fuchs, & Maxwell, 1988).

The inclusion of retelling formats into recently published informal reading inventories (Bader, 2002; Burns & Roe, 2002; Leslie & Caldwell, 2006; Stieglitz, 2002) suggests that asking students to recall and retell text is a valuable assessment tool. Fuchs, Fuchs, and Maxwell (1988) found correlations between .76 and .82 for different forms of retelling scores and standardized reading comprehension tests. Leslie and Caldwell (2006) found significant correlations between retelling and comprehension as measured by questions for upper middle school and high school text.

However, retellings do present difficulties as an assessment tool (Leslie & Caldwell, 2009). They may underestimate the comprehension of some children because of linguistic demands required by the retelling process (Francis, Fletcher, Catts, & Tomblin, 2005; Palincsar, Magnussen, Pesko, & Hamlin, 2005). Retellings are difficult and time-consuming to score because the text must be broken into units to which the student's responses are matched. Because narrative and expository text have different structures, the quality of literal recall is often measured using story maps for narratives and text maps for expository selections (Pearson & Hamm, 2005).

In addition to breaking the text into units of some kind, using retelling as an assessment tool demands attention to extra text comments such as inferences. Students' retellings generally involve more than literal recall. They include inferences, personal comments or observations, and unrelated or erroneous remarks. Limiting recall scoring to literal components may provide an incomplete picture of the student's comprehension. In order to address extra text comments, retellings are often scored according to qualitative rubrics that include such components as gist/main idea statements, details/story elements, interpretive ideas, generalizations, retelling coherence, retelling completeness, use of linguistic/language conventions, and inclusion of additional information not in the passage (Brown, Pressley, Van Meter, & Schuder, 1996; Hall, Markham, & Culatta, 2005; Irwin & Mitchell, 1983; Romero, Paris, & Brem, 2005). Few numerical guidelines exist for describing the completeness of recall. Our pilot data indicated that the amount of retelling comments varied across passages and passage types with higher recall evident in narrative passages.

The *QRI-5* provides retelling scoring grids that are based on narrative and expository structure. Narrative retellings are scored according to recall that addresses setting/background, goal, events, and resolutions. Expository retellings are scored according to recall that focuses on main ideas and details. In all passages, space is provided to record nonliteral elements of recall such as inferences and reader comments.

## Think-Alouds

Asking readers to read a selection and think out loud as they do so provides valuable information about the strategies that readers use as they attempt to comprehend text. It offers the opportunity to gather observations about the thinking that occurs during the reading process. Over a decade ago, Pressley and Afflerbach (1995) provided a comprehensive summary of studies that have examined this think-aloud process.

There is evidence that skilled readers and those with higher levels of prior knowledge employ more and varied think-aloud strategies than poorer readers or those struggling with unfamiliar text. Goldman (1997) noted that individuals who used a variety of strategies and attempted to build relationships between text ideas demonstrated better recall. Similarly, Crain-Thoreson, Lippman, and McClendon-Magnuson (1997) reported that successful comprehension was associated with "knowledge-transforming activities" during the think-aloud process (p. 586). Readers who connected the text to their background knowledge, constructed inferences, and integrated information across

the text demonstrated higher comprehension. Similar results were noted by Chou-Hare and Smith (1982), Kavale and Schreiner (1979), Myers, Lytle, Palladino, Devenpeck, and Green (1990), and Zwaan and Brown (1996). In light of this, we have decided to emphasize a think-aloud component as one possible alternative for identifying why a student may demonstrate inadequate comprehension.

Readers of varying ages have engaged in think-alouds: university students (Garner, 1982); high school students (Olshavsky, 1976–1977; Rogers, 1991); middle school students (Bereiter & Bird, 1985; Caldwell & Leslie, 2003/2004; Chou-Hare & Smith, 1982; Kavale & Schreiner, 1979); and primary school children (Coté & Goldman, 1998; Coté, Goldman, & Saul, 1998; Myers, 1988; Myers, Lytle, Palladino, Devenpeck, & Green, 1990). Apparently, with appropriate instructions and modeling, learners of most ages can effectively engage in thinking aloud while reading.

Researchers use a variety of coding systems to classify think-alouds (Bereiter & Bird, 1985; Coté, Goldman, & Saul, 1998; Crain-Thoreson, Lippman, & McClendon-Magnuson, 1997; Myers, Lytle, Palladino, Devenpeck, & Green, 1990; Olshavsky, 1976–1977). Many of these systems are quite similar, varying only in the language used to describe the think-aloud comment. For our pilot study, we devised a coding system based on those comments identified most often in the literature.

Is there a link between the assessment and identification of think-aloud strategies and reading comprehension instruction? Instruction in many classrooms actively teaches readers to engage in many of the strategies identified in the think-aloud literature, such as predicting, self-questioning, summarizing, and checking understanding (Nist & Kirby, 1986; Pressley & Afflerbach, 1995).

Traditional assessment measures often have predictive validity; that is, good performance on these measures tends to predict average or above-average classroom performance. However, such measures do not assess process and offer no suggestions for increasing learning. Think-aloud data may be a process measure that can suggest instructional directions. For this reason, we have included think-alouds in the *QRI-5*. In our pilot, we carefully modeled each of the think-aloud strategies that we intended to score before asking the students to think aloud on their own. A recent study (Caldwell & Leslie, in press) found that, on the sixth grade and upper middle school level passages, thinking aloud increased the proportion of text-based associative inferences in retelling. However, the more text-based associative inferences made in retelling, the *lower* the student's scores on comprehension questions. This surprising finding indicates that the associative inferences, which tend only to embellish the text, may interrupt students' processing of complex text and therefore, the text base necessary for answering our comprehension questions is impoverished.

## FACTORS RELATED TO WORD IDENTIFICATION

---

### Speed and Automaticity

As readers practice the accurate identification of words, they begin to read these words with less conscious attention to letter–sound matching and therefore more rapidly. Although automaticity and speed in word identification are different constructs (Stanovich, Cunningham, & West, 1981), reading rate can suggest the presence of automaticity.

Perfetti's (1985, 1988) verbal efficiency theory hypothesizes that children who do not develop the ability to read words accurately and quickly will encounter difficulty

in comprehension. Because most of their attention is directed toward identifying individual words, they are unable to access word meanings efficiently and integrate sentence meanings across an entire passage. LaBerge and Samuels (1985) also stressed the importance of fast and accurate automatic word identification. Both word identification and comprehension require attention on the part of the reader. Because attentional resources are limited, if more attention is directed to word identification, less is available for comprehension.

The *QRI-5* measures word-identification speed (and automaticity) in two ways. The timed portions of the word lists provide one measure. Our pilot data showed that the number of words read within one second predicts reading rate in context better than the total number of words read correctly.

Some may argue that the use of word lists for assessment is not an authentic task. We agree. Readers do not curl up with a good list of words. However, the ability to identify words accurately and quickly out of context is a characteristic of the skilled reader (Perfetti, 1985, 1988; Stahl, 2006; Stanovich, 1980, 1991b). Therefore, using word lists may provide an important piece of assessment information. Readers who take more than one second to identify a word accurately may not have achieved automaticity for that word.

The *QRI-5* provides another measure of word-identification speed: rate of reading as measured in words per minute or correct words per minute on the passages. Reading rate is affected by a variety of factors (Carver, 1990). Rate varies with the purpose of the reader and the process chosen. Text difficulty, as determined by the structure of the text, the familiarity of the content, and the difficulty level of the author's vocabulary, can be important determiners of reading rate. Reading rate can also vary according to the mode of reading (oral versus silent), the age of the reader, and reading skill. Finally, reading rate is also determined by individual cognitive processing speeds. The complexity of variables affecting rate suggests that hard-and-fast norms may be impossible to formulate. In Section 9, we offer some general suggestions for evaluating reading rate based on our pilot data.

One might argue that rate of reading as measured in words per minute (WPM) is primarily an indicator of speed. Another measure of rate that is an index of both accuracy and speed is correct words per minute (CWPM) (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Kame'enui & Simmons, 2001). The concept of correct words per minute (often termed oral reading fluency) has its origin in the field of curriculum-based measurement (Deno, 1985; Fuchs & Fuchs, 1992, 1999; Fuchs, Fuchs, & Maxwell, 1988; Good, Simmons, & Kame'enui, 2001). Curriculum-based measurement is the process of using grade-level classroom materials to assess students at regular intervals throughout the year. The number of CWPM has been described as a valid and reliable measure of student progress (Fuchs & Fuchs, 1999; Fuchs, Fuchs, & Maxwell, 1988) and as an indicator of general reading competence (Kame'enui & Simmons, 2001). Good, Simmons, and Kame'enui (2001) and Hasbrouck and Tindal (1992, 2006) have determined classroom norms for correct words per minute.

Fluency has become an important and popular concept. Many articles and even books are devoted to explanations of how to assess and develop it and the majority of early reading assessments contain some measure of oral reading fluency (Caldwell, 2008a; Caldwell & Leslie, 2009; Johns & Berglund, 2002; Kuhn, Schwanenflugel, Morris, Morrow, Woo, Meisinger, Sevcik, Bradley, & Stahl, 2006; Morrow, Kuhn, & Schwanenflugel, 2006/2007; Pikulski & Chard, 2005; Rasinski, Padek, McKeon, Wilfong, Friedauer, & Heim, 2005; Samuels, 2007; Southwest Educational Development Laboratory, 2008; Walczyk & Griffith-Ross, 2007). Why is fluency so popular? Fluency

is highly correlated with comprehension; that is, fluency rates tend to predict comprehension at least on standardized assessment measures. Unfortunately, some have erroneously interpreted prediction as causality and have assumed that assessing and developing fluency will automatically lead to increased comprehension. (Certainly assessing and practicing fluency is an easier task than designing and delivering effective comprehension instruction.) However, increase of fluency may or may not lead to similar increases in comprehension. Although fluency suggests that the student has cognitive capacity for paying attention to comprehension, it does not indicate how or if that capacity is used (Caldwell, 2008b). Paris, Carpenter, Paris, and Hamilton (2005) cautioned that “fast and accurate word identification does not always lead to high levels of comprehension, and neither does slow, less accurate word recognition necessarily imply poor comprehension” (p. 136). The disappointing results of Reading First grants that tended to emphasize fluency assessment and practice suggests that stressing fluency above or in lieu of comprehension instruction may be a mistake (Glenn, 2008).

### Knowledge of Letter–Sound Matching

A common assessment practice in many recent IRIs is to qualitatively evaluate knowledge of letter–sound matching through a process called *miscue analysis*. This was originated by Goodman (1965, 1967), who referred to word pronunciation errors as miscues influenced by three possible cue systems. The graphophonic cue system refers to the relationships between graphemes (letter and letter combinations) and phonemes (units of sound). If a reader pronounces “jump” as “junk,” one can infer that the reader is utilizing sound cues in the initial and medial positions. The syntactic cue system refers to the position of the word within the syntax of the sentence. If the reader reads “Mary sat on his chair” as “Mary sat on a chair,” one can infer that sentence syntax influenced the substitution of an indefinite article for a pronoun that is at variance with the feminine subject of the sentence. Semantic cues are meaning cues obtained from the content of what is being read. For example, if a reader reads the sentence “I received six presents for my birthday” as “I got six presents for my birthday,” one can infer that the reader is using semantic information in saying “got” for “received.” Goodman’s theory was that reader use of context, as exemplified by the semantic and syntactic cue systems, was an important and strong influence in word pronunciation (1965, 1967). He believed that, as readers develop word recognition skill and speed, they use less graphophonic cues. Therefore, miscues that indicate context usage are strengths because they indicate developing expertise on the part of the reader and a focus on meaning. On the other hand, overreliance on letter-sound cues suggests a poor reader or one who may be headed for trouble.

While the majority of IRIs presently on the market offer some form of miscue analysis that focuses on the three cue systems, the viability of extensive reliance on such an analysis has been questioned. First, additional research has called into question Goodman’s theory regarding the role of context in efficient word identification (Nicholson, Lillas, & Rzoska, 1988; Stahl, 2006; Stahl & Hiebert, 2005; Stanovich, 1991b, 2004). Rayner and Pollatsek (1995) believed that the graphophonic cue system takes precedence over the syntactic and semantic ones. Stanovich (2004) explained that when letter–sound knowledge is emerging or deficient, readers compensate by use of context. Word-identification skill does not depend on contextual prediction but rather “the level of word recognition skill determines the extent to which contextual information will be relied on” (p. 466). To put it another way, as readers develop skill in using the graphophonic cue system, they use context less and less to identify words. Finally, context becomes a factor in the comprehension process as opposed to the word-identification process (Stanovich,

1993/1994). In line with this, McKenna and Picard (2006/2007) suggest that meaningful miscues, those that retain meaning but have little similarity to the letters in the text, may not reflect a strength on the part of the reader. They may actually be evidence of inadequate decoding skills as the reader attempts to compensate for weak decoding by use of context. This suggests the need to reevaluate the miscue analysis process and interpretations drawn from it. We believe that miscue analysis can be helpful but only if interpreted in accordance with current research.

Can letter and sound patterns drawn from miscue analysis provide information on a reader's skill needs in the area of decoding? McKenna and Picard (2006/2007) did not believe so, because such needs may be "masked by context" (p. 379). Miscue analysis, as traditionally used, describes miscues made during oral reading of passages. A reader aided by context may be able to offer a pronunciation that is similar to the passage word while unable to identify the same word or word unit in a different situation. To put it another way, the same letter and sound patterns may or may not be an issue if presented in a list devoid of context. This suggests that phonics must be evaluated apart from context.

Users of *QRI-5* can examine oral reading behavior quantitatively and qualitatively. The quantitative criteria used to determine independent, instructional, and frustration levels follow the recommendations of Harris and Sipay (1985) and Betts (1946). Our pilot data suggested that the best predictor of instructional-level comprehension is 95% for Total Acceptability, the measure of accuracy attained when only uncorrected meaning-change miscues are counted. While we include Total Acceptability as an option for those who believe that semantically acceptable miscues should not be counted as errors, we recommend the use of Total Accuracy in determining reading levels, a practice endorsed by McKenna and Picard (2006/2007). We do this because counting all miscues takes less scoring time than deciding whether a miscue did or did not substantively change meaning. In addition, counting all miscues represents a more reliable practice because examiners can vary in their interpretation of what constitutes a meaning-change miscue. For example, while many individuals might not consider the substitution of "a" for "the" as a miscue that changes meaning, others might disagree and distinguish between the indefinite ("a") and definite ("the") articles. In our classes, we have noticed similar disagreements regarding whether meaning is changed by miscues such as the following: "song" for "singing"; "broken" for "old"; "find" for "get"; "drop" for "die"; "special" for "precious"; and "shiny" for "waxy." In Section 9, we offer guidelines for determining whether a miscue did or did not change meaning.

In addition, we encourage qualitative miscue analysis to ascertain how much attention the reader is paying to the graphic elements of the text and to meaning. We caution, however, that miscues focusing on the graphophonemic cue system represent a strength, not a weakness. We also suggest, based on pilot data, that reader self-correction may indicate whether the reader is paying attention to decoding or to overall passage meaning. We examined the self-correction strategies of children reading pre-primer through third grade passages. We distinguished between miscues that changed meaning and were corrected and miscues that did not change meaning but were also corrected. In levels pre-primer through two, there were little differences between the two. Children were as liable to correct a miscue that distorted meaning as they were to correct one that did not. In level three, there was a change; children tended to correct significantly more meaning-change miscues than those that did not change meaning. While our pilot data needs to be validated with larger numbers of children, we do suggest the following. At instructional

levels of pre-primer through grade two, children are still focused on pronouncing words and, as a result, little distinction is made between meaning-change or non-meaning-change correction attempts. However, at the third grade instructional level, developing word-pronunciation skill and increased fluency allow them to focus more on overall passage meaning. Thus they correct more meaning-change miscues. We have revised our miscue analysis worksheet in Section 9 to reflect this alternative interpretation.

At the lower levels (pre-primer through level three), we encourage the examiner to differentiate between words that contain common and regular letter units such as vowel phonograms versus words that contain irregular or uncommon sound patterns. For example, the word “hill” contains a very common vowel pattern, *ill*, found in approximately 26 single-syllable words (Fry, 1998). In *QRI-5*, we have differentiated word list words at the pre-primer through first grade levels into two general kinds: those containing regular sound patterns and those that do not. Why is this important? Children who learn a regular sound pattern in one word can easily transfer this knowledge to other words (Beck, 2006; Cunningham, 2000; Fry, 1998; Gaskins, Ehri, Cress, O’Hara, & Donnelly, 1996; Stahl, Duffy-Hester, & Dougherty Stahl, 2006). The recognition and use of such common patterns can provide a critical element in phonics instruction.

The *QRI-5* provides a list of low-frequency words that contain 18 frequent phonograms (Beck, 2006; Fry, 1998; Gaskins, Downer, & the teachers of the Benchmark School, 1997). These phonograms are present in words on the word lists and in from three to 13 passages at the pre-primer through first grade levels. The words that contain the phonograms in the lists and passages are high-frequency words as is appropriate for those levels. However, Leslie & Calhoun (1995) found that the ability to read a phonogram in a high-frequency word may not indicate that the child has fully analyzed the word into constituent phonemes. To put it another way, the child may recognize the word by sight, not by knowing its spelling pattern. Our list of low-frequency words containing frequent and common phonograms allows the examiner to determine if the child actually knows the phonogram and is approaching economy of word identification that is characteristic of the stage of controlled word recognition.

### Role of Dialect in Coding Miscues

Oral reading accuracy is a determination of how accurately someone reads a specific text, and practitioners must make distinctions between the oral language pronunciation of a word and a reading error (or miscue). Sometimes this is quite difficult. First, humans do not always hear exactly what is said. Also, distinguishing between a miscue and an acceptable pronunciation often depends on whether the listener or teacher is familiar with the dialect patterns of the reader. If not, differences in pronunciation may be mistaken as oral reading errors or miscues.

This becomes seriously problematic when teachers who use Standard American English (SAE) are listening to children who use a nonstandard dialect, such as African American English (AAE), or whose pronunciation reflects the sound system of another language, such as Spanish. Teachers who are not sensitive to variations in the sound patterns of the students’ oral language may tend to score such differences as miscues. Therefore, it is incumbent on all teachers to understand the patterns of a student’s oral language to accurately interpret oral reading pronunciation and discern a student’s miscue patterns in reading. This is not a small task, for teachers typically listen to the meaning of what children say and not the many forms (phonologic, morphologic, and syntactic) in which meaning is communicated.

Much research has focused on the structure and usage of AAE. The context of use determines the form of language and thus, children who speak AAE vary their dialect patterns across language contexts (Thompson, Craig, & Washington, 2004). Preschool and K–5 AAE speakers vary their use of AAE features in free-play contexts versus picture descriptions (Washington, Craig, & Kushmaul, 1998). Additionally, third grade speakers of AAE were found to use AAE dialect patterns in oral language tasks, such as picture description, to a greater degree than when engaged in oral reading or writing tasks. Picture description tasks appear to be a more sensitive method than spontaneous conversation to obtain a representative sample of a student’s oral dialect patterns for comparison of miscues in oral reading (Thompson, Craig, & Washington, 2004).

Craig and Washington (2004b) listed 24 morpho-syntactic examples of AAE that are common in school-aged children. The most frequent seem to be the zero copula and auxiliary form of “to be” (“Where the dog?”) and subject-verb agreement (“Now she need some candy”). However, AAE dialect changes across grade levels (Craig & Washington, 2004b), and as children become more competent readers, use of AAE features decreases during oral reading. This shift has been found to be most dramatic between second and third grade.

Research on child AAE is incomplete in that it has primarily focused on morpho-syntactic characteristics. Less attention has been paid to phonological features, and it is often difficult to distinguish dialect patterns from developing sound production (Craig & Washington, 2004a). In addition, interpreting dialect forms is often considered apart from vocabulary knowledge. For example, does a child mispronounce a word because of dialect or because that specific word is not part of his or her oral vocabulary base? Early research determined that dialect during oral reading was unrelated to reading comprehension; however, Craig and Washington (2004a) questioned acceptance of this conclusion because such research was based on features of adult AAE that are different from the AAE produced by children. Craig and Washington suggest that it may not be the specific types of AAE features produced by the child that are related to literacy but the frequencies with which they are produced. They refer to this as *dialect density* and suggest that there may be a relationship between dialect density and literacy acquisition.

What does this mean for users of *QRI-5*? Certainly, interpretation of miscues during oral reading should reflect the teacher’s awareness of the child’s spoken language whether it be AAE or Spanish. In Section 9 on the DVD we provide some guidelines for doing this.



## FACTORS RELATED TO THE USE OF AN INFORMAL READING INVENTORY

---

### Passage Leveling

What makes one selection more difficult than another? A common method to describe text difficulty is the use of a readability formula. There are a variety of readability formulas. Some are meant to be used on primary or elementary text. Others focus on higher-level text and adult readers. However, all are based on two components. One is word difficulty estimated as the frequency of the word. For example, “matriarch” is much less frequent in our language than “mother,” so a text containing “matriarch” might have a higher readability level. Word length is another aspect of word difficulty. “Received” is a longer word than “got” and its inclusion in a text would increase the readability estimate. Another readability component is sentence complexity, often measured by sentence

length. Thus, “Because she needed sugar, Mary jumped in the car and quickly drove to the store” would increase text readability more than “Mary needed sugar. She jumped in the car. She quickly drove to the store.”

Readability formulas provide a general estimate of difficulty level but many other components contribute to the complexity of a text. Is the text coherent and well written? Does it include pictures? Do headings accurately indicate content? Is it on a familiar topic? Is the topic interesting to the reader? Is the structure of the text clearly signaled? Is the purpose of the text to entertain or to inform? To what extent does the language of the text parallel spoken language or does the language represent a stylized and formal form of writing? Readability formulas do not account for these components and a text leveled as appropriate for the fifth grade level may be very easy for some fifth graders and extremely difficult for others. Readability formulas provide very rough estimates of difficulty and are aptly described by Allington (2001) as “crude.” And yet, he defended readability formulas as “providing a ballpark estimate of text difficulty. Even a ballpark estimate is better than none at all” (p. 48).

In attempting to level passages for multiple editions of the *QRI*, we have subjected each passage to a variety of readability formulas using the technology available at the time. We moved from counting words, syllables, and sentences and using a handheld calculator for applying the formula to a computer program that did all the calculations for us and insured a greater degree of accuracy. Passages were leveled using Readability Calculations Windows version 7.0 (Micro Power and Light Company, 2007) using the following formulas: the Powers-Sumner-Kearl Formula, the Flesch Reading Ease Formula, the Dale-Chall Formula, the Spache Formula, and the Fry Graph. We have also used the Harris-Jacobsen Formula for passages below grade three. At every level, we found wide fluctuations in the grade levels assigned by different formulas. We chose the level agreed on by at least two out of three appropriate formulas and then tested the appropriateness of the level through extensive piloting. Passages were included in the *QRI* if readers at a specific grade level achieved instructional-level scores in word identification or comprehension and if piloting indicated that the passage was more difficult than the next lower level and easier than the next higher level. Thus, third grade passages were more difficult than second grade texts for our pilot sample; similarly, they were easier than the fourth grade selections.

Since the first edition of the *QRI*, additional ways of determining text difficulty have emerged. Beginning texts used in guided reading have been grouped according to such characteristics as length, print size and layout, difficulty of vocabulary and concepts, language structure, genre, text structure, language patterns such as predictability, and support offered by illustrations (Fountas & Pinnell, 2006). Fountas and Pinnell described 16 guided reading levels crossing kindergarten through third grade: nine levels for kindergarten and first grade text, four levels for grade two, and three levels for grade three. Because of many requests from teachers, we have designated guided reading levels for our pre-primer through third grade passages (see page 461 and Section 15 on the DVD). However, we must issue a word of caution. While the guided reading levels assigned to our passages were not piloted and they represent professional judgments on the part of the levelers, we can provide data to show that adjacent passages are easier or harder.

A relatively recent measure, the Lexile scale, has been developed based on the readability measures of word frequency and average sentence length. What makes the lexile unique is that a reader’s ability and text readability are on the same developmental scale (Stenner, Burdick, Sanford, & Burdick, 2006). The Lexile scale is a transformed logit scale where 1 logit equals 180 lexiles (L). Therefore, when we say that a reader has an



L of 500, it means that he or she is predicted to be able to comprehend a text with a 500L with 75% comprehension on a multiple-choice test. The process for determining a reader's lexile is very different from what most of us consider a valid measure of reading comprehension. A text of 125 words of some difficulty is given to a student and comprehension is assessed with one question. This process continues with different segments of 125 words from various texts (there is no indication that these texts are contiguous). In a study of the accuracy of lexiles, Stenner et al. (2006) tested over 1,000 students in grade three, 893 in grade five, and 1,500 in grade eight. Each item writer developed one question on each of 30 text passages, for a total of 90 items. The standard error of measurement (SEM) is reported to be 32L for a short text (i.e., 500 words), but is reduced significantly by a much longer text (e.g., a text of 900 words would have a SEM of 2L).

To determine the lexile of a reader requires the student to read a series of 125-word text segments and answer one question on each. Then, given the lexile of the segments the mean text lexile on which the student can read with 75% accuracy (on a multiple-choice test) is considered to be the student's lexile. The *QRI-5* requested lexile levels from Metametrics Corporation on each of our passages at second grade difficulty and above (see pages 468–469 and Section 15 on the DVD). Second grade was chosen because the lexile scale does not evaluate the role of pictures or predictable text, as well as other graphic features. Therefore, we determined that a lexile is inappropriate for first grade materials. How to interpret these texts is a bit of a challenge. The lexile is an index of the difficulty of a text, but can we say that a student who scores 75% or higher on the most difficult passage has a lexile equal to that of the passage? That, of course, is what people want to conclude, but because we have not obtained student lexiles for our pilot group such a conclusion is highly questionable. Section 15 explains readability issues in more detail.



## Response to Intervention

In 2004 the Individual with Disabilities Education Act (IDEA) was reauthorized to give schools and districts the option of identifying students as learning disabled based on their response to instruction. In the past, identification of a learning disability rested on an observed discrepancy between intelligence as measured by an individually administered IQ test and classroom performance. Now, if a student's performance "is dramatically inferior to that of his peers, the student may be determined to have a learning disability" (Bender & Shores, 2007, p. 7). Why is this important? Schools receive federal money to provide instruction for students with special needs and funding is closely tied to the identification of a learning disability. This new initiative has been termed *Response to Intervention* or *RtI*.

RtI embodies three tiers. The first is classroom instruction. The teacher implements screening measures to identify students who are achieving below the level of their peers. These students then move into the second tier of instruction involving additional assessments to identify specifics of the problem and provision of small-group instruction to address student needs. A student who still fails to respond adequately moves into the third tier of instruction, which is usually individualized in format. At this point, a learning disability is generally identified. It is very probable that an IRI could be used as a screening measure or progress monitoring tool for all three tiers.

At the present time, explanations of student progress in RtI emphasize "short, quick and easy-to-administer probes" (McCook, 2006, p. 14). This does pose a problem for use of an IRI which, unfortunately, is not short, quick, and easy to administer. The most popular RtI assessment is oral reading fluency, the number of words read correctly in

one or two minutes (CWPM). Other RtI probes include the number of sounds and letters recognized and number of words correctly spelled within a one- to three-minute time frame (Wright, 2007). For comprehension assessment the student orally reads a passage for one minute and then retells it with retelling scored as the number of words in the passage offered within a one-minute time frame. A second form of comprehension assessment is a three-minute cloze or maze task where the student supplies missing words in a passage or selects them from three or four choices.

At the present time, the RtI assessment and intervention model primarily centers on early reading performance (Bender & Shores, 2007; Brown-Chidsey & Steege, 2005; McCook, 2006; Wright, 2007) and involves sound-based/phonological measures and reading fluency, which are predictive of future reading performance. However, Gersten and Dimino (2006) stated that while such assessments are not inaccurate, they are incomplete and “only one part of the picture” (p. 104). Other measures also predict future reading performance, such as oral language proficiency, expressive vocabulary, and understanding of narrative structure (Paris, Carpenter, Paris, & Hamilton, 2005). However, these do not necessarily lend themselves to short, quick, and easy-to-administer RtI probes.

RtI is basically quite new and it remains to be seen how various schools and districts will interpret and implement the fairly general federal guidelines. As presently conceptualized, RtI is primarily focused in the early elementary grades; however, there is certainly a chance that the initiative will extend into the upper elementary, middle, and secondary levels, especially given the large numbers of students above grade four who struggle with reading. Perhaps it is at these levels that an IRI would be most effective as a progress monitoring tool and, because an IRI is informal in nature, teachers can adapt the process to fit their needs. In Section 4, we provide descriptions of various ways in which the *QRI-5* can be used to address the increasing accountability issues in our schools.