

READING ASSESSMENT

Looking Ahead

Peter Afflerbach

Conceptualizing Effective Reading Assessment

In this article, I focus on three areas of reading assessment that I believe to be crucial for students' reading development: developing comprehensive formative assessments, assessing the wide array of factors that contribute to students' reading development, and fostering student independence by helping students learn to use reading assessment on their own. I also describe the consistent and negative effects of high-stakes testing on related reading assessment efforts.

When we assess, we make inferences about the nature of a student's reading from a sample of reading behavior. This is a fact common to all reading assessment, and it is worth spending some time with this idea. Pellegrino, Chudowsky, and Glaser (2001) propose that effective assessment is accomplished when we attend to three distinct but related components.

First, we must have a detailed model of the thing we are to assess. Fortunately, we have well-explicated models of reading that describe what students must have to read well. They also describe how these "must-haves" develop. For example, we know from research that phonics and motivation to read must be present for students to succeed. It follows that assessment of student reading should focus on these (and other) demonstrably key contributors to reading growth.

Second, when we determine what to assess, we must design assessment materials and procedures that yield valid and reliable assessment information. With phonics, we have a plethora of options for assessment. We can listen to a child read orally, noting miscues that represent needed work with phonics. We can provide children with flashcards that focus on specific phonics knowledge. We can analyze a child's writing to investigate phonics knowledge in productive language acts. We can administer a

phonics test. Although there are far fewer motivation assessments, we can better understand student motivation through classroom observations, and we can ask students to rate their motivations to read in general and in specific scenarios. We can administer a motivation survey or checklist.

The third step in effective assessment involves the inferences we make based on the assessment data. For example, we may conclude with confidence that a student who can identify five pictures representing words that include the *cl* consonant blend (e.g., *cliff*, *climb*, *clothes*) and who writes stories with a character named Clifford has a working knowledge of the *cl* blend. We can infer that a student who cannot wait to read each day in class is operating with motivation that differs from the student who avoids reading whenever possible.

The inferences we make from reading assessment results are, of course, limited to the types of assessments we use. The ongoing focus on readers' cognitive strategies and skills and on content area knowledge gain results in our ability to make legitimate inferences about these critical aspects of reading development and reading achievement. In contrast, if we do not regularly assess the development of students' motivation and self-efficacy for reading, we cannot make measurement-based inferences about the development of these critical factors.

An Assessment Credo

Reading assessment comes in many forms, and it is intended to serve different purposes and audiences. Whatever the form, audience, or purpose of assessment, I believe that the following credo is appropriate: *Assessment should produce information that is useful*

Peter Afflerbach is a professor of reading education at the University of Maryland, College Park, USA; e-mail afflo@umd.edu.

An assessment credo: "Assessment should produce information that is useful in helping students become better readers, and assessment should do no harm."

in helping students become better readers, and assessment should do no harm. When we use this credo, we can critically evaluate reading assessments and determine their value. We can determine the relative value of high-stakes tests and teachers' daily formative assessments. We can ascertain that our assessments cover all that is important in reading development, provide triangulation to support our inferences about student development, and do not over-assess particular areas of reading. This allows for consideration of questions such as the following: "Given the limited resources of our school and district, are we using the most effective combination of reading assessments?" "Does forcing students to participate in a testing situation in which their failure is guaranteed have negative outcomes for such students?" "Does reliance on an assessment that focuses exclusively on the cognitive aspects of reading development result in the diminution of attention and value given to equally important aspects of reading development, including motivation and self-efficacy?"

The Insidious Nature of Reading Tests

In this section, I set a context for discussing changes in reading assessment: I overview the stranglehold that high-stakes tests have on reading assessment resources and on reading curriculum and instruction in general. We do well to *not* underestimate the pervasive influence of

testing on reading instruction and on all aspects of reading assessment. Federal policy shapes reading assessment and reading instruction in the United States, and the Report of the National Reading Panel (NRP; National Institute of Child Health and Human Development, 2000) continues to shape federal policy. Five cognitive strategy and skill areas—phonemic awareness, phonics, fluency, vocabulary, and comprehension—were recommended by the NRP to be the focus of reading instruction because they "currently reflect the central issues in reading instruction and reading achievement" (p. 1-3). A key conclusion of the NRP report was that strategies and skills should be the focus of reading instruction and reading assessment.

Two things are striking about the NRP report. First, it does not attend to aspects of reading development that are vital for success, including students' motivation to read and students' self-efficacy. Second, the quoted statement anticipates the influence of testing on creating and maintaining reading programs in which cognitive strategy and skill are the overwhelming (if not only) focus.

It is worth a brief overview of the obvious, subtle influences that give reading tests a huge impact here. When deciding that there were five strategy and skill areas that should be the focus of reading instruction (and related assessment), the NRP considered only research that met specific criteria:

To be included in the database, studies had to measure reading as an outcome.

Reading was defined to include several behaviors such as the following: reading real words in isolation or in context, reading pseudowords that can be pronounced but have no meaning, reading text aloud or silently, and comprehending text that is read silently or orally. (National Institute of Child Health and Human Development, 2000, p. 5)

The "measure [of] reading as an outcome" was accomplished using tests. Reading research deemed acceptable by the NRP was designed to identify statistically significant differences that may derive from comparisons of treatment and control groups' test scores. These test scores focused exclusively on students' cognitive strategies and skills. Thus, consideration of readers and their development by the NRP was restricted to the nature of the outcomes that are assessed by tests: Students' attempts at pronouncing pseudowords and isolated real words were defined as "reading" and considered a reading outcome, but other necessary outcomes for successful reading, such as the development and maintenance of student readers' self-efficacy and motivation, were ignored.

Under No Child Left Behind (NCLB) and Reading First initiatives, funds were awarded to states when they promised to purchase reading instruction materials that were based on "scientific evidence" from reading research. Using reading programs that are based on proven instructional approaches and that address students' specific reading

"We do well to not underestimate the pervasive influence of testing on reading instruction."

needs is sensible. However, federal law ensured that the “scientific evidence” supporting effective reading programs derived from standardized reading test scores. In a manner similar to that used to identify acceptable research, reading programs were determined to be effective when they were based on research that found statistically significant differences between experimental treatment and control groups. The dependent variables in this research were reading test scores—a proxy for students’ cognitive strategy and skill use.

Test results are the basis for developing reading programs, and tests of cognitive strategy and skill continue to be the primary measure of schoolwide achievement. Students’ and schools’ adequate yearly progress (AYP) is determined each year in grades 3 through 8 using test scores. Again, these scores represent strategy and skill development as reading progress. There are no standardized test scores for changes in positive student reading affect, for growth in self-efficacy, or for a student’s turn toward intrinsic motivation to read.

To reiterate, test scores earn reading research and reading programs the labels *important* and *evidence-based*, respectively. Tests of reading strategies and skills are ingrained in our educational system—they are used to designate research as worthy, to prescribe the contents of reading curricula, to certify teacher and school accountability, and to measure student learning. The endemic nature of strategy and skill testing makes for a grand challenge if we are to argue for reducing the prevalence of testing or broadening testing’s focus. And this is the very argument that has to be made if reading assessment programs are to become more useful for teachers and students.

“Test scores earn reading research and reading programs the labels important and evidence-based, respectively.”

Formative Assessment and Summative Assessment

Formative and summative assessment must work together in a highly coordinated effort. Never has the need for this coordination of reading assessments been as great, as highlighted by the National Assessment of Educational Progress and the Common Core State Standards (CCSS). Our conceptualization of reading evolves; so too must our conceptualization of reading assessment. The National Assessment of Educational Progress (National Assessment Governing Board, 2013) defines reading as follows: 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)

This definition has significant implications for assessment. For example, when reading is conceptualized as using the meaning that is constructed, assessment must—to maintain construct validity—assess not only how students strategically construct meaning but also how they use the meanings that they construct from reading. At a minimum, this means augmenting the assessment of the “big five” of NCLB and Reading First (phonemic awareness, phonics, fluency, vocabulary, and comprehension) with a measure of what students do with the meaning they construct from reading, including

complex acts of applying, analyzing, critiquing, synthesizing, and evaluating.

The CCSS build on the definition of reading provided by the National Assessment of Educational Progress (National Assessment Governing Board, 2013) and require students to both construct meaning from text and use that constructed meaning. Examining one Standard provides a rich context in which to argue for the necessity of formative assessment to complement summative assessment. Consider RI.2.7, a second-grade English Language Arts Standard for reading informational text: “Explain how specific images (e.g., a diagram showing how a machine works) contribute to and clarify a text” (National Governors Association Center for Best Practices [NGA Center] & Council of Chief State School Officers [CCSSO], 2010, p. 13). If we conduct a task analysis of what a second-grade student must do to meet this Standard, we have the following:

- Effectively use the five strategies and skills of NCLB and Reading First (phonemic awareness, phonics, fluency, vocabulary, comprehension).
- Construct meaning from the text.
- Comprehend a related image, such as a photograph, drawing, or chart.
- Compare the two related understandings (text and image).
- Analyze the two for their separate and joint contributions to understanding.

“Assessment should describe whether students attain a particular Standard through summative assessment and, prior to this, how they are developing on this path to attainment using formative assessment.”

- Explain (through writing or speaking) how the two comprehended parts relate to each other.
- Describe how the image helps overall comprehension of the document.

A second-grade student’s attainment of this Standard is a notable accomplishment and should be demonstrated with an appropriate summative assessment. However, prior to any student taking the summative assessment, there must be a comprehensive series of related and detailed formative assessments. These formative assessments help us determine each student’s development in relation to these bullet points (e.g., Can Ellie provide a written description of how the text and photograph contribute different but related information to the overall understanding? Has Jerome sufficient phonics skills to decode unfamiliar words and construct a literal understanding of the included text? Will Ephraim be able to describe the unique information that is provided by the photograph?)

We certainly don’t want any students attempting the grand performance that is the summative assessment of RI.2.7 without continued checks that they are developing the collection of needed strategies and skills to the point that they can do so. The complexity of the reading text and task that epitomize this Standard demands that we investigate

students’ learning and progress along the way. Given the range of reading development that I have experienced in second-grade classrooms, I would want formative reading assessment that that helps me address students’ individual differences in all the areas implicated by the particular Standard.

I note that each and every Standard introduces levels of complexity in expected student outcomes that warrant assessment attention. Assessment should describe whether students attain a particular Standard through summative assessment and, prior to this, how they are developing on this path to attainment using formative assessment.

A further argument for the necessity of formative assessment can be made in relation to Vygotsky’s (1978) notion of zones of proximal development. Effective instruction within each student’s zone of proximal development depends on the teacher having up-to-date assessment information that describes what a student can do independently and what a student might next do with teacher support. Without formative assessment, critical decisions about what to teach and when to teach in the zone of proximal development are informed by guesswork. I note again that with the CCSS, students’ reading development demands teachers’ attention to multiple zones of proximal development simultaneously. For example,

planning effective instruction that helps students determine main ideas, critically evaluate the texts they read, and then produce a written commentary requires formative assessment information on three related but distinct zones of proximal development.

The example I have just discussed focuses on formative assessment of cognitive aspects of reading development. Equally important is formative assessment of affective aspects of reading. Just as we want to understand how fluency, vocabulary, and the ability to evaluate texts are developing, we want also to understand how (and if) a student’s self-efficacy is developing and how it affects reading. We want to know how student motivation develops and is maintained and whether it varies from reading situation to reading situation and from school subject to school subject.

Using Assessment to Help Students Develop as Independent Readers

A major goal of reading instruction is fostering the development of independent, successful student readers. To be independent and successful, all readers must assume responsibility for self-assessment: setting clear goals for reading, monitoring progress along the reading path, and determining if reading is successful. All independent readers self-assess, but self-assessment

“To be independent and successful, all readers must assume responsibility for self-assessment.”

is not always a focus of reading assessment. It must be.

When we view reading assessment as an important teacher task, we may gather information that informs our instructional decision making. When we view reading assessment as something to teach, we can help students move toward the goal of independence through self-assessment. We want to use assessment that helps shift students from an outward orientation, where there is dependence on the teacher for assessment feedback, to one that looks inward. As students adopt this self-assessment orientation, we must provide them with the tools to do the work. Here is where many reading assessment programs fall short: They may do a decent, even outstanding job of evaluating student progress, but assessment is done *to* or *for* students. Exceptional reading assessment programs help transform assessment so that it is done *with*, then *by* students.

Self-assessment in reading requires specific student mind-sets and strategies, and our instruction here must help students develop both. As for self-assessment mind-sets, we can encourage students to look first to themselves to answer questions that include “How am I doing?” “Is there a problem?” “What is it?” “How can I fix the problem?” and “How am I progressing toward my reading goal(s)?” Self-assessment strategies include comprehension monitoring, detecting problems, identifying problems, fixing problems, and getting back on track. They also include maintaining a focus on purpose and goal throughout acts of reading that are increasingly complex. Fortunately, self-assessment strategies (part of the larger group of metacognitive strategies) are close relatives of reading comprehension strategies and can be introduced, explained, and modeled in similar fashion. For

“Self-assessment in reading requires specific student mind-sets and strategies, and our instruction here must help students develop both.”

example, we can think aloud as we introduce a reading self-assessment checklist, and we can model how to use the checklist as we progress through a reading passage.

As students grow their ability to self-assess, benefits accrue. Independently undertaking, persisting at, and successfully completing reading tasks become possible. As well, students who are metacognitive are meta-affective: They understand the power that comes from successfully using self-assessment strategies, and this contributes to increased self-efficacy, which contributes to positive motivation to read.

Assessment of All Valuable Outcomes of Reading and All Factors That Contribute to Reading Development

Reading assessment in the United States is dominated by the measure of cognitive strategy and skill and the comprehension of text. If we were to construct an account of how students develop as readers by examining high-stakes test items, our description would be limited to students’ achievements in relation to phonics, phonemic awareness, fluency, vocabulary, and comprehension. While each of these makes a vital contribution to successful reading, they do not tell the full story of students’ reading development or reading achievement. Other important factors influence students’ reading development and should be assessed (Afflerbach, 2016).

Successful readers are independent—they monitor their meaning making, they approach reading with an attitude to achieve, and they are motivated and engaged. They attribute their reading outcomes to appropriate causes, such as effort and planning. Successful readers are also highly efficacious—they believe that they can succeed at most reading tasks. If these characteristics were of secondary importance to reading success and development, we might be able to explain their absence on each and every statewide test of reading. But they aren’t of secondary importance. Motivation, self-efficacy, and making correct attributions for reading performance are essential for reading success.

Let’s examine self-efficacy in some depth. Consider the following quote while thinking about your students who find reading challenging:

Among the mechanisms of human agency, none is more central or pervasive than beliefs of personal efficacy. Whatever other factors serve as guides and motivators, they are rooted in the core belief that one has the power to produce desired effects; otherwise one has little incentive to act or to persevere in the face of difficulties. Self-efficacy beliefs regulate human functioning through cognitive, motivational, affective, and decisional processes. (Bandura & Locke, 2003, p. 87)

Bandura and Locke (2003) describe the power of self-efficacy and belief: Students who don’t believe they will succeed have little or no reason to invest

“When our teaching helps students develop and maintain personal efficacy, it is an educational outcome worth noting and worth assessing.”

time and effort in their reading. Our struggling readers often lack of self-efficacy, disengage from instruction, and give up when facing challenges. This is fairly predictable human behavior: Why participate in activities where there is the belief that failure will occur and where the individual is powerless to change such foregone conclusions?

Recent research demonstrates the consequences of self-efficacy. Students' increased self-efficacy is related to enhanced reading comprehension and achievement (Solheim, 2011). It is also related to increases in motivation (McCrudden, Perkins, & Putney, 2005). High-achieving students have high self-efficacy, and they make fewer attributions for their performance to external causes, including luck and task difficulty (Shell, Colvin, & Bruning, 1995).

When our teaching helps students develop and maintain personal efficacy, it is an educational outcome worth noting and worth assessing. We are not wanting for research that demonstrates the power and necessity of positive affect and self-efficacy in acts of reading. They are part and parcel of reading development. That they have been ignored from the assessment perspective for so long needs to change. Reading assessment should focus not only on strategy and skill but also on dispositions that either support or work against students' reading development.

Note also that powerful influences—including self-efficacy, motivation and

engagement, and attributions for performance—are dynamic. For example, self-efficacy will influence a student's approach to and performance on a particular reading text and task. In turn, a positive or negative experience with the particular text and task can further influence self-efficacy. Like reading strategies and skills, students develop affective stances toward their reading. Our assessments must be sensitive to these developments and provide information that describes the status of self-efficacy, the nature of a student's motivation, and the types of attribution that a student makes for reading success or failure. With such assessment information, we can create reading environments and instruction that foster healthy development of students' cognition and affect.

Professional Development in Support of Teachers' Assessment Expertise

The ideas about effective reading assessment I have so far sketched are inextricably linked to teacher expertise with reading assessment. Conducting formative assessment and teaching students how to “do” assessment for themselves demands that teachers are facile with reading assessment materials and procedures. Using assessment to evaluate the broad array of cognitive and affective factors that influence students' reading development also depends on teachers who are assessment experts (Johnston, 1987).

Too often, the assessment of reading may be given cursory attention in reading methods courses, and this does not bode well for beginning teachers who are interested in changing reading assessment culture. Veteran teachers are all too familiar with school resources that are gobbled up by testing. The amount of time and effort that is required to help teachers become experts at administering, interpreting, and using valuable assessments such as Running Records, informal reading inventories, and detailed questioning routines hints at the tasks ahead.

If accomplished teaching includes identification of students' varied zones of proximal development, then a first-order need for professional development is helping teachers use the skills and tools that are central to this identification. For example, informal reading inventories can provide a wealth of information about a student's reading development: phonemic awareness, sight word knowledge, application of phonics, fluency, literal and inferential comprehension. While this menu is limited to cognitive aspects of student reading development, it illustrates the level of detail

“Reading assessment should focus not only on strategy and skill but also on dispositions that either support or work against students' reading development.”

that assessment can describe when the assessment materials and procedures are combined with teacher expertise.

Unfortunately, the resources given to particular reading assessment initiatives and programs cannot be given to others. Economists call this *opportunity cost*, and it is defined as

the added cost of using resources (as for production or speculative investment) that is the difference between the actual value resulting from such use and that of an alternative (as another use of the same resources or an investment of equal risk but greater return). (“Opportunity cost,” n.d.)

It is worth considering the opportunity costs invoked when there is massive expenditure on high-stakes reading tests: purchasing the tests, practicing for them, administering them, scoring them, and reporting scores. A recent estimate of the costs of testing in United States schools determined that “states nationwide spend upwards of roughly \$1.7 billion on assessments each year” (Chingos, 2012, p. 1). I note that the use of *assessments* in the quoted sentence is interchangeable with *tests* and that the report predates the assessments required by the CCSS. That the amount spent on tests has risen since 2012 is most probable.

What is the opportunity cost involved in spending such sums on testing and not spending it to support teachers’ professional development in assessment (reading and otherwise)? I don’t see the means by which consequential

formative assessment will be conducted, with which teachers will be able to address students’ diverse individual differences without this formative assessment.

Conclusions

Reading assessment is at a crossroads, and significant numbers of teachers, parents, students, administrators, and other concerned citizens are questioning the assessment status quo. This status quo is one in which massive amounts of school resources are given to high-stakes tests—tests that offer precious little data that can help teachers understand students’ individual differences and needs and inform instruction. Should the stranglehold of testing be eased, there are clear needs for other reading assessments. Formative assessments are necessary to help teachers and students move toward attainment of increasingly complex standards embodied in summative assessments. Students should be learning to do assessment as they continue to learn to read— independence with assessment is the hallmark of a successful reader. Finally, assessment must describe the full range of students’ reading development—the cognitive and the affective—so that we fully appreciate the accomplishments of students and teachers.

REFERENCES

- Afflerbach, P. (2012). *Understanding and using reading assessment, K–12* (2nd ed.). Newark, DE: International Reading Association.
- Afflerbach, P. (2016). *Handbook of individual differences in reading: Reader, text, and context*. New York: Routledge.
- Bandura, A., & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*, 87–99.
- Chingos, M.M. (2012). *Strength in numbers: State spending on K–12 assessment systems*. Washington, DC: Brown Center on Education Policy, Brookings Institution.
- Johnston, P. (1987). Teachers as evaluation experts. *The Reading Teacher, 40*(8), 744–748.
- McCrudden, M., Perkins, P., & Putney, L. (2005). Self-efficacy and interest in the use of reading strategies. *Journal of Research in Childhood Education, 20*(2), 119–131.
- National Assessment Governing Board. (2013). *Reading framework for the 2013 National Assessment of Educational Progress*. Washington, DC: U.S. Government Printing Office.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.
- 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. Retrieved from <https://www.nichd.nih.gov/publications/pubs/nrp/pages/smallbook.aspx>
- Opportunity cost. (n.d.). In *Merriam-Webster’s online dictionary* (11th ed.). Retrieved from www.merriam-webster.com/dictionary/opportunity%20cost
- Pellegrino, J., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- Shell, D., Colvin, C., & Bruning, R. (1995). Self-efficacy, attributions, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology, 87*(3), 386–398.
- Solheim, J. (2011). The impact of reading self-efficacy and task value on reading comprehension scores in different item formats. *Reading Psychology, 32*(1), 1–27.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. & Trans.). Cambridge, MA: Harvard University Press.