

CHAPTER 5

Performance Assessment

We are surrounded by performances related to our reading. When we read instructions to assemble a child's toy, read a recipe to make dinner, or read a set of directions for taking medicine, reading and action are intertwined. In school, students read earth science articles downloaded from the Internet and craft a three-dimensional diorama representing plate tectonics. Students create a dramatic skit based on their interpretation of characters in a short story. Students read two history texts that pose competing and conflicting accounts of slavery and determine which text is more accurate and trustworthy. All of these students are using the knowledge they gain through reading to perform important tasks.

Starting in elementary school, performance assessment can help us determine not only what students understand from reading but also how they use what they understand. Performance assessment helps describe student development in relation to complex, curricular goals. The knowledge that students gain from reading is not inert, and reading assessment should not treat it as such. Performance assessments situate students' reading in relation to important work in school, as students read to construct meaning and use it. In many classrooms, performance assessment represents a sea change in how assessment is conceptualized and conducted. Performance assessment is demanding of resources and teacher expertise but has many features that are worth the expense and effort.

A Brief History of Performance Assessment

The close connection between reading and performance tasks is as old as reading itself. However, current interest in performance assessment is attributable to several converging factors. First, ongoing research on how people read, think, and learn describes the complexity of reading, and areas of student growth and achievement related to reading (McNamara & Magliano, 2009; OECD, 2010; Pajares & Urdan, 2006; Pressley & Afflerbach, 1995; Snow, 2002). Second, the reading standards developed by professional organizations, states, and school districts in relation to this new research knowledge require assessments that capture and describe complex student achievements, such as those expected by the Common Core State Standards (CCSSO & NGA, 2010). Third, traditional reading tests do not give a full account of how well students read and use what they understand from reading (Wiggins, 1998), and these tests

may be negative influences on curricula and teaching (Frederiksen, 1984). Fourth, efforts to develop large-scale and consequential performance assessment programs have demonstrated their viability (National Center on Education and the Economy, 1998). Finally, performance assessments can be used in the teaching and learning of course content and to help students learn to self-assess (Black & Wiliam, 1998).

Accompanying the development of performance assessments is the belief that reading assessment can better reflect what is known about academic reading and inform educators of students' progress related to reading in college and professional life (PARCC, 2010; SBAC, 2010). There is a considerable disconnect between the breadth and depth of students' learning in a high-quality reading curriculum and the narrow focus of many reading assessments (Davis, 1998). There must be the commitment to reducing the gap between what students learn and how we assess this learning. As we expect reading standards to reflect our most current understandings of reading, we should expect the same of the assessments used to measure student progress. It is difficult to understand the detail of a student's achievement when reading assessment results are a letter grade of B, a raw score of 53, or a ranking in the 82nd percentile, and performance assessments provide illustrative and educative details that complement the general statement made by such single, summative assessment markers.

Most recently, performance assessment has been the focus of broad initiatives to improve teaching and learning in U.S. schools. The Common Core State Standards (CCSSO & NGA, 2011) present an ambitious set of learning goals. A major focus is on students using what they understand from reading in content areas. Thus, the Common Core State Standards represent a challenging curriculum that requires a new generation of reading assessments, and such assessments are promised by two assessment consortia, the PARCC and the SBAC. Assessments developed by these consortia will regularly require student performances to measure and report the complexity of learning that is demanded by the Common Core State Standards.

The SBAC (2010) describes general criteria for reading assessments (i.e., criteria that align with performance assessments): "Assessments must be carefully structured to improve teaching and learning. This means establishing summative assessments that reflect the challenging CCSS [Common Core State Standards] content, emphasizing not just students' 'knowing,' but also 'doing'" (p. 37). Similarly, the PARCC (2010) describes assessment as the following:

In...ELA/literacy...in grades three through high school, students will take focused assessments...and participate in an extended and engaging performance-based task...students will be given extended time to identify or read relevant research materials and compose written essays based on them. Afterwards, students will publicly present the results of that research and writing to their classmates, answering questions or engaging in debate, so that teachers can assess students' speaking and listening skills using a common rubric. (p. 36)

So far, 44 states and the District of Columbia have formally adopted the Common Core State Standards. With this adoption, there is the concomitant acknowledgment that reading assessment must change, and this change will be in the direction of performance assessments.

Characteristics of Performance Assessment

Our knowledge of reading is continually evolving. The vast amount of reading research conducted in the past several decades has added much to our understandings of how reading works, how reading ability develops, and the place of reading in the lives of students (Berkeley et al., 2011; Moje et al., 2004; Rogers, 2003). Performance assessments offer opportunities to describe and measure student reading and learning across the curriculum. In the content domains of science, social studies, math, English, and the arts, effective performance assessments share important characteristics. Performance assessments help us understand student growth and learning. They describe student achievement in detail. They are developed in relation to comprehensive analyses of what students must do to understand and complete the performance. Effective performance assessment uses scoring rubrics and guides that inform teachers and students as to the nature of the performance and gradations of accomplishment related to the performance.

In many cases, the nature and structure of performance assessments reflect classroom practices in the content areas. Performance assessments may focus on how students synthesize information learned from different sources, how they put to use the information learned from reading in classroom tasks, and how reading plays a role in students' growth and development. Thus, the inferences made from performance assessments may emanate from the daily routines of the classroom—a form of authentic assessment (see the Reading Assessment Snapshot for Authentic Assessment at the end of Chapter 4). Further, performance assessments' open-ended response formats can accommodate different student approaches to tasks and problem solving, reflecting students' constructed understanding of what they read.

Effective performance assessment requires the identification of key student learning and related performances, aided by consultation with research related to the demands of the performance. For example, if a performance assessment requires students to collect information from different texts and synthesize that information, the developers of the assessment may consult the research literature on synthesizing information from text (Bråten et al., 2009) and the special requirements for performing well in reading in content domains such as history (VanSledright, 2010; Wineburg, 1997). This type of research helps us understand specific students' abilities that can be tapped with performance assessment.

Effective performance assessment is enabled by a detailed task analysis of the performance that is the focus of the assessment. As the performance task is created, it

is imperative that we establish an informed, a priori understanding of exactly what we are asking students to do. Task analyses involve a meticulous accounting of the things that must be done in order to perform an assessment task satisfactorily. What exactly are we asking students to do when they read the following prompt in an American history performance assessment: “Based on your understanding of the two primary source texts, develop an account of the challenges faced by the Jamestown colonists?” Later in this chapter, we observe that responding to such prompts can involve literal, inferential, and critical understanding; knowledge of source text cues; and the ability to take a critical stance toward authors and their work. Using task analysis, we can check to see that what we ask of students represents an important and legitimate assessment request.

The grounding of performance assessment in relation to reading research and task analysis helps in the development of a performance assessment’s core features, including a task demand or prompt, rubric, and scoring guide. Also, the development of a performance assessment is facilitated by attention to the assessment’s specific purpose, goal, and objectives. Thus, we should regularly ask questions like the following:

- Why are we asking students to do this?
- What will it show us?
- How do students demonstrate their accomplishment of the goal with observable products and performances?

The goal here is building a performance assessment in which we have high confidence and from which we can make legitimate inferences about students’ development and achievement.

Worthwhile performance assessments are designed in relation to logical and understandable rubrics and scoring guides. A rubric is a description of how a student performance may vary in terms of quality and achievement. Popham (1997) describes three essential features of rubrics: (1) the criteria used to evaluate student work, (2) specification of differential quality of student work, and (3) the means to consistently and accurately score student work. An examination of Table 12 illustrates how these three factors operate with a rubric. In this case, the performance assessment focuses on two interrelated history reading goals, or performances: (1) Students read history texts to identify cues in text to determine source text status and then (2) use this source text status to evaluate the trustworthiness of the text. The criteria used in this rubric to evaluate student work are related to ability to identify cues to determine source text status and the ability to use the determination of source text status to evaluate the text for trustworthiness (VanSledright, 2002). The specification of different quality of student work is apparent in the different descriptions of performance. Following this rubric through the different levels of student performance, we are able to trace how students might move from “Not Apparent” to “Developing” to “Proficient”

TABLE 12
History Performance Assessment Rubric: Identifying Cues to Determine Source Text Status and Using Source Text Status to Evaluate Text

Category	Levels of Student Performance			
	Not Apparent (1)	Developing (2)	Proficient (3)	Exemplary (4)
Identifying cues to determine source text status	There is no apparent means for identifying and using cues to determine source text status.	Student identifies and uses one or more cues to determine source text status. For example, student uses archaic spelling and author voice to attribute source text status. However, identification of type of text cue is inaccurate, and/or determination of source text status is erroneous.	Student identifies and uses one or several cues to correctly determine source text status. For example, student uses archaic spelling and author voice to correctly determine text status as primary source.	Student identifies and uses all possible cues to correctly determine source text status. For example, student uses each and every cue contained in text to correctly determine text status as primary source.
Using source text status to evaluate text for trustworthiness	There is no apparent attempt to use understanding of source text status to evaluate the text. There is no apparent attempt to connect source text status with an evaluation of the trustworthiness of text.	Student attempts to use understanding of source text status, demonstrating connection between source text status and evaluation. However, designation of source text as trustworthy or not is erroneous.	Using understanding of source text status, student demonstrates connection between source text status and evaluation. Evaluation of text's trustworthiness is accurate and appropriate.	Using understanding of source text status, student demonstrates connection between determination of source text status and accuracy of evaluation. Evaluation of text's trustworthiness is comprehensive and appropriate, and use of source text status to evaluate text is elaborated.

Note. Adapted from “Teaching and Learning Self-Assessment Strategies in Middle School” by P. Afflerbach and K. Meuwissen, in *Metacognition in Literacy Learning: Theory, Assessment, Instruction, and Professional Development* (pp. 141–164), edited by S.E. Israel, C.C. Block, K.L. Bauserman, and K. Kinnucan-Welsch, 2005, Mahwah, NJ: Erlbaum. Copyright 2005 by Lawrence Erlbaum Associates. Adapted with permission.

to “Exemplary” levels of performance, earning respective scores of 1–4. The means to consistently and accurately score student work is tied to the clarity of the category, the detail of specification of student work of differing quality, and the assignment of that specification to particular scoring points.

Andrade (2000) describes several possible benefits of using rubrics and performance assessment. Well-structured rubrics communicate to students the expectation for learning. In effect, they tell the student what is needed to earn a particular grade. Rubrics also provide students with detailed feedback about their learning. By comparing their work with the set of expectations that a rubric presents, students can determine what has been achieved and what may remain to be accomplished.

The careful development of rubrics for performance assessments provides teachers and students with a clear and consistent means for judging student work. A rubric may also be used to help teach important aspects of the required student performance because the rubric focuses our attention and efforts on key curricular goals. Detailed performance assessment rubrics offer opportunities for scaffolding instruction. Teachers can use the different levels of the rubric to teach in reference to students’ zone of proximal development (Vygotsky, 1934/1978), helping one student move from a level 1 (“Not Apparent” in Table 12) performance to level 2 (“Developing”) and helping another move from level 3 (“Proficient”) to level 4 (“Expert”). Indeed, Popham (1997) characterizes well-designed rubrics as “instructional illuminators” (p. 75). In the case of the history curriculum, the rubric provides a consistent model of the types of cues that students should use as they determine source text status and the manner in which they should use these cues to demonstrate thinking and learning in history.

Popham (1997) also provides two caveats. First, the performance that is being assessed must be worthy of the time and effort that are needed to develop and use truly effective rubrics. We do not need a performance assessment to understand students’ literal comprehension of history facts because a multiple-choice or open-ended question might suffice. The array of student learning that is expected over an entire school year, as represented in statewide and school district learning standards, requires that we be judicious in choosing and developing performance assessments. From the start, performance assessment development must be approached with the knowledge that there are limits to how many assessments might be developed, given the intensive work they demand. Performance assessment tasks must reflect consensus on important student learning. Ideally, these assessments will focus on performances that represent convergences of student learning, incorporating skills, strategies, and the content domain learning that are the goals of effective instruction. A performance assessment rubric that is well designed but not at the core of important student learning may indicate unwise use of limited school resources. Also, the rubric must reflect a detailed understanding of the performance that students undertake, as informed by

continuous comparisons of the performance rubric and task analyses. Otherwise, exceptional student work might not be in the direction of core learning goals and may be missed by assessment.

The effectiveness of performance assessment is enhanced by examples of student performances, for they help make explicit what students' final products can look like. Combined with instruction that helps students learn and perform the processes that yield such products, representative performances can clearly signify accomplishment at specific levels. These sample performances should serve as models of specific levels of achievement for both the scorers of the assessment and the students who are required to create the performance. For example, a scoring guide that allows us to assign a score between 1 and 4 can be accompanied by examples of student work that provide students with a target for their performance. Student work samples, perhaps taken from the previous year's class, help students develop a refined sense of what work scored as a 1 and work scored as a 4 look like and how they differ. Through their public display, these samples can provide students with consistent guidance as they build their performance toward successful completion. The samples in Figure 6 demonstrate a student's ability to locate and use different cues to determine the primary or secondary source status of history texts and then evaluate the trustworthiness of the

FIGURE 6
Example of a Proficient (Level 3) Performance Assessment Score:
Student Reading History Text, Using Cues to Determine Source
Text Status, and Evaluating Text for Trustworthiness

Text excerpt that contains cues used by student:

As a citizen of Jamestown verginia I do my demesticall work each day. Yet, there remains little food for us, the poore distressed subjects. We are in dandger of losing our bretheren to hunger, this hunger that hath no seeming end.

Using cues to determine that it is a primary source text, student writes:

This appears to be a primary source text because it has old spellings like "poore," "dandger," and "hath." Also, I get a good feel for the author of this text as he speaks in the first person—this makes the description more real for me.

Using determination of source text status to evaluate the text's trustworthiness, student writes:

I believe that this is a primary source text because of the archaic spellings and the nature of the author's voice. Therefore, I believe that the text is trustworthy and that it reports an eyewitness account of some of the suffering that the Jamestown colonists experienced.

Note. From *Understanding and Using Reading Assessment, K–12* (p. 99), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

text to earn a score of 3, thus providing useful guidance to students who are learning this strategy. Over time, our use of rubrics and work samples in the classroom contributes to students developing specific schemata for what good work looks like, strategies for progress, and a schema for the ongoing self-assessment of their progress toward performance goals.

A final characteristic of performance assessments is that they often require an expanded time frame to fully realize their potential. Many standardized assessments, including reading tests, are administered in a compact time frame. However, performance assessments require time for learning to use a rubric to support learning, conducting self-assessment, and describing and measuring student learning. Performances take more time than answering simple questions. Performance assessment done well is a result of the consideration of the time needed to participate in and complete a performance, as well as the provision of a realistic time frame that is tailored to and appropriate for eliciting and accommodating students' performances.

Performance Assessment in Ivan's Eighth-Grade History Classroom

Ivan teaches eighth-grade history and has used performance assessment for five years. His enthusiastic use of performance assessment is guided by his strong conviction that students who are finishing middle school and about to enter high school must be able to understand complex texts and then apply the things they learn from reading. This belief dovetails with the intent of the Common Core State Standards (CCSSO & NGA, 2010), to help prepare students for success in college and their careers. Ivan's interest in performance assessment is shared with many teachers across the school district. Across grades and content areas, students' reading is assessed with performance assessments. Students read math word problems that require them to determine rainfall averages for each month of the school year, and science texts to help them interpret their firsthand observations of the schoolyard ecosystem. Students read history texts with competing, and sometimes conflicting, accounts of historical events and characters, from which students must construct an understanding of history and how it is created. The eighth graders are continuing their education in a school system that promotes critical inquiry into what is learned, how it is learned, and why it is learned. The history curriculum in Ivan's class reflects our most current understanding of how people read and learn about history and is replete with performances that require students to demonstrate their understanding (National Council for the Social Studies, 1994; Wineburg, 1997).

Within the history curriculum, students learn to analyze historical documents, assess primary and secondary source status of texts, and combine diverse sources to

construct accounts of history. Students learn to adopt a questioning stance toward the texts they read. Students question the authors of the texts and read critically, considering competing accounts of history and discerning fact and opinion. In this curriculum, students learn how history is written, the materials that are used by authors to write history, and how one can critically read history to try to determine the reliability and trustworthiness of texts. This learning complements students' understanding of historical facts about people, places, dates, and events.

The above tasks and learning goals relate directly to the reading standards for literacy in history/social studies in grades 6–12 that are contained in the Common Core State Standards. Specifically, the performance assessment in Ivan's classroom helps students learn and perform as related to the following categories of the Common Core State Standards (CCSSO & NGA, 2010, p. 14):

Key ideas and details

- Cite specific textual evidence to support analysis of primary and secondary sources.

Craft and structure

- Identify aspects of a text that reveal an author's point of view or purpose (e.g., loaded language, inclusion or avoidance of particular facts).

Integration of knowledge and ideas

- Analyze the relationship between a primary and secondary source on the same topic.

The Nature of the History Performance Assessment

The performance assessment in Ivan's Colonial American history class requires students to read and evaluate history texts like historians. Research demonstrates that historians have specific strategies that help them determine the primary or secondary source status of history texts (VanSledright, 2002, 2010). Primary source texts may include original newspaper articles, diary entries, letters, maps, cartoons, and other documents that are contemporary with the historical period being studied. Secondary sources involve interpretation and reporting of primary sources and may include history textbooks, historical novels, and works representing syntheses of other texts. Students' critical performances in eighth-grade history require that they read three different texts about the same historical events (in this case, the founding and survival of the Jamestown Colony). The texts include a diary excerpt from a Jamestown colonist, an account from a work of fiction of how Jamestown changed the lives of Native Americans, and a newspaper article that describes how the arrival of a second group of colonists saved Jamestown from extinction.

The performance assessment prompt reads as follows:

You are to read the three texts on Jamestown. For each text, identify and use as many cues as possible that help you determine the status of the source text. Once you have identified the cues, determine text status and then evaluate each text for trustworthiness based on these determinations.

Students must locate and use cues in the three different texts to determine when the texts were written, who may have written them, and why the texts were written. Ultimately, students must determine the status of these source texts. Students use particular strategies and cues to make critical evaluations of text, including the trustworthiness of the historical account and the reliability of the author. These critical reading strategies and the intent to read text critically are key aspects of reading history that are goals of the eighth-grade curriculum.

Based on this understanding of reading history, the curriculum is designed to help students understand the texts that they read and teach students how to find and use particular cues in history texts that may help identify a text as a primary or secondary source. The different types of cues, including vocabulary, spelling, syntax, author voice, type of text, age of text, and header material (i.e., where the text comes from), derive directly from research on reading history and are then built into the curriculum and assessment (Afflerbach & VanSledright, 2001). Table 13 lists the types of cues that students can use to help determine source text status.

In the year-end performance assessment, students are responsible for identifying cues and using them to determine if the texts read are primary or secondary source documents. Next, students are expected to evaluate the texts for their trustworthiness, based on students' determination of primary or secondary source status. Ivan believes that consistent use of performance assessments to measure complex student

TABLE 13
Cues Used to Determine the Primary and Secondary Source Status of History Texts

- Age of the text
- Type of text
- Header material in text (displays author attributions)
- Author presence and voice in text
- Spelling in the text
- Syntax
- Vocabulary
- Combinations of two or more of the above cues

Note. From *Understanding and Using Reading Assessment, K–12* (p. 101), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

learning and help students better understand assessment is well worth the time and effort he spends on them. He expects that the use of performance assessments across each school year will reveal important student learning and help students build an appreciation for both their accomplishments and their increasing ability to self-assess.

Consequences and Usefulness of Performance Assessments

Ivan uses performance assessment for several reasons. First and foremost, it allows him to assess his students in a manner that honors the complexity of their learning and his teaching. Recall that this performance assessment derives from the detailed understanding of how accomplished readers read history. This understanding moves our expectations for assessment in history beyond the traditional inventory of students' memory of historical dates, places, people, and events. Learning goals and performance assessment converge as students demonstrate their knowledge about why history is written, how it is written, who writes it, and how it is read. The assessment does not discount the importance of historical dates, places, people, and events but seeks to describe how students learn and understand history. Because the performance assessment reflects the detail of what we know to be critical aspects of reading history, the assessment provides a direct view of student achievement in relation to this. The performance assessment used in Ivan's classroom demonstrates not only students' understanding of the content of history texts but also students' ability to determine primary or secondary source status of those texts.

Second, the use of this history performance assessment encourages Ivan to continually reflect on the content and process of schooling. As noted by Jamentz (1994), the establishment of clear performance goals and outcomes (and their alignment) provides educators with a window on their standards and their teaching craft intended to help students meet those standards. The performance assessment provides Ivan with detailed information about his students' progress toward major goals of schooling. For example, the source text status assessment helps him understand students' critical reading development within history and critical reading in general. Ivan is able to infer, with high confidence, what his students learn and the effectiveness of his teaching.

Third, the performance assessment provides formative and summative assessment information on student learning in relation to the rich construct of reading history. Ivan uses class time to focus on each cue that readers can use to determine the primary or secondary source status of the history texts they read. He constructs assessments, based on the unit-end performance assessment, that describe each student's ability to recognize and use the different cues. These assessments provide formative feedback on the effectiveness of Ivan's instruction and the extent of student learning. Across the school year, he knows each student's achievement in relation to learning the different cues that help the reader determine the source text status.

A fourth use of performance assessments is their means for teaching important content and strategies for reading like a historian. The detailed information included in Ivan's performance assessment rubrics, scoring guides, and sample student papers is useful for communicating to students what they must do to receive particular performance assessment scores. The rubrics and scoring guide help him focus on important curricular content, materials, and processes. A task analysis of locating cues to determine the status of history texts is used to build the performance assessment, and this analysis also reminds Ivan of the level of detail that is necessary in his instruction. For example, he knows that many of the primary source texts in Colonial American history are marked by archaic spellings, such as *wee* for *we* and *bye* for *by*. The close match between instruction and scoring rubrics allows Ivan to focus on helping students develop a strategy for identifying archaic spelling during class readings, which helps students with their task at hand and prepares them for the performance assessment.

The rubric itself focuses on the particular means for students to demonstrate their learning and is utilized as an outline for learning. For example, the rubric's focus on cues of archaic vocabulary, spelling, and syntax prompts Ivan to locate and use samples of historic writing that contain clear examples of such language. During classroom instruction, he uses texts with archaic spelling and vocabulary words that are associated with the Colonial era. Guided by the rubric, Ivan develops instruction that helps his students understand the nature of archaic language and how the presence of archaic language may be an important clue as to the age and source status of a history text. He uses this approach to scaffold instruction with each of the cues mentioned in the rubric. Thus, the instructional focus mirrored in the rubric and scoring guide moves from archaic spelling to syntax, and then on to author voice and text attributions (i.e., the information that may help students identify a text as a primary or secondary source text).

A fifth use of the performance assessment is teaching students self-assessment. Independence and success with increasingly complex reading performances are foci of the Common Core State Standards. As students develop their performances, they learn to self-check their work against the detailed information in the rubric. Here, the rubric provides the necessary scaffold for students who are learning to self-assess and judge their performances. Ivan further pursues his goal of teaching students how to read like a historian and assess their work by creating student checklists. He prepares a checklist that matches each of the particular cues that students may use and that is tied to the rubric. Students are required to use the checklist to determine if a particular type of cue is present in the history texts they read. Figure 7 is a checklist used by Ivan and his fellow history teachers to help students learn and attend to the different text cues in primary and secondary source texts.

Ivan's goal here is to help students practice using the teacher-provided checklist on a regular basis so that they eventually internalize the criteria it contains. This provides

FIGURE 7
**Sample Checklist for Cues Used to Signal Primary
and Secondary Source Texts**

The texts we read in history class may be primary or secondary source texts. Use the following checklist to help you remember particular cues and practice using these cues to determine if a text is a primary or secondary source.

Does this text contain cues that are normally found in primary source texts?

- Yes
 No

Check the following cues as you find them in the text:

- Spelling
 Vocabulary
 Author voice
 Print material attribution
 Combination of two or more of the above cues

Based on your reading and search for cues in this text, which do you determine it to be?

- Primary source text
 Secondary source text

Note. From *Understanding and Using Reading Assessment, K–12* (p. 103), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

students with the means to read and critically evaluate history texts independently. By using checklists, Ivan is helping his students build a useful schema for conducting self-assessment. When we examine national, state, and district standards, we note that students are expected to read with superior comprehension. We anticipate that they will apply what they learn through reading, and we hope that self-assessment will develop as a necessary part of their reading achievement. Ivan believes that students who do not learn self-assessment will forever be dependent on another to tell them how they are doing in reading. The use of checklists in performance assessment is not confined to the eighth-grade history classroom. In fact, the districtwide use of checklists in different content areas and grades breeds a familiarity with assessment that helps students build independent self-assessment routines.

A final use of performance assessments relates to their communicative ability. Performance assessment helps Ivan portray the richness of student learning and the quality of his teaching. As described previously, the assessment is used regularly to communicate to students both the goals of instruction and the means to reach these goals. The performance assessment also communicates to students the ways to self-assess. Seeing the benefits of performance assessment and the ability to describe in detail student learning and teacher accountability, teachers and administrators become supporters of performance assessments. These assessments demonstrate to parents and other school community members the complexity of learning and the success of

teaching. The parents who fully understand the importance of their children developing as critical readers appreciate how the history performance assessment helps them move toward this ability. The detailed examples provided by performance assessment help parents understand the advantage of performance assessment relative to more traditional reading tests.

Roles and Responsibilities Related to Performance Assessment

Ivan has four priorities when using performance assessments: (1) conducting reliable assessment using rubrics and scoring guides, (2) using rubrics and scoring guides as teaching tools in history, (3) using scoring guides and rubrics to help students learn self-assessment, and (4) helping others learn and appreciate the value of performance assessment as he advocates for these assessments. The teacher who demonstrates how performance assessments assist instruction, guide student work, and provide a rich measure of students' accomplishments converts parents to this manner of assessment. Each of these priorities creates a series of roles and responsibilities.

Useful performance assessments derive from the careful consideration of what exactly students do as they undertake and complete a performance. As described earlier in this chapter, Ivan uses task analysis to create a detailed account of what students must know and do to perform assessment tasks. Also, teachers must regularly check to see that their instruction matches the learning and work that is demanded of students by particular performances. Teachers should check to determine the alignment of their instruction, learning standards, and the performance assessment and should be alert to the many different components of student performance. Otherwise, an accurate interpretation of students' reading may be confounded with other factors. A performance assessment that is created to have students demonstrate their content area learning may also involve reading skills and strategies, prior knowledge, motivation, and social interactions. Should the performance assessment be assumed to measure only reading and content area knowledge, the possibility of mismeasuring student achievement increases.

Performance assessment tasks and student performances are not, by themselves, necessarily transparent. For the student, parent, or teacher who does not understand how the score was derived, a score of 3 on a performance assessment is no less opaque than a score of 78 on a standardized, norm-referenced test. Students must have guidance in the process of conceptualizing and doing the performance so that they are working in the direction of the goal of the performance assessment. Ivan wants both the sample performances and anchor papers (i.e., papers that illustrate the essence of a particular score and are used by scorers to guide their work) to provide details to students about the differences in performance, or what determines a score of 4 and what determines a score of 1. In addition, the samples should help describe the space

between, or the zones in which students can work and develop their abilities toward a better performance. Thus, a key responsibility for Ivan is to learn to use rubrics, scoring guides, and example papers productively; to teach performance; and to help uncover the “black box,” or the relatively unknown ways and means of assessment, for his students (Black & Wiliam, 1998, p. 139). He works with performance assessments across the school year and has developed the ability to teach *with* the assessment and not *to* it. That is, the rubrics and examples of student work become tools for teaching in Ivan’s hands. As the performance assessment mirrors important classroom learning and achievement, teaching with the assessment is an important daily routine.

Ivan’s work with performance assessment, his task analysis of what it means to read history, and his knowledge of the curriculum allow him to teach in relation to curricular goals, the score points on the history source text rubric, and the Common Core State Standards. Note that the difference between a score of 2 and a score of 3 (in Table 12, these are “Developing” and “Proficient,” respectively) on the rubric relates to becoming familiar with using particular strategies and then using those same strategies accurately and successfully. Ivan matches this important difference in student performance scoring points to his knowledge of the curriculum. He determines that his teaching, in relation to the score of 2 on the performance assessment, should focus on accurate identification of the type of text cues and accurate designation of the trustworthiness of the text. Here, performance assessment helps Ivan define and meet an important series of responsibilities.

Performance assessments often assume student reading ability, rather than directly assessing this ability. Although some assessments, such as standardized reading tests, can provide information about the degree of development of students’ decoding, word recognition, reading comprehension strategies, and vocabulary, performance assessments may take this development as a given. Thus, performance assessments most often require an existing level of reading proficiency for entry into the task. This special aspect of performance assessments in reading demands that Ivan place his students in performance situations where they can succeed at the reading work. Students asked to perform complex tasks when they do not have the appropriate reading ability will experience prolonged failure. This undesirable scenario reminds us that task analyses should be conducted to determine what a particular performance might demand of students. Performance assessment practice must be supplemented by careful assessment of individual students’ achievements and needs. Ivan does not want to misunderstand the information that is provided by the performance assessment. The fact that the history performance assessment demands that students can read at grade level is incentive for him and his colleagues to keep a close watch on individual students’ reading ability. This helps prepare students for the assessment and allows for the allocation of class time to help build reading skills and strategies that are prerequisites to adequate performance on the history reading performance assessment.

Over time, students who are involved with performance assessment should experience growth in their ability to self-assess their reading. Performance assessment rubrics provide a degree of transparency to the assessment and represent an opportunity for students to gradually control the assessment of their work. The increased control of reading and performance derives from students who are active participants in their reading. The gradual assumption of control contributes to students' increased independence and success, which can stimulate motivation to read (Aarnoutse & Schellings, 2003). Ivan's responsibilities in relation to teaching self-assessment are guided by his understanding of the need to make clear the ways and means of self-assessment. In the daily routines of the classroom, what provides opportunities to model, discuss, and demonstrate self-assessment for students? As Ivan's students master using different cues to detect source text status, he introduces checklists that help students build toward independent and reliable routines for assessing their knowledge.

Ivan is fortunate to have colleagues with whom he can discuss and plan performance assessments. The performance assessment group uses a list of performance assessment goals and objectives (Moskal, 2003) that describes the responsibilities related to developing and using these assessments effectively and anticipates the assessments being developed by the Race to the Top assessment consortia:

1. The statement of goals and accompanying objectives should provide a clear focus for both instruction and assessment....
2. Both goals and objectives should reflect knowledge and information that is worthwhile for students to learn....
3. The relationship between a given goal and the objectives that describe that goal should be apparent....
4. All of the important aspects of the given goal should be reflected through the objectives....
5. Objectives should describe measurable student outcomes....
6. Goals and objectives should be used to guide the selection of an appropriate assessment activity. (p. 2)

Attention to the roles and responsibilities of performance assessment means that the school district realizes its promise and assigns suitable resources to do so.

States and school districts should use reading research to develop current and well-informed standards for reading across grade levels and content areas. The complexity of the reading performances related to the Common Core State Standards (CCSSO & NGA, 2010) reflects the complexity of reading, as described by contemporary research. One result is that administrators must attend to the call for more authentic assessment, or assessment that allows us to know how students work when they are given complex assignments or learning tasks (Wiggins, 1998). Ivan's building principal and the district curriculum administrators work with teachers to build an understanding of the challenges and benefits of performance assessments.

We are fortunate to have detailed suggestions for developing effective performance assessments, whether the development effort is local or national (Airasian, 1991; Brualdi, 1998; Stiggins, 2002). For example, Roeber (1996) provides helpful advice on administrators' critical roles and responsibilities that contribute to an effective performance assessment program. These include the development of an assessment framework in concert with a detailed assessment plan, establishment of assessment resources, and creation of the assessment blueprint. Detailed guidelines for performance assessment development in relation to administrators' and teachers' roles and responsibilities are presented in Table 14. Any school or school district that sees the value in performance assessment and is planning performance assessments as part of classroom practice needs to examine each of these points carefully. Without attention to each, the best intended performance assessment program may fail.

Performance assessment and the activity it involves often represent a significant change for students, given the passive nature of many student–assessment relationships. Students must be involved and invested in performance assessments, as with portfolios and other active assessments. To do so, students must learn the assessment. The history performance assessments in Ivan's classroom present a set of responsibilities for his students. The responsibilities are framed by students' developing understanding that performance assessments are helpful and that they hold no surprise in

TABLE 14
Guidelines for Developing and Using Performance Assessments

- Determine why the performance assessment is needed.
- Specify the important reading knowledge and outcomes that will be assessed by the performance assessment.
- Propose a specific performance that is composed of the reading knowledge and outcomes and conduct a task analysis.
- Based on the task analysis and in relation to research findings, specify the performance that allows students to demonstrate their learning and achievement.
- Specify the aspects of performance that will receive assessment attention and enable the identification and determination of student success.
- Determine the degrees of students' performance that will be identified using rubrics.
- Set performance levels and different levels of proficiency in relation to instructional goals and standards.
- Use the performance assessment to evaluate student learning and work, perhaps using checklists geared to important (i.e., "must have") aspects of the task, using rating scales that represent the continuum of expected student performance. Note that piloting helps fine-tune scoring guides to best cover and represent what students might do.

Note. From *Understanding and Using Reading Assessment, K–12* (p. 108), by P. Afflerbach, 2007, Newark, DE: International Reading Association. Copyright 2007 by the International Reading Association.

how students and their work will be evaluated. The students know that they must attend to the detailed rubrics to guide their work, that this attention provides guidance on the detail of good performance. For example, the students know that superior performances in the history performance assessment will involve the use of different cues to determine source text status and the evaluation of the text in relation to this determination. The students are familiar with taking responsibility for knowing where they are in a task and how they are doing. They are comfortable with being in control of assessment, providing their own feedback to ongoing efforts. Ivan's students appreciate the control and agency that result from this involvement. They understand that responsibilities taken and met can bring tangible rewards.

Like students, parents need to pay attention to the special features of performance assessments, as these differ from the assessments with which parents are most familiar. They must learn the new performance assessment and understand its benefits. When parents are introduced to a new assessment and learn of its possible benefits to their children, they can become knowledgeable and strong supporters of the assessment (Shepard & Bliem, 1995). The development of appreciation for performance assessment and advocacy for it depend on the ability of the teacher and student to demonstrate its usefulness.

Reliability of Performance Assessment

The complexity of performance assessment creates a set of reliability issues. Teachers and students must attend to the special nature of student performances, work to develop consistent routines for evaluating performances, and be aware of the numerous types of confounds that are possible with performance assessments. These assessments usually do not restrict student answers to a single correct response. Rather, the assessments reflect the fact that student performance may be individualistic while still representative of important learning. Performance assessments often set the parameters for appropriate student performances and do not dictate a single acceptable response. When there is more than one acceptable answer, response, or performance, it is imperative that Ivan and his colleagues develop the ability to reliably score them. Here, teachers' knowledge of the parameters of acceptable performance must be well developed. The scorer of a performance assessment must be able to interpret students' varied responses to a single prompt, accommodate variations on a theme, and consistently evaluate these diverse responses in relation to a scoring rubric.

Students' learning of performance assessment is developmental, and we should expect, over time, the increased ability to use performance assessment rubrics to guide student learning. When students are expected to develop self-assessment strategies, we must consider the learning curve that is involved. Ivan and his students are fortunate that school district policy focuses on the consistent use of performance

assessments across grades and subjects. This contributes to the development of reliable assessment routines on the part of the students. They use rubrics and scoring guides in each subject, including history, science, and literature, to help them achieve and self-assess their performances. At the same time, teacher judgments of performance assessments provide regular and accurate information about student progress. A student's ability to judge his or her own work may be fledgling in September but better developed in May, nascent in kindergarten but maturing in fourth grade.

There are potential confounds in a performance assessment. Performance assessments, by definition, demand performances. We must be clear about what skills, strategies, and knowledge are involved in student performances so that we can anticipate how student work related to reading may influence the performance. For example, a performance assessment that requires students to write accounts based on their understanding of primary and secondary source texts introduces writing ability into the performance. A laboratory procedure that requires students to read instructions and then perform a series of measures using a balance beam involves both mathematical ability and fine motor skills. The creation of a diorama to portray what a student learns from reading about the culture of the Iroquois involves a creative, artistic component. Each of these performances not only demands successful reading but also engages other aspects of students' learning and ability. In each case, our ability to identify and anticipate the nonreading aspects of a performance assessment will provide the opportunity to evaluate students' performances in an accurate and useful manner and contribute to the reliability of the assessment. (For more on this issue, see the Reading Assessment Snapshot for Confounds in Reading Assessment at the end of Chapter 9.)

Validity of Performance Assessment

Effective performance assessment programs are the result of a careful alignment between research, standards, the curriculum, and assessment. Performance assessments offer the possibility of a rich representation of the construct and learning that we believe to be of value for our students. For example, when we determine that reasoning about the trustworthiness of a particular history text and its author is important, we must invest in assessment that accurately describes the effectiveness of our teaching and students' learning. A performance assessment allows us to create a situation in which students can demonstrate their learning in close relationship and alignment with the construct of reading as a historian. In contrast, we would be hard-pressed to make such inferences from most multiple-choice assessment items. The ability of the performance assessment to honor the construct of student learning provides the possibility of attaining high construct validity.

Earlier in this chapter, we examined rubrics, scoring guides, and sample performances. We determined that they can serve as tools of teaching, learning, and

assessment. The close proximity of teaching and learning to assessment is a hallmark of performance assessment. When Ivan's students use the rubric to guide their independent work and when he teaches critical history reading strategies in relation to the rubric, these important classroom routines anticipate the end-of-unit performance assessment. The assessment tasks that students engage in to demonstrate learning are the very tasks that they have been engaged in as part of the learning process. The performance assessment's rubric and sample performances serve as a scaffold and superstructure for learning, and when assessment is this involved in teaching and learning, ecological validity is high.

Summary

Performance assessments are especially effective in describing the manner in which students use reading within the content domains of school learning. Performance assessments reflect the complexity of learning and using knowledge in school and anticipate students' uses and applications of reading beyond school. These are two of the reasons for the keen interest in performance assessments shown by curricular and assessment reform efforts (CCSSO & NGA, 2010; PARCC, 2010; SBAC, 2010). Effective performance assessments reflect a clear and detailed understanding of the things that we would like students to learn and do. This understanding is established through consultation with relevant research and task analyses that promote detailed understanding of proposed student performances.

Performance assessment can promote excellence in teaching because it presumes our attention to important curricular goals and a detailed understanding of how to reach those goals. Performance assessments come with user's guides, or rubrics, which present students and teachers with clear goals and paths for attaining those goals. Rubrics provide a means for scoring performance assessments in a consistent manner in relation to students' complex learning. Rubrics also provide a means for making transparent our curricular and assessment goals. Teachers who are familiar with rubrics can use them to teach well because they help direct attention to important learning goals. Teachers also establish consistent and reliable scoring of student performances with rubrics.

The promise of performance assessment is accompanied by different challenges. Performance assessment demands teacher and student involvement in assessment that may be unprecedented. Although performance assessments have the potential to measure and describe student growth in complex reading tasks, they must be supported by substantial commitments of school resources. We must complement our understanding of how to do a performance assessment well with the sustained commitment of resources that promote the development and use of successful performance assessment programs.

ENHANCING YOUR UNDERSTANDING

1. In relation to your personal goals for students and state and district goals, identify important school performances that may be the focus of a performance assessment.
2. Conduct a task analysis of your proposed performance assessment. What is demanded of students? What types of reading behavior are expected? What related behaviors, such as writing, creating, and discussing, are also involved?
3. Develop a rubric that delineates at least three different levels of student achievement on a selected performance.
4. Create a lesson that focuses on how students can use performance assessment rubrics to learn to self-assess their reading and reading-related work.

Advocating for Reading Assessment

Advocacy of useful reading assessment is important because high-quality reading assessments do not always find their way into classrooms. The tradition of using particular reading assessments, such as paper-and-pencil tests, creates a level of familiarity and comfort for many educators. Assessments, including high-stakes tests, are traditional in nature and firmly established. This creates an environment in which reading assessment may be conducted by habit rather than by informed choice. Thus, the most appropriate reading assessment for measuring student and teacher accomplishments may be overlooked. When particular reading assessments have the potential of increasing the value of the information we obtain, we need to consider how to best advocate for such assessments.

The advocacy process presents an opportunity to educate those whose support is needed to implement assessment change: parents, administrators, and others in the school community who have the voice and energy to advocate for change. This process may be difficult when we deal with assessments and rationales for assessments that are firmly entrenched in tradition. However, our advocacy may be welcomed. Parents who learn that performance and portfolio assessments serve the dual purpose of describing their children's achievement while supporting it may be enthusiastic in their support of new assessments. Administrators who understand that a new reading inventory provides detailed, immediately useful information that helps to both support and describe student growth will support the reading inventory.

Detailed knowledge of the different types of assessment and clear communication of their strengths and weaknesses is required to advocate for a new reading assessment. Advocacy succeeds when we are able to describe how the advocated-for assessment represents an improvement over the existing assessments. Advocacy is a critical part of a thriving reading assessment program. The requirements of advocacy must be anticipated and included in any plan that seeks to engage the school community and change an existing program, implement a new program, or maintain a successful one.