

CHAPTER 1

Important Issues and Concepts in Reading Assessment

Reading is a key to students' success in school and in life, and reading assessment done well guides our efforts to foster this success. As we strive to understand and use reading assessment, it is important to consider three questions: First, why do we assess reading? All reading assessment should be conducted with the purpose of helping students achieve in reading. Second, what do we assess when we assess reading? Asking this question allows us to focus on reading program goals and outcomes and what we hope for our students as we teach and support their reading development. Third, how, where, and when do we assess reading? This query anticipates the array of reading assessment materials and procedures that are examined throughout this book in individual chapters.

THREE KEY QUESTIONS

Why Do We Assess Reading?

Reading assessment helps us understand the strengths and needs of each of our students: their reading development and reading achievement. Although all reading assessments should share this purpose, the manner in which individual assessments provide information and the manner in which we use the particular assessment information are varied. Consider the different formative and summative purposes for assessment that are demonstrated in the following scenario.

Hannah, a third-grade teacher, uses a reading inventory to gather detailed information about a student's oral reading strategies and skills. The reading inventory provides information for ongoing analysis of student reading. She determines that the student reads with high confidence but also reads through sentence

boundaries. The student does not reread after obvious meaning-changing miscues. The teacher uses this new assessment information to update her understanding and determines that the student needs to concentrate on developing self-awareness in general, and comprehension-monitoring strategies in particular. Hannah uses the assessment information in the next day's instruction, the goal of which is to encourage the student to regularly monitor the meaning-making process that is reading. Using think-alouds, Hannah models the types of questions that good readers ask themselves as they read, including "Why am I reading?" and "Does that make sense?" In this case, the answer to the question of why we assess reading is that it provides detailed and timely information that is used by a talented teacher to shape instruction to the student's needs.

Late in the school year, Hannah administers a statewide high-stakes reading assessment. The test provides information on students' reading strategies and skills. Results of this test may be used for several purposes. The mean student scores, derived from multiple-choice and short fill-in items, will be used to determine if the school meets federally mandated levels of adequate yearly progress at reading achievement. The test is considered by some to be a judge of accountability, helping to determine if the teacher, school, and school district are working successfully to help students develop as readers and meet state standards in reading. The results of this test are also reported at the individual student level, and parents receive their child's raw scores and percentile rankings in vocabulary knowledge and literal and inferential comprehension. Thus, test results inform parents of their children's general reading achievement levels.

In each of the preceding scenarios, reading assessment is conducted for specific purposes and specific audiences. One assessment is more direct: The classroom teacher is accomplished at using the reading inventory to understand the nature of a student's reading, how it relates to a model of highly efficient reading, and how it anticipates the instruction and learning that the teacher plans for the student. The process orientation of the reading inventory provides a window into the reading strategies and skills that the student uses, or needs. The reading assessment information is immediate and fleeting, and the teacher knows how to focus on and interpret the information that the reading inventory produces. The teacher's knowledge of the nature of students' self-monitoring of reading is matched with the teacher's ability to use the reading inventory to provide information related to this important instructional goal.

In contrast, the end-of-year test is composed of items that describe the students' vocabulary knowledge and text comprehension. The test focuses on reading comprehension products. The results signal that a certain percentage of students meet state and federal reading benchmarks and communicate to particular audiences that the teaching of reading in the district is going well and that taxpayers' money has been well spent.

Throughout this book, the question of why we assess reading frames our consideration of the diverse purposes for assessing reading. These purposes include determining students' reading development, informing instruction, demonstrating teacher and school accountability, describing a reading program's strengths and weaknesses, motivating and encouraging students, and teaching students how to self-assess. Representative purposes for reading assessment and the audiences that use assessment information are presented in Table 1.1. If you are interested in an accounting of the different assessment audiences and purposes in your school or district, you can use the Reading Assessment Inventory: Audiences and Purposes reproducible form in the Appendix.

The question of why we assess reading is answered in different ways because reading instruction and reading assessment are influenced by the larger society in which students, teachers, administrators, and schools work. Consider that diverse theories and bodies of research inform the successful teaching and learning of reading (Tierney & Pearson, 2024). These theories emanate from domains of knowledge that include cognitive psychology, developmental psychology, linguistics, pedagogy, sociology, anthropology, critical race theory, and critical theory. Each theory may suggest different priorities for reading instruction and reading

TABLE 1.1. **Representative Audiences and Purposes for Reading Assessment**

Audience for assessment	Purpose for assessment
Students	<ul style="list-style-type: none"> • To report on learning and communicate progress • To motivate and encourage • To learn about assessment and how to self-assess • To build independence in reading
Teachers	<ul style="list-style-type: none"> • To determine the nature of student learning • To inform instruction • To evaluate students and construct grades • To diagnose students' strengths and weaknesses in reading
School administrators	<ul style="list-style-type: none"> • To determine reading program effectiveness • To prove school and teacher accountability • To determine resource allocation • To support teachers' professional development
Parents	<ul style="list-style-type: none"> • To be informed about children's achievements • To help connect home efforts with school efforts to support children's reading development
Politicians	<ul style="list-style-type: none"> • To establish accountability of schools • To inform the public of school progress
Taxpayers	<ul style="list-style-type: none"> • To demonstrate that tax dollars are well spent

assessment, which will signal different purposes for doing reading assessment. For example, research on reading strategies describes their importance for constructing meaning, whereas research in motivation provides evidence of the need to engage students as they develop into independent, committed readers (Afflerbach et al., 2008; Guthrie & Klauda, 2016). A successful reading program has varied, important outcomes that should include students' growth in the ability to use reading strategies and skills as well as students' increased motivation to read. Reading assessment must have strong connections to these outcomes and describe them well.

The assessment of reading takes place in a context that is influenced by social and political forces. There are assessment practices that may be favored politically, practiced locally, widely supported, or widely questioned. Legislators, taxpayers, parents, school administrators, teachers, and students may all claim legitimately that part of the question of why we assess reading is answered: to provide useful information. However, "useful information" varies, from the parent seeking assessment information that will help coordinate classroom and home reading efforts, to the legislator seeking districtwide reading assessment information in anticipation of an upcoming vote for school funding, to the administrator interested in documenting reading program effectiveness. In these contexts, each purpose for reading assessment must have the potential result of the betterment of student reading. Ideally, one group's need for particular reading assessment information should not displace another group's need. The goal of improving the teaching and learning of reading should help us determine our reading assessment priorities in all cases.

What Do We Assess When We Assess Reading?: The Focus of Assessment

Asking what we assess helps us focus on the goals of the reading instruction program and the relationship of reading instruction to reading assessment. The answer may demonstrate that our conceptualization of reading achievement, as reflected in the reading assessments used, is broad or narrow. The answer may help us determine whether the diverse goals of reading instruction are adequately reflected in the assessments that are intended to measure progress toward those goals. Or, the answer may indicate that whereas school district standards and the curriculum conceptualize reading development broadly, reading assessment measures it narrowly. We should plan to assess what we plan to teach.

Effective instruction contributes to the development of students' reading strategies and skills, motivation, and commitment to reading. Effective instruction broadens students' conceptualization of reading as contributing to success in life. Given the characteristics of successful readers, the array of reasons for reading, and the diverse outcomes of successful reading instruction, should we expect reading assessment to be similarly broad? Does our assessment describe the many

beneficial outcomes of becoming a better reader? How are the outcomes of reading instruction weighted in relation to the assessment that is conducted in states, districts, schools, and classrooms? An examination of popular reading assessments reveals that there are clear gaps between the rhetoric of why reading is important and what is assessed.

Most reading assessments focus narrowly on one set of important reading outcomes: the cognitive strategies and skills of reading. We are familiar with these outcomes because of our experiences with them in school as teachers and former students. Phonemic awareness, phonics, sight word recognition, and fluency (National Institute of Child Health and Human Development, 2000), as well as vocabulary knowledge (Stahl & Bravo, 2010) and literal and inferential comprehension (Snow, 2002), contribute to reading success. Although these are essential elements of successful student reading, they do not fully represent the growth and development that students experience in exemplary reading programs (Afflerbach, 2022). Related, few reading assessments measure student motivation to read or the range of students' social uses of reading. Many reading assessments sample a small portion of student accomplishment and growth—and by implication, teacher and school success. Bracey (2001) notes that standardized tests regularly miss the following outcomes of effective teaching and student learning: “creativity, critical thinking, resilience, motivation, persistence, humor, reliability, enthusiasm, civic-mindedness, self-awareness, self-discipline, empathy, leadership, and compassion” (p. 158). If we want reading assessment to mirror students' accomplishments, we must avoid reading assessment practice that provides, at best, only a partial reflection of those accomplishments.

The question of what we assess when we assess reading must be asked because it can help us become better at assessment. This question helps us prioritize our reading instruction goals and focus on the most appropriate assessment materials and procedures. Schools use an array of assessments conducted across the school year, from reading inventories at the beginning of the year to standardized, norm-referenced tests at the end of the year. An accounting is necessary to optimize this variety of assessments that are intended to serve different audiences and purposes. Our reading assessments include those mandated by the district, the state, and the federal government and those selected by teachers and administrators in schools. An assessment inventory can help us better understand the relationship between the things that a school community values in relation to students' reading development and what is actually assessed. A sample reading assessment inventory, which may be used to investigate the variety, breadth, and focus of assessment, is presented in Figure 1.1. A reading assessment inventory allows us to rank assessment in terms of the match between our teaching and learning priorities and time demands. An assessment inventory helps us compare what is with what could be. This information may be used to create an action plan with the goal of achieving

Assessment is a measure of students' . . .						
Assessment type	<i>Cognitive reading strategies and skills</i>	<i>Motivation for reading</i>	<i>Social uses of reading</i>	<i>Independence in reading</i>	<i>Using reading in collaborative learning environments</i>	<i>Choosing reading over attractive alternatives</i>
Tests and quizzes	X					
Portfolios	X	X	X	X	X	X
Performance assessments	X		X	X		
Teacher questions	X	X	X		X	X
Reading inventories	X					
Teacher observations	X	X	X	X	X	X

FIGURE 1.1. A sample reading assessment inventory. X = demonstrated ability of a particular type of reading assessment to serve the indicated purpose.

better alignment among valued and agreed-upon outcomes of reading instruction, what is taught, and what is assessed. If you are interested in an accounting of the different assessments you use and the focus of these assessments in your school or district, you can use the Reading Assessment Inventory: What Is Assessed? reproducible form in the Appendix.

How, Where, and When Do We Assess Reading?

The determination of why we assess and what we assess must be followed by informed decisions of how best to examine and evaluate students' reading development. Indeed, the majority of this book addresses the different means for assessing students' reading. Part and parcel of a description of how to assess reading is the determination of where and when such assessment should occur. This is where the logical relationship between why we assess, what we assess, and how we assess should be evident. If we assess students' reading comprehension strategies and skills to determine the general success of a districtwide reading program, standardized and norm-referenced tests may be the first choice of school administrators and

other educational decision makers. In contrast, if we assess students' progress to gauge the effectiveness of daily reading lessons, our assessment must be sensitive to the detailed goals of the lessons, and the information provided by the assessment must be immediately useful. Here, we could focus on questions about the contents of the chapter being read, with students' responses providing formative assessment information.

Just as reading assessment should be matched to particular purposes and audiences, how we assess students' reading achievement must be informed by the nature of the reading we expect of them. Much is known about the complexities of reading and the manner in which student readers develop (Kim et al., 2021; Tierney & Pearson, 2024; van den Broek et al., 2016; VanSledright, 2010). Reading is described in detail, and research reminds us that reading is a stunning human accomplishment. Although we are far from any claim that we know all we need to know about reading, what we know should be reflected in our assessments. For example, the necessity of learning phonics and comprehension strategies is well documented, and we have many useful approaches to assessing phonics and comprehension.

In some districts, these assessments are regularly called on to tell the story of students' reading achievement. In other districts, the aggregate results of reading assessments may underrepresent how students have developed as readers. Doing our best work with reading assessments demands that we understand the available assessment materials and procedures and that we use them expertly.

DEFINING READING IS CENTRAL TO USEFUL READING ASSESSMENT

Reading assessment must be clearly linked to a definition of what reading is. The accuracy and detail of this definition will figure largely in the process of determining the validity and usefulness of all reading assessments. This definition of reading should consistently serve as a touchstone as we consider different reading assessment materials and procedures. As we proceed through this book, the assessments we examine must map clearly onto a definition of reading. If not, then the definition or the choice of assessment is faulty. The definition of reading should inform the goals of reading instruction and the reading assessment program that is developed for a particular classroom, grade, school, school district, or state. I note that developing a definition and description of reading can be challenging for one individual. Developing a consensus definition—for a district, school, or grade—is more challenging when many stakeholders are involved, as indicated by the ongoing and contentious process of defining “science of reading.” We are not wanting for definitions and descriptions of reading, but it is exceedingly difficult to create a consensus definition with which all teachers, administrators, parents, students,

legislators, and the general public agree. This creates a complex situation in which there is universal agreement about the importance of reading, but not universal agreement on what reading is, how children learn to read, and how reading is best taught.

Our professional knowledge of reading should inform our conceptualization of reading and students' reading development. Thus, I believe that it is imperative that teachers compare their understandings of what reading is with others' definitions. Reading assessment can be narrowly focused, missing aspects of students' development that are keys to lifelong, accomplished reading. Following is an overview of how I construct an understanding of reading, based on research and professional knowledge. I consider reading as defined in the frameworks of two major national and international reading assessments. First, I focus on the definition of reading comprehension that is provided by the framework of the National Assessment of Educational Progress (NAEP):

Reading comprehension is making meaning with text, a complex process shaped by many factors, including readers' abilities to

- engage with texts in print and multimodal forms;
- employ personal resources that include foundational reading skills, language, knowledge, and motivation; and
- extract, construct, integrate, critique, and apply meaning in activities across a range of social and cultural contexts. (National Assessment Governing Board, 2023, p. 10)

The NAEP reading framework portrays reading as a dynamic and goal-oriented process that involves strategies, skills, and prior knowledge, as well as motivation. I note that the inclusion of motivation in the definition of reading comprehension is a watershed moment; it serves as acknowledgment of the essential nature of motivation. I also appreciate that the definition reminds us that students read across different media. Furthermore, the definition includes what readers may do with what they comprehend (i.e., “integrate, critique, and apply meaning”) and notes that reading is situated in social and cultural contexts.

A dynamic, strategic, and goal-oriented conceptualization of reading also serves as a foundation for the Progress in International Reading Literacy Study (PIRLS; Mullis & Martin, 2021), which assesses reading achievement across the globe. The PIRLS definition of reading provides further details on the nature of reading and anticipates the types of reading assessment that are necessary to gauge student growth in reading across the school year: “*Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school*

and everyday life, and for enjoyment” (italics in original; Mullis & Martin, 2021). The further explanation of reading in relation to the PIRLS definition is critical, in my opinion, for built into this definition of reading are reader motives and the subtexts for why we read and what we read.

We can appreciate the reading that the NAEP and PIRLS reading frameworks describe. However, there is one major wrinkle. The NAEP (conducted in the 4th, 8th, and 12th grades) and the PIRLS (conducted with 9- and 10-year-old students) are based on the premise that the majority of students who take the tests are capable of reading. These students can decode language, recognize words, read fluently, understand the concepts represented in text, and comprehend. The challenge, then, is to create a definition of reading that reflects the development that students undergo as they learn to read. We are fortunate to have considerable research information related to the development of reading ability in young children (Clay, 1979; Heath, 1983; Metsala & David, 2016; Rasinski et al., 2021; Slesman et al., 2022). We also have recent research syntheses that describe the importance of language knowledge in helping students learn to read and continue their development as readers (National Institute of Child Health and Human Development, 2000). This research demonstrates that success in reading is attributable, in part, to the development of strategies and skills related to phonemic awareness, phonics, fluency, vocabulary, and comprehension. Within each of these areas are developmental benchmarks and trajectories that are helpful to me as I construct an account of successful student reading and consider reading assessment suitable for different students.

At this point, my understanding of reading includes the ideas that we read to construct meaning and that we use strategies and skills to do so. These strategies and skills are developmental in nature. While cognitive strategies and skills are essential to success in reading, they do not guarantee success. When I compare my evolving definition with known and valued outcomes of teaching readers, I find that there are several missing pieces. These include students' motivation to choose to read, whether it is in the face of what might be attractive alternatives that may (or may not) involve reading (e.g., TikTok, Instagram, sleeping, soccer), and students' motivation to persevere with reading when the going gets tough. Self-efficacy is a powerful influence on student performance (Bandura, 1998), but it is rarely considered in models of reading development. Another missing piece is how reading experience and reading accomplishment make ongoing contributions to a student's personality development and sense of self. Fortunately, there is research that describes these essential elements of reading development. Reading consists of identifiable cognitive components (e.g., word recognition, comprehension) that interact to make reading successful (Alexander, 2005). Reading development and success are influenced by students' metacognition (Xie et al., 2023), motivation (Barber & Klaua, 2020), self-efficacy (Schunk & Bursuck, 2016), and prior

reading experiences (Anderson & Pearson, 1984). These powerful factors are operating not only in individual acts of reading but also on students' reading development. Students who are not motivated, who struggle to be metacognitive, and who lack self-efficacy find that it is difficult to become a better reader (Afflerbach, 2022).

As I contemplate the nature of reading and what is currently assessed, I am reminded of the need to develop assessments that measure the complexity of student achievement. Too often, reading assessment is a thin account of a robust phenomenon (Davis, 1998). For example, we might ask students to determine the main idea of a text that describes the economic concept of opportunity cost. Determining the theme or main idea of a text is an important reading ability, but it can be complemented with assessment that asks students to apply the main idea related to opportunity cost to an economic decision that the student makes (e.g., How should Reg spend his allowance? Should Maryann buy running shoes or a bicycle helmet with the money she received from her parents for her birthday?) to explain the relation of the main idea to other important economic concepts, or to critique the author's stance toward the main idea.

As to my definition of reading, I am confident that it includes a breadth of conceptualization that informs my reading instruction goals and the nature of the reading assessments I will use:

Reading is the act of constructing meaning from text. We use cognitive and metacognitive skills, strategies, and prior knowledge, all of which are developmental in nature, to understand what we read. The act of reading is supported by reader motivation and positive self-efficacy, which are also developmental. We read to help us achieve our goals, within and outside of school.

In summary, the definition of reading that we construct must reflect an accurate understanding of what reading is, for this definition becomes a benchmark for determining the reading assessments in our classrooms. We must assess our assessments to determine that they get into the nooks and crannies of students' strengths and needs, and that they describe students' immediate and long-lasting achievements in reading. We must build, maintain, and revise our understandings of what reading is to make informed decisions about the quality of our reading assessments. Our detailed definition and characterization of reading will help us vet reading assessments.

A MODEL OF READING ASSESSMENT

Just as we need a clear definition of reading to help us determine what we must assess, we also need a clear understanding of how assessment works to develop

effective assessments. Pellegrino et al. (2001) describe three components of useful assessment: cognition, observation, and interpretation. We engage these aspects whenever we assess, although we may not have considered them in such formal terms.

To the degree that we understand what developing students do when they read, we can use this information to specify the things we would like to assess. The cognition component of reading assessment focuses on the strategies and skills that students use to decode, understand words, and construct meaning. Related research contributes to the building of detailed theoretical constructs that reflect successful reading, which in turn informs our instruction. For example, we know that the ability to summarize a text is an important comprehension strategy (Snow, 2002) that is frequently applied by student readers in school reading tasks. Research provides considerable detail on the nature of summarization strategies (Pressley & Afflerbach, 1995), the usefulness of the strategies, and how they are used with traditional and electronic texts (Cho, 2014; Pečjaka & Pirceb, 2018). We know that students must be able to ignore unimportant text information and recognize and synthesize important text information, using processes that determine connections, similarities, and repetitions within the text. Students must be able to clearly synthesize and state the content and the purpose of the text read. This detailed understanding of the phenomenon to be assessed—summarization, in this case—is the first component in Pellegrino et al.’s (2001) model of assessment. Our model of reading tells us that summarization is important and worthy of teaching and assessing.

Our ability to describe the detailed nature of summarization informs the second component of the reading assessment model: observation. Another way to view observation is that it is what we get from the assessment itself, and a related challenge is to design an assessment task that will provide us the opportunity to observe students as they summarize text. The observation component of reading assessment must accurately represent our knowledge related to the domains of reading *and* assessment. As research contributes to our evolving understanding of how reading works, a concurrent evolution is taking place in educational measurement. That is, theories of how to evaluate student progress are informed by research. This research, in turn, informs our conceptualization of how we can assess those things that we deem critical for reading success. We can propose reading assessments that reflect classroom reading practice and require students to use summarization strategies:

- Do we ask students to write and construct a summary as opposed to choosing the correct summary in a multiple-choice format?
- Do our assessments allow for different interpretation and summary of the text, or is there a single, “correct” meaning?

- Do we ask for retellings of text that can be checked against a detailed list of the text's contents?
- Do we ask for summary accounts of several related texts?
- Are summarization assessment tasks related to the types of summarization regularly done in the classroom?

The observation component of the assessment model (Pellegrino et al., 2001) reminds us that an effective assessment reflects our understanding of how students read in relation to a particular task, text, and setting. How can we evoke, record, and evaluate student summaries? What are our options? We may consider constructed-response questions, think-alouds, short fill-in responses, and answers to questions as the means to gather and then observe students' work with summarization. Or we may require students to perform tasks in which success is contingent, in part, on their ability to summarize text. An assessment task worthy of our intent to observe students' summarization strategies and skills is to create a summary statement of the author's argument for addressing global warming.

Interpretation is the third component of Pellegrino et al.'s (2001) assessment model. So far, we have designed an assessment that focuses on students' summarization ability, combining our understandings of reading and assessment. We expect that careful development of the assessment will yield valuable assessment results. Our faith in our understanding of the nature of summarization and in the assessment materials and procedures that we have developed to assess summarization figure largely in the confidence we have in the inferences that we then make about student achievement and ability.

All reading assessment involves interpretation, and assessment is always related to acts of inference (Johnston, 1987). Reading assessment done well allows us to make inferences about students' needs and strengths. This basic process of reading assessment, generalizing to students' reading performances from a sample of their reading, demands that our inferences be accurate and born of high confidence. The importance of accurate inferences from assessment information cannot be understated, given that we assess to help students become better readers. However, we can make inferences only about those things that we sample. An incomplete assessment agenda, including one that ignores how students develop in terms of motivation and self-efficacy in relation to reading, or their ability to construct meaning from multiple texts on the internet, will limit the inferences that we can make about students' reading development, our teaching effectiveness, and the value of the reading curriculum. To make valid inferences from our summarization assessment, we use a rubric that indicates the degree of student success at summarizing the author's argument for addressing global warming.

The assessment model proposed by Pellegrino et al. (2001) focuses on reading. However, we should also assess related aspects of reading performance and

development. We should conceptualize assessment of reading in addition to—and other than—the cognitive. We know that students' motivations and self-efficacy matter in single acts of reading and across time as student readers develop. We must construct assessments with the rigor and attention to detail that are equivalent to those that measure students' cognitive strategy and skill growth. For example, an assessment of students' motivations for reading must demonstrate our understanding of the construct of motivation. This understanding allows us to specify what aspects of motivation are to be assessed, inform the development of the observation instrument, and guide our interpretation of results.

Pellegrino et al.'s (2001) assessment model includes the three elements of cognition, observation, and interpretation. In the previous extended example, we considered each of these components separately, and then in relation to one another. We contemplated what happens when students summarize text, and we considered the nature of an assessment that would capture and describe this important reading ability. Also, we considered the inferences that we can make about students' reading and summarization, based on our understanding of the reading assessment we have planned. We examined the factors involved in successful student performance, beyond cognitive strategies and skills.

DETERMINING THE SUITABILITY OF READING ASSESSMENT: THE CURRV MODEL

Our final consideration for this chapter is the suitability of reading assessments:

- What is the optimal mix of reading assessments that we use across a school year?
- How do we choose one assessment over another, given what are often limited resources?
- On what basis can decisions about the suitability of a reading assessment be made?

The CURRV model (Leipzig & Afflerbach, 2000) encourages us to examine a reading assessment using five criteria: consequences, usefulness, roles/responsibilities, reliability, and validity. This model supports us in determining if a reading assessment is appropriate for measuring and describing our students' learning. The knowledge gained from applying the CURRV model brings us closer to an informed use of reading assessment.

The CURRV model was developed, in part, as a reaction to the historical practice of using only the criteria of reliability and validity to argue for the quality of reading assessments. Reliability and validity are traditional and critical aspects

of assessment. Yet, they are psychometric principles that cannot help us ultimately determine whether an assessment is suitable for particular teachers, students, and reading situations and assessment tasks. The CURRV model retains the criteria of reliability and validity and adds three necessary considerations:

1. What are the consequences of the reading assessment?
2. What is the usefulness of the reading assessment?
3. What are the roles and responsibilities related to effectively using the reading assessment?

The five different components of the model allow us to analyze different reading assessments and make choices and suggestions based on our understandings of the nature, strengths, and weaknesses of these assessments. The model allows us to judge the situational appropriateness of an assessment. A sampling of the questions that the CURRV model allows us to ask of a reading assessment is presented in Figure 1.2. If you are interested in using the CURRV criteria to evaluate the reading assessments in your school or district, you can use the Using the CURRV Model to Evaluate Reading Assessment reproducible form in the Appendix.

Consequences of Assessment

All reading assessments have consequences. If we return to the question of why we assess, we are reminded that a reading assessment must have the primary consequence of helping students continue their development as readers. Yet, not all reading assessments may effect this change. It is essential to consider all the possible consequences of a reading assessment, positive or negative. Students may

- What are the positive consequences of the use of this assessment?
- What are the negative consequences of the use of this assessment?
- What is the usefulness of this assessment to teachers, students, administrators, and others?
- What are the specific roles and responsibilities for the teachers, students, and administrators associated with this assessment?
- What are the reliability issues related to this assessment?
- What are the validity issues related to this assessment?

FIGURE 1.2. The CURRV model's framework (Leipzig & Afflerbach, 2000): Questions to help determine the suitability of a reading assessment. CURRV = consequences, usefulness, roles and responsibilities, reliability, and validity.

experience consistent support in their reading development as a consequence of careful classroom-based assessment. High-quality reading assessment will help them become better readers. Students may feel increased self-esteem when their high-stakes test scores demonstrate learning. Students could become motivated to read as a result of encouraging teacher assessment feedback. In contrast, students may lose class reading time as school resources are allocated to test preparation. Inappropriate assessment will not provide the type of information that best shapes classroom instruction to students' immediate and long-term needs. A teacher's insensitivity to a student's response to a question stifles student engagement. Additionally, a history of low test scores may teach students to avoid reading. Ultimately, their motivation to read suffers.

The positive or negative consequences of different types of reading assessments influence teachers. Reading inventories and careful teacher questioning provide important information with which accomplished teachers practice the art of teaching. These assessments allow teachers to adjust instruction and influence student learning in a dynamic manner. Performance assessments allow a teacher to better understand the depth and breadth of student achievement related to content-area reading and learning. High test scores garner a salary increase for some teachers and may help build parental and community support for teachers and schools. In contrast, inappropriate assessments take valuable class time from the teaching of reading without yielding valuable information. Decisions made in relation to high-stakes test scores may constrict the curriculum: The content of what is taught in reading blocks and the time to teach it both shrink. Assessing our assessments from the perspective of their intended and unintended consequences will help us determine their suitability.

Usefulness of Assessment

A second aspect of the suitability of a reading assessment, closely related to consequences, is the usefulness of the assessment. If the criterion of usefulness were applied to the mix of reading assessments selected and mandated in schools, the assessment landscape might look different. The array of assessments found in classrooms represents a legacy of tradition and habit, insight and oversight. Reading assessments accompany districtwide initiatives and are mandated under federal and state laws. Some are developed by teachers, some are bought off the shelf, and others are inherited from earlier times. There may be no strategy for coordinating reading assessment efforts. Thus, it is important to take stock of available assessments to consider their usefulness.

A useful assessment is one that allows teachers to gather accurate and actionable information about students' reading. As teachers, we need reading assessments that help us address the different audiences for the information garnered.

We need assessments that provide both formative and summative information. We need assessments that focus on the processes and products of student reading. We also need assessments that are sensitive to the breadth and depth of students' accomplishments in reading at different levels of reading achievement.

Criteria for usefulness include how well the assessment describes student achievement, how easily the assessment information is communicated, and how well the assessment works with curricular goals. Using such criteria, we can create a ranking system that informs and allows us to make sometimes difficult decisions about which assessments are first-order and keepers, which are optional, and which we might do well without. Please note that in some of the chapters that follow, I combine our consideration of the consequences and uses for particular assessments as guided by the CURRV framework because they are tightly interwoven.

Roles and Responsibilities Related to Assessment

The third component of the CURRV framework (Leipzig & Afflerbach, 2000) reminds us that reading assessments come with roles and responsibilities. For example, performance assessments offer distinct advantages over many machine-scored, multiple-choice tests because the performance assessments can describe detailed student learning and achievement. Not all parents are aware of this fact, and administrators and teachers may be charged with informing parents about the potential advantages of performance assessments. In addition to the need to communicate these potential advantages to parents, we must become familiar with the important components of different reading assessments. If we adopt a series of performance assessments to measure students' reading and learning in the content areas, then we must be able to use rubrics to score students' complex performances. In addition, we should be able to use rubrics to help students anticipate the nature of the performance expected of them and to provide models for student learning. We should also be prepared to use performance assessments and rubrics to help students develop their self-assessment abilities (Afflerbach, 2002a).

Reliability of Assessment

The reliability of a reading assessment relates to the consistency and precision of the assessment instrument and process (Kerlinger, 1986). When we assess students, we want to make inferences about the students' learning and performances. Reliability theory posits that the information we gather through assessment is composed of two components: (1) the true component, which reflects the student's real reading achievement, and (2) the error component, which signals "noise" and is the component of an assessment result that does not reflect the student's reading achievement. We must be vigilant in recognizing and controlling the error component. If we

recall the model of assessment presented by Pellegrino et al. (2001), we can immediately appreciate the need for high reliability. If our assessments are unreliable, then the inferences we make about our students' learning and achievements may be erroneous, or worthless. We may miss a student's need for developing critical reading strategies, mistakenly teach decoding skills to a student who already has them, or fail to recognize an increase in a student's motivation to read.

We must strive to determine that assessment practices are consistent and focused on important aspects of student learning. The goals of evenhandedness in dealing with different students and of clear and fair communication with our reading assessments are imperative. When we assess student reading, we must have confidence in the reliability of our assessment. Otherwise, assessment is not worth administering and the results are not worth considering.

Validity of Assessment

The fifth component of the CURRV model (Leipzig & Afflerbach, 2000) is validity (Messick, 1989). We want our assessment efforts to matter, and we must ask questions related to validity before we invest valuable time in any assessment. There are several types of validity. For our purposes here, it is important to consider the construct validity and ecological validity of a reading assessment. (Within the chapters, I only discuss the particular forms of validity that pertain to the chapter topics.)

How do we conceptualize reading? The construct of reading represents our best theory of what reading is. If we view reading as a series of strategies and skills, then we likely believe that phonemic awareness and reading comprehension are critical to students' development as readers. We should make every effort to assess students' growth and achievement related to their comprehension and phonemic awareness. If we believe that reading achievement can be influenced by student motivation, then we should use a reading curriculum that addresses student readers' motivation, as well as assessments that help us understand growth in student motivation.

If we believe a student's self-efficacy influences reading development, we should signal that importance with appropriate assessments. When we invest time in ascertaining the links between our assessment, curriculum, standards, and constructs, we may arrive at the determination of construct validity for an assessment without surprise.

An additional consideration is ecological validity, or the degree to which assessment items and tasks reflect what students do when they read in the classroom. Schmuckler (2001) describes ecological validity as a test of "whether or not one can generalize from observed behavior in the laboratory to natural behavior in the world" (p. 419). This description leads us to two questions:

1. Does student work on an assessment generalize to what is normally done in the classroom?
2. Does student work in the classroom generalize to important tasks and accomplishments in the world outside the classroom?

Contrasting a reading inventory conducted while a student reads orally from a self-chosen text and a series of comprehension questions that follow a two-paragraph reading selection on a standardized, norm-referenced reading test helps us consider ecological validity. A talented teacher conducting a reading inventory can, in this instance, gather information from a student reading texts that are part of the school curriculum and that are read in a normal manner, reflecting the classroom routine. The assessment focuses on students' real-time use of strategies and skills. Compare this with the ecological validity of a multiple-choice, machine-scored reading test. There may be a very limited relationship between daily classroom reading instruction routines and students' reading behaviors, except for those classrooms where test preparation is a focus, with test-like reading materials and assessments used regularly. When the reading and reading-related tasks demanded on an assessment vary greatly from the reading and reading-related tasks done regularly in the classroom, we may see considerable challenges to ecological validity.

THE PLAN FOR THIS BOOK

The CURRV framework (Leipzig & Afflerbach, 2000) described in this chapter is used as an organizing principle for many of the chapters in the book. After an examination of assessments for early reading in Chapter 2, Chapters 3 through 6 focus on particular types of reading assessment, including reading inventories, teacher questioning, portfolios, performance assessments, and high-stakes tests. These chapters begin with a brief introduction and historical overview of the particular assessment, followed by a detailed accounting of the characteristics of the assessment. We then examine the consequences, usefulness, roles and responsibilities, reliability, and validity of each type of assessment as it is situated in a particular teacher's classroom. The consideration of these different types of reading assessment is done in relation to the reading development that most students experience across their school careers. Thus, we determine the suitability of an assessment in relation to students' development as readers. Chapters 7 and 8 focus on important issues that are not given the attention they deserve: the accommodation of English learners (ELs) and students with learning disabilities in reading assessment, and how reading development other than cognitive strategy and skill growth may be assessed. Chapter 10 focuses on using reading assessment to help

students learn to self-assess and become truly independent readers, and Chapter 11 examines assessment related to digital reading and critical reading.

The array of reading assessments now available to teachers and students is broad. To this end, I acknowledge that successful reading assessment programs sample and choose from this wide assortment and tailor these assessments to the programs' particular needs. It is not uncommon to encounter individualistic approaches to portfolios and performance assessment, or hybrid assessments that combine positive features of checklists and performance assessment, and issues that surround high-stakes testing. Thus, my chapter-by-chapter approach to understanding and using reading assessments may appear artificial in some respects. Guiding my plan here is the goal of presenting each assessment separately and providing details related to the assessment, while noting the ways in which it might be complemented by other, valuable reading assessments.

Throughout this book, issues are framed in relation to the educational, social, and political factors that exert varied degrees of influence on reading assessment. There are sharp divides in how different stakeholders, from teachers to legislators, conceptualize reading, the teaching of reading, and reading achievement. It follows that there are disparate ideas related to the nature and role of reading assessment. I attempt to represent reading assessment in relation to the frames of reference and agendas that different people bring to the assessment arena, for consideration of any reading assessment divorced from the school and societal contexts in which it is used does not pass the reality test. Furthermore, I attempt to represent the thinking and rationale behind particular reading assessment initiatives and programs. My purpose here is to anchor reading assessment to the classrooms, the schools, and the society in which our students read and are assessed. Each chapter ends with a section called "Enhancing Your Understanding," in which I provide questions and tasks that invite the readers of this book to apply the knowledge they gain from each chapter to their own assessment practices.

Each chapter is followed by a section called "Reading Assessment Snapshot." Each snapshot addresses an important reading assessment issue that pertains to some or all of the assessments covered in this book. For example, the reading assessment snapshots include examination of the confounds in reading assessment, technology and assessment, and task analyses of our assessments as a check on their suitability.

Throughout this book, I stress the need to examine reading assessment in relation to our current understanding of the reading process, students' development, and the culture of schooling. I reflect on Huey's (1908) observation, made over 100 years ago: "To completely analyze what we do when we read would almost be the acme of a psychologist's achievements, for it would be to describe very many of the most intricate workings of the human mind" (p. 6). When we are successful in our attempts to assess the range of students' development in reading and use this

information to help our student readers thrive, we will have accomplished a similarly remarkable, and necessary, feat.

THE PURPOSE AND INTENDED AUDIENCES FOR THIS BOOK

My purpose is to help us understand and use reading assessments. Through reading this book, readers will become familiar with different types of reading assessments, and together we examine important issues in reading assessment. Never has there been a more promising time for the implementation and productive use of assessments to help us understand students' growth in reading. It is my hope that this book will help readers become familiar with the characteristics of different types of reading assessments and become accomplished in the assessment of reading. Herein, we consider the suitability of different reading assessments for particular purposes and audiences. In doing so, we examine the means for developing, conducting, and using reading assessments to help foster students' reading achievements.

This book is intended for those who are interested in developing a more detailed understanding of different reading assessments, their characteristics, their usefulness and possible consequences, and their requirements. As such, this book can be used in undergraduate and graduate teacher-preparation courses that focus on reading assessment. This book may also be useful in graduate courses that include a comprehensive overview of reading assessment materials and procedures. Finally, this book is inspired by and intended for K–12 teachers and the administrators that support them.

SUMMARY

Using assessments well demands our knowledge and vigilance. The informed use of reading assessments may be accomplished when we attend to the issues discussed in this chapter. First, we must regularly ask the following questions:

- Why do we assess reading?
- What do we assess?
- How do we assess reading?

Second, all of our work in reading assessment must be guided by a detailed understanding and definition of what reading is and a clear conceptualization of reading assessment. We are fortunate that our evolving understanding of reading parallels an evolving understanding of how to best assess reading. A model of

reading assessment provides useful guidelines for us to assess assessments. Finally, we must examine the suitability of a reading assessment. The psychometric standards of reliability and validity are central to any successful reading assessment. Yet, these aspects of assessment must share the stage with our consideration of the consequences, usefulness, and associated roles and responsibilities of particular reading assessments. Equipped with these important understandings of assessment, we are now ready to begin our consideration of the different types of reading assessments.

Enhancing Your Understanding

1. An important use of reading assessment is the inferences we make about students' reading from assessment information. Chart an inference that you make from a reading assessment about a student's reading development. Where does the information that you use to make the inference come from? What degree of confidence do you have in the inference? How could you gather complementary information about the student and the inference? How does the inference help you plan and deliver instruction?
2. Assess your assessments. Are there assessments that provide information about student reading that is not otherwise obtainable? Are there assessments that you cannot live without? Are there assessments that are not worth the time and effort put into them in relation to the quality or type of information they provide?
3. Talk with your students about a particular assessment. Do they understand what it is? Do they understand how it works and why it is valuable?
4. Working with administrators and teachers, develop an assessment inventory. Identify the different types of reading assessments that are used in your classroom, grade, and school. Describe their frequency of use and their usefulness.
5. Use the CURRV model to analyze an assessment that you use or are considering using.

Reading Assessment Snapshot

ASSESSMENT AND EQUITY

Assessment and equity are in a dynamic relationship. We conduct assessment to learn about students' needs and strengths and to then provide suitable instruction. We want all assessment to be fair and reliable—to treat students equitably. We want to use the results of assessment to establish equity of opportunity and to promote achievement in all schools, for all students.

Assessment plays several roles related to educational equity. First, assessment results can be used to identify inequities. For example, reading scores on the NAEP indicate a consistent achievement gap between students from higher and lower socio-economic conditions. The NAEP assessment results can be used to address inequity, as they provide evidence for the claim that economically disadvantaged students should be better supported in schools and communities. We attain educational equity for all students when we provide instruction and support that helps them reach their potential. In doing so, assessment is in service of achieving equity.

In addition to using test scores to evaluate the equity of students' opportunity, assessment itself may be equitable or inequitable. For example, a reading comprehension test that privileges prior knowledge of particular groups of students will result in assessment results that are biased and that give less-than-accurate portrayals of students (or groups of students). Teacher bias may be unconsciously triggered by a student's spoken dialect. Students recently arrived from another country may speak little English, hindering our ability to fully understand their language competencies. These examples remind us that we should examine our assessment materials and procedures to determine if bias is present, and then work to eliminate that bias. This examination helps us determine if an assessment positions students to do their best and if the assessment provides accurate information in relation to reading development. Promoting equity in all assessments, be it tests, teachers' questions, or worksheets and quizzes, results in information that best serves the goal of helping every student. When assessment is fair and reliable, bias against (or for) particular students is lessened or removed, and equity is encouraged.