UNDERSTANDING AND USING

READING ASSESSMENT

K–12

3RD EDITION

Peter Afflerbach

Copublished simultaneously by ASCD and the International Literacy Association, 800 Barksdale Road, Newark, DE 19711, USA.

All web links in this book are correct as of the publication date below but may have become inactive or otherwise modified since that time. If you notice a deactivated or changed link, please e-mail books@ascd.org with the words "Link Update" in the subject line. In your message, please specify the web link, the book title, and the page number on which the link appears.

26 25 24 23 22 21 20 19 18   1 2 3 4 5 6 7 8 9 10 11 12

ADVANCE UNCORRECTED COPY - NOT FOR DISTRIBUTION

TOC Treatment TK

IMPORTANT ISSUES AND CONCEPTS IN READING ASSESSMENT

Three Key Questions

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.

Why Do We Assess Reading?

Reading assessment helps us understand the strengths and needs of each of our students. 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 scenarios (noting that all teacher and student names are pseudonyms).

Hannah, a 3rd 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 her ongoing analysis of student reading. Hannah determines that the student reads with high confidence but also reads through sentence boundaries. The student does not reread after obvious meaning-changing miscues. Hannah uses this new assessment information to update her understanding of the student and determines that the student needs to concentrate on developing self-awareness in general, and comprehension-monitoring strategies in particular. She 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 test. 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 in students' reading achievement. The test results are considered by some to be a measure of accountability, helping to determine if teachers, schools, and school districts 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 children'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 (Hannah) 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 by her 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 results of this test focus 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 Figure 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 form "Reading Assessment Inventory: Audiences and Purposes" 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. 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 assessment; these different priorities will signal different purposes for reading assessment. For example, research on reading strategies describes their importance for constructing meaning, whereas research on motivation provides evidence of the need to engage students as they develop into independent,

committed readers (Afflerbach, Pearson, & Paris, 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.

Figure 1.1

Representative Audiences and Purposes for Reading Assessment

Assessment Audience	Assessment Purpose
Students	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	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	To determine reading program effectiveness
	To prove school and teacher accountability
	To determine resource allocation
	To support teachers' professional development
Parents	To be informed about children's achievements
	To help connect home efforts with school efforts to support children's reading development
Politicians	To establish accountability of schools
	To inform the public of school progress
Taxpayers	To demonstrate that tax dollars are well spent

Source: Modified from *Understanding and Using Reading Assessment, K–12* (p. 6), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

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, supported broadly, or questioned widely. Legislators, taxpayers, parents, school administrators, teachers, and students may all legitimately claim that we assess reading in part to provide useful information. However, the meaning of "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 regimen of assessments that is 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 can broaden 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 widely used reading assessments reveals clear gaps between the rhetoric of the important outcomes of learning to read 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 [NICHD], 2000), as well as vocabulary knowledge (Stahl & Bravo, 2010) and literal and inferential comprehension (Snow, 2002), contribute to reading success. Although these are important elements of successful student reading, they do not fully represent the growth and development that students experience in exemplary reading programs (Afflerbach, Cho, Kim, & Clark, 2010). Related, few reading assessments measure changes in or maintenance of 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 1.2. 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 and enact an action plan with the goal of achieving 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 form "Reading Assessment Inventory: What Is Assessed?" in the Appendix.

Figure 1.2
A Sample Reading Assessment Inventory

	Assessment is a measure of students'…					
	Cognitive Reading Strategies and Skills	**Motivation for Reading**	**Social Uses of Reading**	**Independence in Reading**	**Reading in Collaborative Learning Environments**	**Choosing Reading Over Attractive Alternatives**
Tests and quizzes	✓					
Portfolios	✓	✓	✓	✓	✓	✓
Performance assessments	✓		✓	✓	✓	
Teacher questions	✓	✓	✓		✓	✓
Reading inventories	✓					
Teacher observations	✓	✓	✓	✓	✓	✓

Note: ✓ = demonstrated ability of a particular type of reading assessment to serve the indicated purpose.

Source: From *Understanding and Using Reading Assessment, K–12* (p. 9), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

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 with a description of how to assess reading is the determination of where and when such assessment should occur. This is where the logical relationship among 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, then standardized and norm-referenced tests might be the first choice of school administrators and other educational decision makers. Such tests, administered toward the end of the school year, seem well suited to the task. Test scores tell us whether certain cognitive strategies and skills have been learned. In contrast, if we assess students' progress to gauge the effectiveness of daily reading lessons, then 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. We are beneficiaries of considerable knowledge about the complexities of reading and the manner in which accomplished student readers develop (Guthrie, Wigfield, & You, 2013; van den Broek, Mouw, & Kraal, 2016; VanSledright, 2010; Veenman, 2016b). Much of the act of reading has been examined and described in detail, and research reminds us that reading is a stunning and sometimes arduous human accomplishment. Although we are far from any claim that we know all we need to know about reading, our detailed understanding of reading and how it develops should be reflected in our assessments. For example, the value and necessity of learning phonics, or the correspondence between speech sounds and printed letters, is well documented. It follows from this fact that reading assessment should include materials and procedures that allow students to demonstrate their knowledge and use of phonics, where developmentally appropriate. We have many useful approaches to conducting the assessment of students' phonics knowledge. We can

ask students to tell us the sounds that different letters make. We can ask students to circle letters or groups of letters that correspond to speech sounds that they hear. We can provide students with exercises in which they circle the words that begin with the same sound as a printed letter. We can have students identify words contained in the classroom word wall. We can analyze the spelling in students' stories.

In contrast, how we assess other reading strategies and skills is less well developed. We know that reading history, reading science, and reading literature share common strategies (Pressley & Afflerbach, 1995), such as determining important vocabulary and synthesizing information. Yet reading in each of the content areas demands unique reading strategies (Bråten, Strømsø, & Britt, 2009; VanSledright, 2002). Reading assessment does not regularly reflect our most recent knowledge of accomplished reading in content areas. We know that performance assessments of reading in the content areas can provide valuable information related to students' reading achievement, teachers' accountability, and schools' goodness (Baxter & Glaser, 1998). We know that authentic assessments of reading are developed, in part, as a reaction to the limitations of multiple-choice, machine-scored tests of reading (Afflerbach, 2005). Such promising assessments have made varied inroads. 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 Valid and Useful Reading Assessment

Reading assessment must demonstrate clear linkage with a detailed definition of what reading is. The detail and richness of this definition will figure largely in the process of determining the validity and usefulness of all reading assessments. In this section, my intent is twofold: to describe the process with which I construct a definition of *reading,* and to provide a sufficiently detailed definition of *reading*. This description of reading will serve as a reality check for my discussion of various reading assessment materials and procedures. Throughout this book,

these assessments must map clearly onto my definition of *reading.* If not, then my definition or my 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, even for one individual. Developing a consensus definition—for a district, school, or grade—is more challenging when many stakeholders are involved. 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 challenging 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 and fostered.

I construct my definition of *reading* from several sources. These sources include reading research, state and district standards for reading, recent federal law, and frameworks from national and international assessments of reading. My definition also reflects the knowledge that I have developed as a teacher and researcher of reading, as a parent, and as an experienced reader. In forming this definition, I create the necessary means for critiquing and choosing the types and forms of reading assessment that I examine in this book. I encourage readers of this book to take the time to create their own definition of *reading,* for this personalization of the reading assessment context allows us to focus on the valuable aspects of instruction and learning, which then provides us the opportunity to compare our reading assessment plans with those used in our schools' reading programs. Further, this personalization allows us to develop reading assessments that are best suited to the goals we bring to our reading instruction. It also places us in a strong position to advocate for the assessment of aspects of reading that may be missing or underrepresented in an existing school, district, or state reading assessment program.

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 my definition of *reading* and with others' definitions. Reading assessment can be narrowly focused, missing aspects of students' development that are keys to lifelong, accomplished reading.

As mentioned previously, I use several sources of information to build a definition of *reading* that I believe is sufficient to guide my efforts. First, I refer to my own understanding of reading: reading is a dynamic and complex process (Afflerbach, 2016; Afflerbach & Cho, 2009; Pressley & Afflerbach, 1995) that involves strategies, skills, and prior knowledge. Reading is developmental in nature (Alexander, 2005) and consists of identifiable cognitive components (e.g., word recognition, comprehension) that interact to make reading successful. Reading development and success are influenced by motivation (Guthrie & Klauda, 2016), self-efficacy (Schunk & Bursuck, 2016), and prior reading experiences (Anderson & Pearson, 1984). Adding to this developing definition, I next consult *The Literacy Dictionary: The Vocabulary of Reading and Writing* (Harris & Hodges, 1995). Under the entry for *reading*, there are no less than 20 different takes on the meaning of the word, from Plato to the present. As noted in this dictionary, reading has been conceptualized as a "perceptual act" (p. 206), with this understanding more recently augmented by our knowledge that cognition occurs when we read and that readers actively use strategies and skills to construct meaning. In addition, reading is socially situated. We read, using strategies and skills, in relation to our intellectual and social goals.

Next, I consider reading as defined in the frameworks of two major national and international reading assessments. These frameworks have the advantage of not only describing reading but also describing it in relation to reading assessment. The *Reading Framework for the 2015 National Assessment of Educational Progress* (National Assessment Governing Board, 2015) derives from expert consensus on the nature of reading and defines *reading* in the following manner:

> Reading is an active and complex process that involves:
> - Understanding written text.
> - Developing and interpreting meaning.
> - Using meaning as appropriate to type of text, purpose, and situation. (p. iv)

The National Assessment of Educational Progress (NAEP) reading framework portrays reading as a dynamic and goal-oriented process that involves strategies, skills, prior knowledge, and the reader's purpose. I am encouraged when I read this portrayal of reading because the

framework appears to cover important aspects of reading and shares much with what I believe is an accurate representation of reading.

A dynamic, strategic, and goal-oriented conceptualization of reading also serves as a foundation for the Programme for International Student Assessment (PISA), which is used to assess reading achievement in 74 countries (Organisation for Economic Co-operation and Development [OECD], 2014). The PISA framework 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 understanding, using, reflecting on and engaging with written texts, in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society. (OECD, 2014, p. 61)

The further explanation of reading in relation to the PISA framework 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. The definition acknowledges that we read to do things and that we must read critically in a world that is swimming in information.

We can appreciate the definitions of reading that the NAEP and PISA reading frameworks provide. However, there is one major wrinkle. The NAEP (conducted at 4th, 8th, and 12th grades) and the PISA (conducted with 15-year-old students) are based on the premise that the majority of students who take the related tests are capable of reading. The majority of these students can decode language, recognize words, read fluently, understand the concepts represented in text, comprehend, and critically evaluate. 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; Gamse, Jacob, Horst, Boulay, & Unlu, 2008; Heath, 1983; Metsala & David, 2016; Wagner & Torgesen, 1987). We also have recent research syntheses that describe the importance of particular reading and language knowledge in helping students learn to read and continue their development as readers (NICHD, 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.

At this point, my definition of *reading* includes the ideas that we read to construct meaning and that we must use particular strategies and skills to do so (Afflerbach et al., 2008). These strategies and skills are developmental in nature. My experiences as a teacher and as a parent remind me that cognitive strategies and skills are keys to success in reading, but they do not guarantee success. When I compare my evolving definition with what I know about reading and what I value as important outcomes of reading instruction, I find that there are several missing pieces. They 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., Facebook, 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. My hope for all students is that they become lifelong, committed readers, and if reading instruction helps them do so, that is an important accomplishment. I certainly would like part of my reading assessment plan to describe this feat.

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 main idea is an important reading ability, but it can be complemented with an assessment that asks students to apply the main idea 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 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 reader affect. 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 judging 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.

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 so that we can develop informed reading assessment materials and procedures. *Knowing What Students Know: The Science and Design of Educational Assessment* (Pellegrino, Chudowsky, & Glaser, 2001) examines successful assessment. The authors describe three requisite components of useful assessment: cognition, observation, and interpretation. We touch on aspects of this model whenever we assess, although we may not have considered them in such formal terms.

The cognition component of reading assessment focuses on the strategies and skills used by students as they develop as readers. 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. We have research related to particular cognitive aspects of reading: the strategies and skills that students use to decode and understand words and construct meaning. This 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 successful student readers frequently apply in school reading tasks.

Research provides considerable detail on the nature of summarization strategies (Pressley & Afflerbach, 1995), the usefulness of the strategies, and the manner in which they are used with traditional and electronic texts (Berkeley, Mastopieri, & Scruggs, 2011; Cho, 2014). 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. Once a text is summarized, it can be used in the creation of student work or to challenge another text (Luke, 1994). Thus, reading strategy research clearly describes the cognitive details of summarization.

Our ability to describe the detailed nature of summarization and other important aspects of reading should inform the second component of the reading assessment model: observation. 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. For instance, we can propose reading assessments that reflect classroom reading practice and require students to use summarization strategies:

- Do we ask students to 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 a single, "correct" meaning?
- Do our assessments provide an optimal mix of formative and summative assessment?
- Does this mix help us guide students across the grades toward consistent attainment of educational goals, including the Common Core State Standards?
- 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. The observation component also reminds us to create an efficient combination of assessment materials and procedures. 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 the students' ability to summarize text. Our analysis of the text that students read tells us that their summaries should include specific information, along with information that is attributed to individual differences in prior knowledge related to the text topic.

Interpretation is the third component of Pellegrino and colleagues' (2001) assessment model. Assessment is always related to acts of inference (Johnston, 1987). So far, we have designed an assessment that focuses on students' summarization ability, combining our understandings of reading and assessment. One product of this process is the assessment results themselves. Our faith in our understanding of the nature of summarization and our faith in the assessment materials and procedures that we have developed to assess summarization figure largely in the faith we have in the inferences that we then make about student achievement and ability.

All reading assessment involves interpretation. Reading assessment done well allows us to make inferences about students' needs and strengths. We may move from the specific to the general—a student's performance in a particular place and time on an assessment task, and the more broad abilities and talents that are related to the performance. 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 making accurate inferences from valid assessment information cannot be overstated, given that we assess reading to help students become better readers. However, we are able to 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.

The helpful model of assessment proposed by Pellegrino and colleagues (2001) focuses on cognition as the important aspect of learning and instructional outcomes. We can also assess related and important aspects of reading development. We can determine that students' motivations, self-efficacy, and developing sense of the self as a reader are important. We must develop assessments with the rigor and attention to detail that are equivalent to those that measure 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, informs the development of the observation instrument, and guides our interpretation of results.

Pellegrino and colleagues' (2001) assessment model includes the three components 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 and cannot make about students' reading and summarization, based on our understanding of the reading assessment we have planned. Finally, we considered the factors involved in successful student performance beyond cognitive strategies and skills.

Determining the Suitability of Reading Assessment: The CURRV Framework

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 our limited resources?
- On what basis can decisions about the suitability of a reading assessment be made?

The CURRV framework (Leipzig & Afflerbach, 2000) encourages us to examine a reading assessment using five criteria to determine if it is appropriate for measuring and describing our students' learning. This framework supports us in understanding and using reading assessment and allows us to examine the (1) C*onsequences,* (2) U*sefulness,* (3) R*oles and responsibilities,* (4) R*eliability,* and (5) V*alidity* of a reading assessment. The knowledge gained from applying the CURRV framework brings us closer to an informed use of reading assessment.

The CURRV framework 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 do not help us determine whether an assessment is suitable for particular teachers, their students, and specific reading situations. The CURRV framework 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 the effective use of the reading assessment?

The five components of the framework 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 framework allows us to judge the situational appropriateness of an assessment. Figure 1.3 presents a sampling of the questions that the CURRV framework allows us to ask of a reading assessment.

If you are interested in using the CURRV criteria to evaluate the reading assessments in your school or district, you can use the form "Using the CURRV Framework to Evaluate Reading Assessment" 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 of the possible consequences of a

reading assessment, positive or negative. Students may 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 receiving 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 cause students to avoid reading. Ultimately, their motivation to read suffers.

Figure 1.3

The CURRV Framework[a]: Questions to Help Determine the Suitability of a Reading Assessment

CURRV = Consequences, Usefulness, Roles and responsibilities, Reliability, and Validity

- 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?

[a]Leipzig and Afflerbach (2000).

Source: From *Understanding and Using Reading Assessment, K–12* (p. 18), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

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 help teachers to adjust instruction and influence student learning. Performance assessments allow a teacher to better understand the depth and breadth of student achievement related to content area reading and learning. Performance assessment results also help administrators choose instructional programs related to the Common Core State Standards. 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: both the content of what is taught in reading blocks and the time to teach it 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. In schools and districts, there may be no strategy for coordinating reading assessment efforts. Thus, it is important to take stock of available assessments to consider their usefulness. For example, administrators should be familiar with the range of reading assessments that are used throughout a building or throughout the district. Given the frequency of educational initiatives (both local and imposed from afar), it is not surprising to find complicated mixes of assessment from classroom to classroom.

A reading assessment may or may not be broad enough to cover the range of student ability and achievement that exists in our classrooms. A useful assessment would be one that allows teachers to gather accurate and usable 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.

A focus on the usefulness of reading assessment allows us to determine necessary assessments, as well as those that are expendable. Criteria for usefulness, including how well the assessment describes student achievement, how easily the assessment information is communicated, and how well the assessment works with a particular teacher's immediate and long-term goals, can help us rank reading assessments. This usefulness ranking then 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. In some of the chapters that follow, I combine my consideration of the consequences and uses of 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 reminds us that our use of reading assessments must be accompanied by an understanding of the related roles and responsibilities. It is especially important to focus on the roles and responsibilities created by reading assessments when we are considering moving beyond the traditional and familiar forms of reading assessment. For example, teachers' roles and responsibilities related to testing include reminding students to take the tests seriously, familiarizing students with the tests' formats, and following the tests' standardized administration procedures. Students are responsible for adopting a serious approach to the tests and preparing to take them.

All assessments demand attention to specific roles and responsibilities. Teachers must play the role of educator when new assessment materials and procedures are used. 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 comprises two components: (1) the true component, which reflects the reality of the student's reading achievement; and (2) the error component, which signals the part 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 and colleagues (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.

Standardized testing provides a paradigm, of sorts, for reliability. Considerable effort is expended in the development of such tests to reduce the error of measurement that is present in all assessments. The application of reliability theory to assessments other than standardized tests is an important and necessary practice. We must strive to determine that our 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, our assessment information is not worth consulting.

Validity of Assessment

The fifth component of the CURRV framework is validity (Messick, 1989). In addition to the voluminous research and theory related to the validity of assessment, there is the intuitive test: 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 the ecological validity of a reading assessment. (Within the chapters, I discuss only 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 consider a reading curriculum that addresses student readers' motivation. We should also develop assessments that help us understand growth in student motivation, in relation to the construct of reading. If we believe a student's self-efficacy or mindset influences reading development, then we should signal that importance with appropriate assessments. When we invest time in ascertaining the links among our assessment, the curriculum, standards, and constructs, we are likely to find that our assessment has construct validity.

An additional consideration is ecological validity. Neisser (1976) uses the term to refer to the extent that a theory can tell us something about what people do in real and significant situations. An application to assessment requires us to focus on 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 typically done in the classroom?
2. Does student work in the classroom generalize to important tasks and accomplishments in the world outside the classroom?

An example of reading assessments with contrasting ecological validity might include 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. 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. One challenge to ecological validity might be that if a student does not regularly read orally, then the demand of doing so compromises the assessment's validity. 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 the notion of ecological validity.

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 consider the following questions:

- Why do we assess reading?
- What do we assess when we assess reading?
- How, where, and when 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. The CURRV framework of reading assessment

provides useful guidelines for assessing assessments, enabling us to examine the suitability of the reading assessments we use. 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. Making inferences about students' reading is an important use of the information we gain from reading assessment. Chart an inference that you make about a student's reading development using reading assessment. 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 you cannot live without? Are there assessments that provide information about student reading that is not otherwise obtainable? Are there assessments that are not worth the time and effort expended on them, in relation to the quality or type of information they provide? Do your assessments provide a comprehensive portrait of your student readers or an incomplete sketch?

3. Talk with your students about a particular assessment. 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.

Reading Assessment Snapshot

Task Analysis

How can we determine that student work within an assessment is clearly related to the specific reading skills, strategies, and attitudes that we teach and hope to measure? How can we determine that our assessment is clearly focused on the knowledge that students gain from reading? How can we be certain that what we ask students to do in a reading assessment is a legitimate request? Conducting an assessment task analysis can provide confidence that we are getting what we ask for from our assessments.

Effective assessment is aided by the process of piloting (or trying out), the assessment. The pilot phase of an assessment allows us to determine if the assessments we create or choose actually elicit particular student thinking and student work. We can walk through the very assessment tasks that we demand of our students. By taking the assessment ourselves, we become aware of what the assessment task demands of our students. We must also be sensitive to the developmental differences that influence our work on a task and our students' work on the same task.

A common result of a detailed task analysis is a realization of the necessary commitment of time and resources to create worthwhile assessments. We may determine that our "reading" assessments involve student capabilities that are not necessarily specific to reading. If students write to demonstrate what they have understood in a textbook chapter or are called on to verbalize their constructed understanding of a poem, we need to consider how writing and speaking figure in the assessment.

In summary, task analysis is a critical part of assessment development and determining the congruence of an assessment with its ostensible purpose, or verifying that we are getting what we ask for. If you are interested in task analyses related to the reading assessments used in your school or district, you can use the form "Conducting a Task Analysis to Enhance Reading Assessment" in the Appendix.
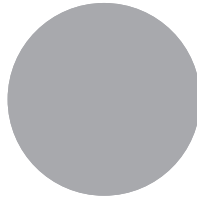
REFERENCES

Aarnoutse, C., & Schellings, G. (2003). Learning reading strategies by triggering reading motivation. *Educational Studies, 29,* 387–409. doi:10.1080/0305569032000159688

Abedi, J. (2012). Validity issues in designing accommodations. In G. Fulcher & F. Davidson (Eds.), *Handbook of language testing* (pp. 48–61). New York: Routledge.

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice, 19,* 16–26. doi:10.1111/j.1745-3992.2000.tb00034.x

Abedi, J., Lord, C., & Plummer, J. R. (1997). *Final report of language background as a variable in NAEP mathematics performance.* (CSE Technical Report No. 429). Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.

Afflerbach, P. (Ed.). (1990). *Issues in statewide reading assessment.* Washington, DC: American Institutes for Research.

Afflerbach, P. (2002a). Teaching reading self-assessment strategies. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 96–111). New York: Guilford.

Afflerbach, P. (2002b). The road to folly and redemption: Perspectives on the legitimacy of high-stakes testing. *Reading Research Quarterly, 37,* 348–360. doi:10.1598/RRQ.37.3.5

Afflerbach, P. (2005). National Reading Conference policy brief: High stakes testing and reading assessment. *Journal of Literacy Research, 37,* 151–162. doi:10.1207/s15548430jlr3702_2

Afflerbach, P. (2016). An overview of individual differences in reading: Research, policy, and practice. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 1–12). New York: Routledge.

Afflerbach, P., & Cho, B. (2009). Identifying and describing constructively responsive comprehension strategies in new and traditional forms of reading. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 69–90). New York: Routledge.

Afflerbach, P., & Cho, B. (2010). The classroom assessment of reading. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 487–514). New York: Routledge.

Afflerbach, P., Cho, B., & Kim, J. (2011). The assessment of higher order thinking in reading. In G. Schraw & D. H. Robinson (Eds.), *Assessment of higher order thinking skills* (pp. 185–218). Charlotte, NC: Information Age.

Afflerbach, P., Cho, B., & Kim, J. (2015). Inaccuracy and reading in multiple text and Internet/hypertext environments. In D. Rapp & J. Braasch (Eds.), *Processing inaccurate information: Theoretical and applied perspectives from cognitive science and the educational sciences* (pp. 403–424). Cambridge, MA: MIT Press.

Afflerbach, P., Cho, B., Kim, J., & Clark, S. (2010). Classroom assessment of literacy. In D. Wyse, R. Andrews, & J. Hoffman (Eds.), *The Routledge international handbook of English, language and literacy teaching* (pp. 401–412). New York: Routledge.

Afflerbach, P., Cho, B.-Y., Kim, J.-Y., Crassas, M., & Doyle, B. (2013). Reading: What else matters besides strategies and skills? *The Reading Teacher, 66*, 440–448.

Afflerbach, P., & Meuwissen, K. (2005). Teaching and learning self-assessment strategies in middle school. In S. E. Israel, C. C. Block, K. L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 141–164). Mahwah, NJ: Erlbaum.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61,* 364–373. doi:10.1598/RT.61.5.1

Afflerbach, P., & VanSledright, B. (2001). Hath! Doth! What? Middle graders reading innovative history text. *Journal of Adolescent & Adult Literacy, 44,* 696–707.

Airasian, P. W. (1991). *Classroom assessment.* New York: McGraw-Hill.

Alexander, P. A. (2005). The path to competence: A lifespan developmental perspective on reading. *Journal of Literacy Research, 37,* 413–436. doi:10.1207/s15548430jlr3704_1

Alexander, P. (2010). Reading into the future: Competence for the 21st century. *Educational Psychologist, 47,* 259–280.

Allington, R. L. (1983). The reading instruction provided readers of differing reading abilities. *The Elementary School Journal, 83,* 548–559. doi:10.1086/461333

Allington, R., & Gabriel, R. (2016). Classroom influences on individual differences. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 196–208). New York: Routledge.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards*

for educational and psychological testing. Washington, DC: American Educational Research Association.

Anderson, L. W., & Bourke, S. F. (2000). *Assessing affective characteristics in the schools* (2nd ed.). Mahwah, NJ: Erlbaum.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (pp. 255–291). White Plains, NY: Longman.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57,* 13–18.

Athey, I. (1985). Reading research in the affective domain. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 527–557). Newark, DE: International Reading Association.

Baker, L., & Brown, A. (1984). Metacognitive skills in reading. In P. Pearson, R. Barr, M. Kamil, & P. Mosenthal (Eds.), *Handbook of reading research* (Vol. 1, pp. 353–394). New York: Longman.

Bandura, A. (1998). Personal and collective efficacy in human adaptation and change. In J. G. Adair, D. Belanger, & K. L. Dion (Eds.), *Advances in psychological science: Vol. 1. Personal, social and cultural aspects* (pp. 51–71). Hove, UK: Psychology Press.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Charlotte, NC: IAP.

Baxter, G. P., & Glaser, R. (1998). Investigating the cognitive complexity of science assessments. *Educational Measurement: Issues and Practice, 17,* 37–45. doi:10.1111/j.1745-3992.1998.tb00627.x

Beck, I. L., McKeown, M. G., Hamilton, R. L., & Kucan, L. (1997). *Questioning the author: An approach for enhancing student engagement with text*. Newark, DE: International Reading Association.

Berkeley, S., Mastopieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities, 44,* 18–32. doi:10.1177/0022219410371677

Berliner, D. C., & Biddle, B. J. (1995). The manufactured crisis: Myths, fraud, and the attack on America's public schools. Cambridge, MA: Perseus.

Betts, E. A. (1946*). Foundations of reading instruction, with emphasis on differentiated guidance*. New York: American.

Bhattacharya, A. (2016). The influence of poverty on individual differences in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 305–317). New York: Routledge.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80,* 139–144.

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain*. New York: David McKay.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology, 54,* 199–231. doi:10.1111/j.1464-0597.2005.00205.x

Boekaerts, M., Pintrich, P., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego, CA: Academic.

Borkowski, J. G., & Turner, L. A. (1990). Transsituational characteristics of metacognition. In W. Schneider & F. E. Weinert (Eds.), *Interactions among aptitudes, strategies and knowledge in cognitive performance* (pp. 159–176). New York: Springer-Verlag.

Bracey, G. W. (2001). The 11th Bracey report on the condition of public education. *Phi Delta Kappan, 83,* 157–168.

Bråten, I., Strømsø, H. I., & Britt, M. A. (2009). Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly, 44,* 6–28. doi:10.1598/RRQ.44.1.1

Bråten, I., Strømsø, H. I., & Ferguson, L. (2016). The role of epistemic beliefs in the comprehension of single and multiple texts. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 67–79). New York: Routledge.

Brown, J. E., & Sanford, A. (2011). *RTI for English language learners: Appropriately using screening and progress monitoring tools to improve instructional outcomes*. Washington, DC: National Center on Response to Intervention, Office of Special Education Programs, U.S. Department of Education.

Brozo, W., & Afflerbach, P. (2010). *Adolescent literacy inventory*. New York: Pearson.

Brualdi, A. (1998). Implementing performance assessment in the classroom. *Practical Assessment, Research & Evaluation, 6*(2). Retrieved from http://pareonline.net/getvn.asp?v=6&n=2

California Department of Education. (1998). *English language arts content standards for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Author. Retrieved from http://www.cde.ca.gov/be/st/ss/documents/elacontentstnds.pdf

California Department of Education. (2016). *Common Core ELA parent overview: Grades 3–8.* Sacramento, CA: Author. Retrieved from http://www.scoe.net/castandards/Documents/parent_overview_ela_3-5.pdf

Calkins, L., Montgomery, K., & Santman, D. (with Falk, B.). (1998). *A teacher's guide to standardized reading tests: Knowledge is power*. Portsmouth, NH: Heinemann.

Cazden, C. B. (1986). Classroom discourse. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 432–462). New York: Macmillan.

Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology, 87,* 154–167. doi:10.1037/0022-0663.87.1.154

Cho, B. (2014). Competent adolescent readers' use of Internet reading strategies: A think-aloud study. *Cognition and Instruction, 32,* 253–289.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review, 24,* 205–249.

Clay, M. M. (1979). *Reading: The patterning of complex behaviour* (2nd ed.). Portsmouth, NH: Heinemann.

Clay, M. M. (2000). *Running records for classroom teachers*. Portsmouth, NH: Heinemann.

Clay, M. M. (2002). *An observation survey of early literacy achievement* (2nd ed.). Portsmouth, NH: Heinemann.

Council for Exceptional Children. (2000). *Making assessment accommodations: A toolkit for educators*. Reston, VA: Author.

Council of Chief State School Officers & National Governors Association. (2010). *Common core standards for English language arts and literacy in history/ social studies, science, and technical subjects*. Retrieved from http://www. corestandards.org/wp-content/uploads/ELA_Standards1.pdf

Davis, A. (1998). *The limits of educational assessment*. Oxford: Blackwell.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14,* 481–533. doi:10.1598/ RRQ.14.4.2

Dweck, C. S. (1999). Self-theories: *Their role in motivation, personality, and development*. Philadelphia: Psychology.

Fairbairn, S. (2007). Facilitating greater test success for English language learners. *Practical Assessment, Research & Evaluation, 12*(11). Retrieved from http://pareonline.net/pdf/v12n11.pdf

Fletcher, J. M., Francis, D. J., Boudosquie, A., Copeland, K., Young, V., Kalinowski, S., et al. (2006). Effects of accommodations on high-stakes testing for students with reading disabilities. *Exceptional Children*, *72*, 136.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist, 39,* 193–202. doi:10.1037/0003-066X.39.3.193

Gambrell, L. B., Codling, R. M., & Palmer, B. M. (1996). *Elementary students' motivation to read.* (Reading Research Report No. 52). Athens, GA: National Reading Research Center.

Gamse, B. C., Jacob, R. T., Horst, M., Boulay, B., & Unlu, F. (2008). *Reading First Impact Study final report.* (NCEE 2009-4038). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Goodman, K. S. (with Flurkey, A., Kato, T., Kamii, C., Manning, M., Seay, S., et al.). (Ed.). (2006). *The truth about DIBELS: What it is, what it does*. Portsmouth, NH: Heinemann.

Goodman, K. S., & Goodman, Y. M. (1977). Learning about psycholinguistic processes by analyzing oral reading. *Harvard Educational Review, 47,* 317–333.

Guszak, F. J. (1967). Teacher questioning and reading. *The Reading Teacher, 21,* 227–234.

Guthrie, J. (2001). Contexts for engagement and motivation in reading. *Reading Online, 4*. Retrieved from http://www.readingonline.org/articles/handbook/ guthrie/

Guthrie, J., & Klauda, S. (2016). Engagement and motivational processes in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 41–53). New York: Routledge.

Guthrie, J. T., McGough, K., Bennett, L., & Rice, M. E. (1996). Concept-oriented reading instruction: An integrated curriculum to develop motivations and

strategies for reading. In L. Baker, P. Afflerbach, & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 165–190). Mahwah, NJ: Erlbaum.

Guthrie, J., Wigfield, A., & You, W. (2013). Instructional contexts for engagement and achievement in reading. In S. Christensen, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 601–634). New York: Springer Science.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (Eds.). (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.

Harris, T. L., & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: The vocabulary of reading and writing*. Newark, DE: International Reading Association.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. New York: Cambridge University Press.

Henk, W., & Melnick, S. (1995). The Reader Self-Perception Scale (RSPS): A new tool for measuring how children feel about themselves as readers. *The Reading Teacher, 48,* 470–482.

Hiebert, E. H., Winograd, P. N., & Danner, F. W. (1984). Children's attributions for failure and success in different aspects of reading. *Journal of Educational Psychology, 76,* 1139–1148. doi:10.1037/0022-0663.76.6.1139

Hildebrandt, D. (2001). "But there's nothing good to read" (in the library media center). *Media Spectrum: The Journal for Library Media Specialists in Michigan, 28,* 34–37.

Huey, E. B. (1908). *The psychology and pedagogy of reading, with a review of the history of reading and writing and of methods, texts, and hygiene in reading*. New York: Macmillan.

Institute of Education Sciences. (2017a). Children and youth with disabilities. In *The condition of education.* Washington, DC: Author. Retrieved from http://nces.ed.gov/programs/coe/indicator_cgg.asp

Institute of Education Sciences. (2017b). English language learners in public schools. In *The condition of education.* Washington, DC: Author. Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp

International Reading Association. (1999). *High-stakes assessments in reading: A position statement of the International Reading Association*. Newark, DE: Author.

Jamentz, K. (1994). Making sure that assessment improves performance. *Educational Leadership, 51,* 55–57.

Johnston, P. (1987). Teachers as evaluation experts. *The Reading Teacher, 40,* 744–748.

Johnston, P. H. (2004). *Choice words: How our language affects children's learning*. Portland, ME: Stenhouse.

Johnston, P. H., Afflerbach, P., & Weiss, P. B. (1993). Teachers' assessment of the teaching and learning of literacy. *Educational Assessment, 1,* 91–117. doi:10.1207/s15326977ea0102_1

Johnston, P., & Ivey, G. (2016). Discursive contexts, reading, individual differences. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 209–222). New York: Routledge.

Johnston, P. H., & Winograd, P. N. (1985). Passive failure in reading. *Journal of Reading, 17,* 279–301.

Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York: Holt, Rinehart and Winston.

Kieffer, M., Lesaux, N., Rivera, M., & Francis, D. (2009). Accommodations for English language learners on large-scale assessments: A meta-analysis of effectiveness and validity. *Review of Educational Research, 79,* 1168–1201.

Kober, N., Chudowsky, N., & Chudowsky, V. (2008). *Has student achievement increased since 2002? State test score trends through 2006–07.* Washington, DC: Center on Education Policy.

Koenig, J. A., & Bachman, L. F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments.* Washington, DC: National Academies Press.

Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools.* Portsmouth, NH: Heinemann.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41,* 212–218. doi:10.1207/s15430421tip4104_2

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15,* 309–328.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* New York: Cambridge University Press.

Lee, S., & Jonson-Reid, M. (2016). The role of self-efficacy in reading achievement of young children in urban schools. *Child and Adolescent Social Work Journal, 33*, 79–89.

Leipzig, D. H., & Afflerbach, P. (2000). Determining the suitability of assessments: Using the CURRV framework. In L. Baker, M. J. Dreher, & J. T. Guthrie (Eds.), *Engaging young readers: Promoting achievement and motivation* (pp. 159–187). New York: Guilford.

Lemann, N. (1999). *The big test: The secret history of the American meritocracy.* New York: Farrar, Straus and Giroux.

Leslie, L., & Caldwell, J. S. (2017). *Qualitative reading inventory* (6th ed.). New York: Pearson.

Leu, D., C. Kiili, C., & Forzani, E. (2016). Individual differences in the new literacies of online research and comprehension. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 259–272). New York: Routledge.

Li, Y., & Wang, C. (2010). An empirical study of reading self-efficacy and the use of reading strategies in the Chinese EFL context. *Asian EFL Journal, 12*, 144–162.

Luke, A. (1994). *The social construction of literacy in the classroom.* Melbourne, Australia: Macmillan.

Mansell, W., James, M., & the Assessment Reform Group. (2009). *Assessment in schools. Fit for purpose? A commentary by the Teaching and Learning Research Programme.* London: Economic and Social Research Council, Teaching and Learning Research Programme.

Mathewson, G. C. (1994). Model of attitude influence upon reading and learning to read. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (4th ed., pp. 1131–1161). Newark, DE: International Reading Association.

McCracken, R. A. (1966). *Standard reading inventory*. Klamath Falls, OR: Klamath.

McKenna, M. C., & Kear, D. J. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher, 43,* 626–639. doi:10.1598/RT.43.8.3

McNamara, D., Jacovina, M., & Varner, L. (2016). Higher order thinking in reading. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 164–176). New York: Routledge.

McNamara, D. S., & Magliano, J. P. (2009). Self-explanation and metacognition: The dynamics of reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 60–81). New York: Routledge.

McTigue, E. M., Washburn, E. K., & Liew, J. (2009). Academic resilience and reading: Building successful readers. *The Reading Teacher, 62,* 422–432. doi:10.1598/RT.62.5.5

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.

Merriam-Webster. (2016). *Merriam-Webster online dictionary*. Retrieved from http://www.merriam-webster.com

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Metsala, J., & David, M. (2016). Individual differences in word recognition. In P. Afflerbach (Ed.)., *Handbook of individual differences in reading: Reader, text, and context* (pp. 93–106). New York: Routledge.

Moje, E. B., Ciechanowski, K. M., Kramer, K., Ellis, L., Carrillo, R., & Collazo, T. (2004). Working toward third space in content area literacy: An examination of everyday funds of knowledge and discourse. *Reading Research Quarterly, 39,* 38–70. doi:10.1598/RRQ.39.1.4

Moore, K. A., Redd, Z., Burkhauser, M., Mbwana, K., & Collins, A. (2009). *Children in poverty: Trends, consequences, and policy options* (Child Trends Research Brief No. 2009-11). Washington, DC: Child Trends.

Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation, 8*(14). Retrieved from http://pareonline.net/getvn.asp?v=8&n=14

Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muspratt, S., Luke, A., & Freebody, P. (1997). *Constructing critical literacies: Teaching and learning textual practice*. Creskill, NJ: Hampton.

National Assessment Governing Board. (2015). *Reading framework for the 2015 National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board, U.S. Department of Education. Retrieved from https://www.nagb.org/publications/frameworks/reading/2015-reading-framework.html

National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education. (2003). *Where we stand on curriculum, assessment, and program evaluation*. Washington, DC: Authors.

National Center for Education Statistics. (2016). *Inclusion of special-needs students*. Retrieved from https://nces.ed.gov/nationsreportcard/about/inclusion.aspx

National Center on Education and the Economy. (1998). *New standards: Performance standards and assessments for the schools*. Washington, DC: Author.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.

National Council for the Social Studies. (1994). *Expectations of excellence: Curriculum standards for social studies*. Washington, DC: Author.

National Council of Teachers of English. (2000). *Resolution on urging reconsideration of high stakes testing*. Retrieved from http://www.ncte.org/positions/statements/highstakestestrecons

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. (NIH Publication No. 00-4769)*. Washington, DC: U.S. Government Printing Office.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.

New York State Education Department. (2009). *English language arts learning standards*. Retrieved from http://www.p12.nysed.gov/ciai/ela/elastandards/home.html

New York State Education Department. (2012). *NYS common core ELA curriculum: G3: Curriculum map*. Retrieved from https://www.engageny.org/resource/grade-3-ela-curriculum-map

Nidus, G., & Sadder, M. (2016). More than a checklist. *Educational Leadership, 73,* 62–66.

Ogle, D. M. (1986). K-W-L: A teaching model that develops active reading of expository text. *The Reading Teacher, 39,* 564–570. doi:10.1598/RT.39.6.11

Olson, B., Mead, R., & Payne, D. (2002). *A report of a standard setting method for alternate assessments for students with significant disabilities.* (NCEO Synthesis Report No. 47). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota. Retrieved from https://nceo.info/Resources/publications/OnlinePubs/Synthesis47.html

Organisation for Economic Co-operation and Development. (2011). *Participating countries/economies*. Paris: Author. Retrieved from www.pisa.oecd.org/pages/0,3417,en_32252351_ 32236225_1_1_1_1_1,00.html

Organisation for Economic Co-operation and Development. (2014). *PISA 2012 results: What students know and can do: Student performance in reading, mathematics and science (Vol. 1)*. Paris: Author.

Pajares, F., & Urdan, T. (Eds.). (2006). *Self-efficacy beliefs of adolescents*. Greenwich, CT: Information Age.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1,* 117–175. doi:10.1207/s1532690xci0102_1

Paris, S. G. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*(3), 293–316. doi:10.1016/0361-476X(83)90018-8

Paris, S. G., & Carpenter, R. D. (2003). FAQs about IRIs. *The Reading Teacher, 56,* 578–580.

Partnership for Assessment of Readiness for College and Careers. (2010). *The Partnership for Assessment of Readiness for College and Careers (PARCC) application for the Race to the Top Comprehensive Assessment Systems Competition*. Retrieved from http://www.fldoe.org/parcc/pdf/apprtcasc.pdf

Partnership for Assessment of College and Career Readiness. (2017). *Grade 4 English language arts/literacy test booklet: practice test*. Retrieved from https://parcc.pearson.com/resources/Practice-Tests/TBAD/Gr4ELA/PC1105810_Gr4ELATB_PT.pdf

Payne, D. A. (1974). *The assessment of learning: Cognitive and affective*. Lexington, MA: D. C. Heath.

Pearson, P. D., & Johnson, D. D. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart and Winston.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.

Pennock-Roman, M., & Rivera, C. (2011). Mean effects of test accommodations for ELLs and nonELLs: A meta-analysis of experimental studies. *Educational Measurement: Issues and Practice, 30,* 10–28.

Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education, 7*(2), 93–120. doi:10.1207/s15324818ame0702_1

Pikulski, J. J., & Shanahan, T. (1982). Informal reading inventories: A critical analysis. In J. J. Pikulski & T. Shanahan (Eds.), *Approaches to the informal evaluation of reading* (pp. 94–116). Newark, DE: International Reading Association.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63,* 167–199.

Pitcher, S., Albright, L., DeLaney, C., Walker, N., Seunarinesingh, K., Mogge, S., et al. (2007). Assessing adolescents' motivation to read. *Journal of Adolescent and Adult Literacy, 50*, 378–396.

Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practice, 10,* 12–15. doi:10. 1111/j.1745-3992.1991.tb00211.x

Popham, W. J. (1997). What's wrong—and what's right—with rubrics. *Educational Leadership, 55,* 72–75.

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Erlbaum.

Raphael, T. E., & Wonnacott, C. A. (1985). Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions. *Reading Research Quarterly, 20,* 282–296. doi:10.2307/748019

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction, 33,* 86–112.

Roeber, E. (1996). Guidelines for the development and management of performance assessments. *Practical Assessment, Research and Evaluation, 5,* 1–3.

Schiefele, U., Schaffner, E., Moller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly, 47,* 427–463.

Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy, 2*(4), 419–436. doi:10.1207/S15327078IN0204_02

Schunk, D., & Bursuck, W. (2016). Self-efficacy, agency, and volition. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 54–66). New York: Routledge.

Schunk, D., & Zimmerman, B. (2008). *Motivation and self-regulated learning: Theory, research, and applications*. Hillsdale, NJ: Erlbaum.

Shell, D. F., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attributions, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology, 87,* 386–398.

Shepard, L. A., & Bliem, C. L. (1995). Parents' thinking about standardized tests and performance assessments. *Educational Researcher, 24,* 25–32.

Sireci, S. G. (2004). *Validity issues in accommodating NAEP reading tests.* (Center for Educational Assessment Research Report No. 515). Amherst, MA: University of Massachusetts Amherst.

Slack, J. B. (1998). *Questioning strategies to improve student thinking and comprehension*. Unpublished manuscript, Southwest Educational Development Laboratory, Austin, TX.

Smarter Balanced Assessment Consortium. (2010). *Race to the Top assessment program application for new grants: Comprehensive assessment systems*. Retrieved from http://www.scribd.com/doc/53320952/Sbac-Final-Narrative-20100620-4pm

Smarter Balanced Assessment Consortium. (2014, May 16). *English/language arts practice test scoring guide: Grade 4 performance task*. Retrieved from http://www.smarterbalanced.org/wp-content/uploads/2015/11/G4_Practice_Test_Scoring_Guide_ELA_PT.pdf

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Solheim, O. (2011). The impact of reading self-efficacy and task value on reading comprehension scores in different item formats. *Reading Psychology, 32*, 1–27.

Stahl, K., & Bravo, M. A. (2010). Contemporary classroom vocabulary assessment for content areas. *The Reading Teacher, 63,* 566–578. doi:10.1598/RT.63.7.4

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

Stevens, R. (1912). *The question as a measure of efficiency in instruction: A critical study of classroom practice*. New York: Teachers College Press.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83,* 758–765.

Stretch, L. S., & Osborne, J. W. (2005). Extended time test accommodation: Directions for future research and practice. *Practical Assessment, Research & Evaluation, 10*(8). Retrieved from http://pareonline.net/pdf/v10n8.pdf

Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change, 10,* 211–227. doi:10.1007/s10833-009-9105-2

Taylor, K., & Harris, E. A. (2015, August 11). New York State test questions tricky for 3rd graders, and maybe some adults. *New York Times*. Retrieved from https://www.nytimes.com/2015/08/11/nyregion/new-york-state-test-questions-tricky-for-3rd-graders-and-maybe-some-adults.html

Thorndike, E. L. (1917). Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology, 8,* 323–332. doi:10.1037/h0075325

Thurlow, M., & Bolt, S. (2001). *Empirical support for accommodations most often allowed in state policy*. (NCEO Synthesis Report No. 41). Minneapolis, MN: National Center on Educational Outcomes, University of Minnesota.

UNICEF Office of Research. (2014). *Children of the recession: The impact of the economic crisis on child well-being in rich countries*. (Innocenti Report Card 12). Florence, Italy: Author.

U.S. Census Bureau. (2015). *Poverty thresholds* (2016). Retrieved from http://www.census.gov/hhes/www/poverty/data/threshld/

van den Broek, P., Mouw, J., & Kraal, A. (2016). Individual differences in reading comprehension. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 138–150). New York: Routledge.

Van Keer, H., & Verhaeghe, J. P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *Journal of Experimental Education, 73,* 291–329. doi:10.3200/JEXE.73.4.291-329

VanSledright, B. (2002). *In search of America's past: Learning to read history in elementary school*. New York: Teachers College Press.

VanSledright, B. A. (2010). *The challenge of rethinking history education: On practices, theories, and policy*. New York: Routledge.

Veenman, M. (2016a). Metacognition. In P. Afflerbach (Ed.), *Handbook of individual differences in reading: Reader, text, and context* (pp. 26–40). New York: Routledge.

Veenman, M. (2016b). Motto [translated]. Metacognition Research Institute home page. Retrieved from http://www.instituutvoormetacognitieonderzoek.nl/home

Verhoeven, L., & Snow, C. E. (Eds.). (2001). *Literacy and motivation: Reading engagement in individuals and groups*. Mahwah, NJ: Erlbaum.

Virginia Department of Education. (2015). *Limited English proficient students: Guidelines for participation in the Virginia Assessment Program.* Richmond, VA: Author. Retrieved from http://www.doe.virginia.gov/testing/participation/lep_guidelines.pdf

Vygotsky, L. S. (1978). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. and Trans.), *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press. (Original work published 1934.)

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101,* 192–212. doi:10.1037/0033-2909.101.2.192

Walpole, S., & McKenna, M. C. (2006). The role of informal reading inventories in assessing word recognition. *The Reading Teacher, 59,* 592–594. doi:10.1598/RT.59.6.10

Wang, M. C., Haertel, G., & Walberg, H. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research, 84,* 30–43.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review, 20,* 158–177. doi:10.1037/h0074428

Webb, N. (2002). Depth of knowledge: Four content areas. Retrieved from http://schools.nyc.gov/NR/rdonlyres/2711181C-2108-40C4-A7F8-76F243C9B910/0/DOKFourContentAreas.pdf

Wells, G. (1989). Language in the classroom: Literacy and collaborative talk. *Language and Education, 3,* 251–273. doi:10.1080/09500788909541266

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25,* 68–81. doi:10.1006/ceps.1999.1015

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.

Wineburg, S. (1997). Beyond "breadth and depth": Subject matter knowledge and assessment. *Theory into Practice, 36,* 255–261. doi:10.1080/00405849709543776

Wolf, D. P. (1987, Winter). The art of questioning. *Academic Connections,* 1–7. Retrieved from http://www.exploratorium.edu/IFI/resources/workshops/artofquestioning.html

Zuriff, G. E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. *Applied Measurement in Education, 13,* 99–117.
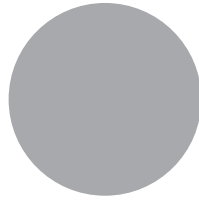
ABOUT THE AUTHOR

Peter Afflerbach is a professor of education at the University of Maryland, College Park. He received his doctorate in reading from the University of Albany. His research interests include reading assessment, individual differences in reading, reading strategies, and the verbal reporting methodology. Afflerbach formerly taught K–6 remedial reading, junior high school remedial reading, and high school English in New York public schools.

Afflerbach serves on the Standing Reading Committee of the National Assessment of Educational Progress (NAEP). He recently completed terms on the Literacy Research Panel of the International Literacy Association, and the National Academy of Education/NAEP Committee to investigate reading achievement–level descriptors. Previously, he was chair of the Assessment Committee of the International Literacy Association and served on the Common Core State Standards Review and Feedback Panels.

Afflerbach, who was inducted into the Reading Hall of Fame in 2009, is the editor of the *Handbook of Individual Differences in Reading: Reader, Text, and Context* (2016) and co-editor of the *Handbook of Reading Research,* 4th edition (2010) and 5th edition (in press). He has published in numerous theoretical and practical journals, including *Reading Research Quarterly, Cognition and Instruction, Elementary School Journal, Journal of Reading, Journal of Reading Behavior, Language Arts, Theory into Practice,* and *The Reading Teacher.*

Related ASCD Resources

At the time of publication, the following resources were available (ASCD stock numbers appear in parentheses):

Print Products

Educational Leadership: Using Assessments Thoughtfully (March 2014) (#114023)

Educational Leadership: Getting Students to Mastery (December 2013) (#114021)

Checking for Understanding: Formative Assessment Techniques for Your Classroom, 2nd Edition by Douglas Fisher and Nancy Frey (#115011)

Assessment and Student Success in a Differentiated Classroom by Carol Ann Tomlinson and Tonya R. Moon (#108028)

The Formative Assessment Action Plan: Practical Steps to More Successful Teaching and Learning by Nancy Frey and Douglas Fisher (#111013)

Taking Action on Adolescent Literacy by Judith L. Irvin, Julie Meltzer, and Melinda S. Dukes (#107034)

For up-to-date information about ASCD resources, go to www.ascd.org. You can search the complete archives of *Educational Leadership* at www.ascd.org/el.

ASCD EDge® Group

Exchange ideas and connect with other educators on the social networking site ASCD EDge at http://ascdedge.ascd.org/.

ASCD myTeachSource®

Download resources from a professional learning platform with hundreds of research-based best practices and tools for your classroom at http://myteachsource.ascd.org/.

For more information, send an e-mail to member@ascd.org; call 1-800-933-2723 or 703-578-9600; send a fax to 703-575-5400; or write to Information Services, ASCD, 1703 N. Beauregard St., Alexandria, VA 22311-1714 USA.