



## Understanding Assessment: Putting Together the Puzzle

SHEILA W. VALENCIA

UNIVERSITY OF WASHINGTON, SEATTLE

*"No one test or assessment should be asked to serve all the assessment purposes. We need, at this point, a system made up of articulated components, glued together by their adherence to content standards and serving explicit purposes for assessment."*

—National Council for Education  
Standards and Testing, 1992

As far back as 1992, when the standards movement was launched, the National Council for Education Standards and Testing, a team of experts in education, assessment and policy, reminded our nation of the different purposes for assessment—from public accountability to creating individualizing instructional plans for children. They also cautioned us to use multiple measures to fit these purposes including standardized tests, running records, informal reading inventories, classroom projects, portfolios, writing samples, debates, literature circle

discussions, and more. Now, 10 years later, we are hearing the same reminder (Brennan, Kim, Wenz-Gross, Siperstein, 2001; Herman, 2001; International Reading Association, 1999). In fact, in July 2001 members of the National Education Association, the nation's largest teacher's union, endorsed a policy calling for a combination of standardized tests and other assessment tools such as teacher designed assessments when making important educational decisions (Blair & Archer, 2001).

Assessment has always been a part of the educational landscape. However, because assessment can serve so many different purposes and can come in so many different forms, it has been confusing and, sometimes, it has been the subject of contentious debate. Unfortunately, as a result, many of us have come to view assessment as a necessary evil, a requirement rather than a helpful part of instruction. But assessment IS a critical part of instruction and it CAN be useful if we understand the pieces of the puzzle.

# A Balanced Approach to Assessment

A balanced assessment system consists of three parts (see Figure 1): standard assessment, classroom-based assessment, and student self-assessment.

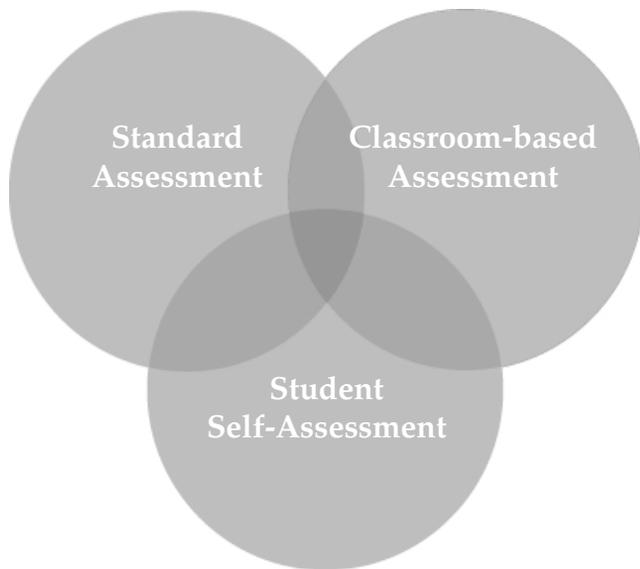


Figure 1: A Balanced Assessment System

Each of these parts serves a different purpose and each is a different and important piece of the puzzle. For example, assessments are used to qualify students for special services; to report to school boards, states, and parents; to evaluate program effectiveness; to monitor student learning and adjust teaching strategies; to evaluate students' growth over time; to engage students in self-assessment; and to understand students' strengths and needs. Each of these different purposes and audiences may require different kinds of assessments and different types of information (Haney, 1991; National Research Council, 1999; Office of Technology Assessment, 1992; Pearson & Valencia, 1987). One type of assessment cannot meet the needs of all the audiences. State departments of education and administrators, for example, will want to know about school programs or the per-

formance of large numbers of students. They generally need assessments that can be easily administered, and they rarely need information more than once a year. They do not need information about individual students' strengths and needs. On the other hand, teachers, parents, and students need more specific information geared to what students are learning, and they need that information more often.

The glue that holds these assessment pieces together is content standards—statements about what students should know and be able to do. Content standards grew out of the concern that students in different parts of the country and in different schools were not being held to the same high standards nor were they being provided with the same quality instruction. At the same time, educators were concerned that what students should learn was often indirectly determined by what appeared on standardized tests. Instead of defining what students should know in terms of a test score, educators wanted to be sure that students were assessed on important and worthwhile learning goals that were a part of classroom instruction (Wiggins, 1998). In the past several years, teams of teachers have participated in many state and local efforts to determine content standards for student learning, and to develop or review assessments that align with the standards.

Standards answer the question, What do we want students to know and be able to do? Next, we must ask, How well are students progressing toward those standards? To answer this question, we must look at the pieces of the assessment system. Each piece has strengths and weaknesses; if we understand what each type of assessment can and cannot contribute to our understanding of student learning, we can do a better job of teaching as well as assessing students (Hiebert & Calfee, 1989; Pearson & Valencia, 1987).

## Standard Assessment

Standard assessment is the term we use here to refer to assessments that are given to all students in a state, school district, or school. We used to think

of these simply as norm-referenced tests, such as the Stanford Achievement Test or the Gates-MacGinitie Reading Test, that compare student performance to the performance of a national sample of students at the same grade. But these days, many states have constructed their own tests, such as the ISAT in Illinois, the FCAT in Florida, or the direct writing assessment in California, that are not norm-referenced. Instead, they have used criteria for student performance set by the state. Regardless of whether a state uses a norm-referenced test or a state-developed test, these standard assessments are designed to evaluate students in a uniform, systematic way against some established standard. All 50 states now administer some type of standard assessment to students at targeted grade levels (Orlofsky & Olson, 2001). These assessments are particularly useful to state legislators, school boards, and administrators because they are easily administered to large numbers of students, and being consistently administered and scored across sites adds a level of reliability that may not be possible on other assessments.

Many current standard assessments have begun to include more authentic literacy tasks, such as reading longer, more complex selections, answering both multiple choice and open-ended questions, making connections across different texts, and using the writing process to persuade or to describe an important event. In fact, more than 90% of the states now require students to write extended responses on their standard assessments (Education Week, 2001). Even the Graduate Record Exam, the standard assessment used for admission to masters and doctoral programs in education, recently moved beyond multiple choice items by adding two essays to the exam (Gonzales, 2001)! It is difficult, however, for standard assessment to be as authentic or performance-based as many classroom-based assessments—there are simply too many constraints (i.e. cost, time, scoring) on this kind of testing. Nevertheless, the changes that have been made DO represent important shifts from standardized tests of the past. They signal closer alignment with our current visions of reading and writing performance, and they suggest that we can help prepare students

for these tests by engaging in good quality teaching. So, for example, it would be both good teaching and good preparation to have students read, discuss, and write about the connections between two texts or to provide students with models and opportunities to write persuasive pieces. At the same time, it is important to prepare students for the format of the standard tests they will take. They should be familiar with the kinds of tasks, questions, and answer sheets they will face.

There are several cautions to keep in mind related to standard assessments. First, although familiarizing students with the test format is important, studies indicate that if students only practice with the format of a test, they are less likely to actually learn (Linn, 2000; Popham, 1999). Preparing students to do well involves more than test preparation; it involves helping students learn to apply important reading and writing strategies. In fact, several recent studies reveal that if test preparation is focused only on a specific standardized reading test, students' scores drop dramatically when they are given a different standardized reading test assessing similar reading abilities. In other words, students in these studies had learned only to do well on a particular test; they hadn't learned reading skills that transferred to other situations. Similarly, if teachers focus instruction on only what is tested, there is evidence that the curriculum narrows, eliminating other important learning from students' experiences (Herman & Golan, 1993; Smith, 1991).

A second caution relates to interpreting test scores. Standard assessments, by their very nature, are not precise, but rather rough approximations of student performance (Popham, 1999). Furthermore, they are not good measures of students who are performing substantially below or above their grade placement. For example, most of the reading selections on a standard reading assessment designed for fourth-grade students would be at the third, fourth or fifth-grade level. It would be unlikely for such a test to include passages at first or second grade level or at eighth or ninth grade level. As a result, students who are reading at these levels will

be unable to demonstrate their abilities on the standard grade-level test. What's more, even if these students were to make gains from September to June, it would be very difficult to show gains using these tests—they simply don't have enough items at the lower and higher levels.

Similarly, grade equivalent scores are often misunderstood. For example, a student score of 8.6 on a third-grade test does not mean that student can read as well as an eighth-grader nor does it mean she can read eighth-grade level material. As we have explained above, there is unlikely to be eighth-grade material on this test nor were eighth-grade students included in the norming sample. Simply, this 8.6 means that the student reads substantially better than other students her age who took this test; precisely how well she reads can't be determined from this test. For these reasons and others, the International Reading Association has taken a strong stand against the use of grade equivalent scores (International Reading Association, 1982).

Finally, recent surveys have documented that only 10 of 50 states provide teachers or students with feedback on how individual students perform on particular test items found on standard assessments (Education Week, 2001). They simply provide overall scores, and often they provide those scores after students have moved on to another grade and another teacher. Boser (2001) concludes that "states rarely provide feedback needed for teachers and students to learn from their mistakes." So, the lack of specific and timely feedback makes it unlikely that teachers or students could use the results to direct future learning.

In summary, standard assessments provide important systematic information about student learning in relation to other students or to a pre-established standard of performance. This information is particularly useful to people outside of the classroom such as legislators and administrators. Standard assessments work much like a thermometer, taking students' temperature to evaluate their academic health or abilities. However, thermometers don't help us know exactly what is

causing our illness or how to get better. For that we need finer-grained assessments and well-trained physicians. That's where classroom-based assessment comes in.

## Classroom-based Assessment

In the past several years, classroom-based assessment has enjoyed renewed support from policymakers (i.e. National Research Council, 1999), assessment experts (Shepard, 1999) and teachers (International Reading Association, 1999) alike, giving it a central position in all assessment discussions. Recent studies suggest that teachers, themselves, are the most important assessment tool. This makes perfect sense when you realize that teachers spend 1/3 to 1/2 of their classroom time in assessment-related activities (Stiggins & Conklin, 1992) and that they make decisions about what and how to teach approximately every 2-3 minutes (Shavelson & Stern, 1981)! Teachers must be able to develop assessment strategies, gather evidence, analyze what they see, and ultimately, make instructional adjustments to respond to student needs. This is precisely why classroom assessment is so powerful. Classroom-based assessment is conducted close to actual learning and to children; as a result, it is most likely to be aligned with instruction, provide immediate feedback to teachers and students, engage students in assessment of their learning, and influence instructional decisions. Classroom assessment also occurs more frequently than standard or norm-referenced testing, and it can be more precisely tailored to individual children and to instruction. With classroom-based assessment, assessment and instruction are melded. Both teachers and students become learners. Teachers become more focused on what and how to teach, and students become more self-directed, motivated, and focused on learning (Graue, 1993; Wolf, 1989).

Classroom-based assessment includes a wide range of tools and strategies. Because the assessments grow out of actual classroom activities, they are more likely to resemble authentic reading and writing and they tend to be more performance-based than stan-

standard assessments (Hiebert, Valencia & Afflerbach, 1994; Wiggins, 1993). For example, students might demonstrate their literacy abilities by conducting research and writing a report, developing a character analysis, debating a character's motives, dramatizing a favorite story, drawing and writing about a nonfiction piece, or reading aloud and discussing a portion of text with the teacher. These assessments can range from relatively short assessments to long-term projects. They often require students to apply their skills and strategies to new reading and writing tasks, and they often value the thinking behind work—the process—as much as the finished product (Pearson & Valencia, 1987; Wiggins, 1989; Wolf, 1989).

Some classroom assessments are somewhat formal; others are more informal. Those that are more formal provide teachers with a systematic way to evaluate how well students are progressing. For example, after completing a 4-6 week theme, teachers will want to know how well students have learned the theme skills and concepts. They may give all the students a **theme test** in which students read, answer questions, and write about a similar theme concept, and in which they apply the skills and strategies taught in the theme to a new reading and writing task. This type of assessment allows the teacher to systematically evaluate all the students on the important skills in the theme and to do so using reading and writing experiences that fit with the instruction. In other situations, or for specific students, teachers might use a **skills test** to examine specific skills or strategies taught in the theme. Finally, teachers may want to know how well students are reading and writing at the beginning and end of the year or how they have progressed compared with other children at the same grade level. Two points of comparison would be available—the student's growth over time, and the student's performance as compared with his grade-level peers. For this type of assessment, teachers could use **leveled reading passages** or **benchmark tests** that have been carefully calibrated. Because these types of classroom assessment are more flexible than standard assessments, teachers can

select the assessments that best fit the needs of individual students, choosing the appropriate level of difficulty and observing students as they work to modify the assessments as needed.

By far, the majority of classroom-based assessments are more informal, including evidence such as **anecdotal notes** on students' participation in literature discussion groups, samples of **student writing**, or artifacts from classroom **projects**. Some informal assessments may be drawn from typical classroom activities such as practice book pages, journals, essays, reports, or reading logs. Other times, it will be difficult to show student progress using actual work so teachers will need to keep **notes** or **checklists** to record their observations from student-teacher conferences or informal classroom interactions. Sometimes informal assessment is as simple as stopping during instruction to observe students, use **diagnostic checkpoints**, or discuss with the students how learning is progressing. Any of these types of assessment can be made more formal by specifying guidelines for how to do them, or they can be quite informal, letting students and teachers adjust to individual needs. In some situations, the teacher will want all students to complete the same assessments; in others, assessments will be tailored to individual needs.

Just because assessment is conducted in the classroom, doesn't make it good assessment. Research suggests that classroom assessment must have three critical features. First, and most obvious, it must be aligned with instruction. Although this seems obvious, teachers sometimes inadvertently hold students accountable for things they haven't adequately taught or students haven't adequately practiced (Valencia, 1998). At the same time, teachers must make decisions about the most important things to assess rather than treating all learning as equally valuable (Wiggins, 1989). There is nothing worse than collecting lots of information that you don't use or that targets unimportant learning. Classroom time is too precious to waste. So, teachers must be strategic and focused as they implement classroom assessment. State and district curriculum

guides, published instructional materials, national standards documents, and professional colleagues are good resources for determining important learning outcomes for students (Education Week, 2001; Valencia & Place, 1994).

Second, classroom assessment needs to be ongoing. This implies that teachers must continually re-evaluate student learning and then use that information to adjust instruction. Evidence from several studies suggests that when classroom assessment is ongoing, student achievement on standardized assessments actually improves (Black & Wiliam, 1998)! The best explanation of this finding is that when teachers are constantly attending to what students are learning, documenting their assessments, and then adjusting their instruction based on the assessments, their instruction is more focused and thus student achievement improves. So it isn't simply the administration of classroom assessments that improves student learning, it is that teachers actually use assessment evidence to continually re-focus and adjust instruction.

Finally, good classroom assessment must rely on a variety of forms of assessment. For some students, written work is difficult, so too much reliance on written work will put them at a disadvantage (Jenkins, Johnson & Hileman, 2000). Similarly, particular activities or topics will inspire excellent performance in some students and frustrate others. And, work supported by teachers or completed collaboratively with peers may give a different impression of students' capabilities than work completed independently. Including a variety of types of assessments over time will insure that students are provided with ample opportunities to demonstrate their abilities and that teachers' conclusions are well-founded.

In sum, classroom-based assessment has the advantage of growing out of classroom instruction and the flexibility to adjust to individual student and teacher needs. Preparing students for classroom assessments is not separate from good instruction; it IS good instruction. Because teachers are in charge of both assessment and instruction, there is greater

likelihood that assessment results actually will be used for instructional decisions. And, as a result, students are more likely to learn. However, the potential of classroom assessment will not be realized unless we are systematic about what and how we assess, and unless we use the information well.

## Student Self-Assessment

Student self-assessment may seem like an extravagant addition to the assessment system. However, both scholars and classroom experience suggest that it is an important piece of the puzzle. Students who are engaged in self-assessment do not become dependent on teachers to determine how well they are doing or where they need more work (Reif, 1990). They see learning as within their control and gain a sense of responsibility and ownership. They move from passive learners (Johnston & Winograd, 1985), unengaged and uninspired, to active learners. As a result, these students become more focused on their work. They learn the qualities of good work, how to judge their work against those qualities, and how to assess their own efforts and feelings of accomplishment (Reif, 1990; Wolf, 1989). They are more likely to set goals and to accomplish them, and consequently, their learning improves (Andrade, 2000; Stiggins, 1997). These students are also more likely to share common goals and expectations with their teachers (Valencia, 1998). That means that teachers and students can work together, rather than at cross-purposes, because they have a shared understanding of what they want to accomplish. Anecdotal evidence from classroom teachers suggests that when students and teachers share goals and values, there is less conflict over grades and better communication among teacher, parent, and student. Naturally, this leads to improved learning.

Self-assessment can take several forms. First, students can use samples of their work as a springboard for self-assessment. For example, they can reflect on their reading preferences and habits by reviewing their reading journals; they can judge the quality of their research report by analyzing the content, voice, and writing conventions; they can

evaluate their progress over time by comparing similar work samples produced at the beginning, middle, and end of the school year; they can even engage in self-assessment as part of a more formal assessment such as a theme test. Relying on classroom artifacts is helpful because students can literally see and take time to carefully study their work. In the process, students are encouraged to think about what went into that work and what strategies seemed most useful or problematic. Having the actual artifact also makes it easier to think about work that was completed in the past.

A second strategy for engaging students in self-assessment is to make self-assessment a natural part of classroom conversation. All too often, self-assessment takes the form of written reflections rather than discussion. Although written reflections are useful and can encourage a bit more reflection time, discussions help students become part of a reflective community whose members are willing and able to talk about their strengths and needs. This is an important “habit of mind” (Meier, 1995) for any learner. For example, after reading a story, students can discuss what parts they found personally engaging and what sections were particularly difficult or confusing for them. They can also help one another develop strategies for clarifying anything that confused them. Other times, it is helpful to think about work while in the midst of doing it. For example, while in the process of creating a piece of writing, students can step back from writing and discuss how the process is going and what they are learning about themselves as writers. Or while students are reading, teachers can stop them briefly to discuss how they are understanding or have them use sticky notes to mark spots of confusion.

A final strategy for engaging students in self-evaluation is to involve them in developing and using rubrics or criteria for their work (Andrade, 2000, Spandel, 1996; Valencia, 1998). For example, process writing instruction can be more effective if students have models and if the criteria for good work are presented using student language. Such

models give students an idea of how to go about their work and what a high quality product will look like when they are finished. Too often, these criteria become clear only at the end of a project, as students see others’ work or receive a grade. When teachers work collaboratively with students to develop criteria, students are more likely to understand the value of self-evaluation and transfer this understanding to other learning situations.

We caution that self-assessment can sometimes overemphasize superficial aspects of students’ work (i.e. handwriting, drawings), efforts (i.e. I worked hard), or unexamined feelings (i.e. I like it, it’s good). In fact, studies suggest that without support to go beyond the superficial, students tend not to develop a more reflective and analytic stance toward their learning (Valencia, 1998). A related caution is that self-assessment can easily become routine and uninspired if it is overused or used in the same way regardless of the kind of work. Students can grow as weary of self-assessment as any mundane activity. The antidote for such problems is to provide instruction in self-assessment (modeling, guidance, practice), time (self-assessment cannot be rushed), and many opportunities for students to discuss insights about their own learning. Like any skill or strategy, self-assessment needs support to develop.

# How Can Teachers Become More Effective at Balanced Assessment?

Compared to a jigsaw puzzle, it may not seem difficult to put together only three pieces of the assessment puzzle. But, it is. On one hand, we must struggle continually to overcome historic tradition and the current inclination to rely on a single, standard score. As educators, we need to counter the illusion of a simple score and the almost exclusive confidence those outside education ascribe to standard assessments. On the other hand, we must learn to deal with multiple indicators (i.e. indicators from standard, classroom-based, and student self-assessment as well as multiple indicators within each type of assessment). Sometimes information from multiple sources will converge, providing a consistent evaluation of student performance; other times the information may be discrepant because of differences in assessment formats, the skills and strategies tested, or simply inconsistencies in student learning. But, it is important to value all the information and to remember that the more samples of student learning we collect, the more trustworthy and informative our results.

We close with several suggestions for implementing the pieces of a balanced assessment system.

- When making important educational decisions or sharing information with parents, be sure to use information from all three pieces – standard assessment, classroom assessment, and student self-assessment. Be sure you understand the purpose and focus of each assessment as well as the strengths and limitations of each. Help parents understand them as well. Use the information to describe students' strengths and needs, rather than to label them with a grade or a number, and to plan for instruction.
- Focus assessment on the most important outcomes in the curriculum. Although teachers informally assess every time they interact with students and every time students work on an activity, you do not have to document every interaction or every lesson. Daily lessons and

activities are often building blocks to more complex goals. Determine the most important goals you have for each unit. Then select a couple of artifacts or focal points for your anecdotal notes or checklists. Use these assessments judiciously depending on the situation, your goals, and the particular students. Collecting too much information is as problematic as not collecting enough.

- “Front-load” instruction. This simply means that you should be clear about the goals of instruction and make those explicit to the students. For example, if students are going to read about environmental issues and be asked to take a position, they will need to learn how to distinguish fact from opinion, synthesize information, and draw conclusions. Both you and the students have a better chance of achieving your goals if you make clear to them the relationship between the skills they are learning and the task they are completing.
- Help students understand what good reading and writing look like by providing them with examples, examining work together, and discussing criteria. For example, help the class develop criteria for a good research report or book talk and then have children evaluate their work according to the criteria. Use criteria and scoring rubrics provided with instructional materials WITH the children instead of using them just for grading.
- Make self-assessment a dependable, integral part of your classroom. Begin with non-academic activities such as judging how well the class is working in groups, or how sharing time is working, or discussing favorite artwork. Some of these activities require students to consider qualities of good performance; others require judgments based on personal criteria. Both, however, require students to step back from their work or their behavior to think reflectively about it. You will need to develop these abilities over time with your students.
- When assessing growth over time, be sure to consider both the assessment task and the individual student. For example, if you want to assess students' ability to read and summarize, the difficulty of the text and the type of text will be important to consider. A student's summary of a second-grade text at the beginning of the year may be better than his summary of a fifth-grade text at the end of the year, but the change in

difficulty level would signal growth. Similarly, the topic or text type (narrative vs. information) will influence the quality of students' summaries. Multiple measures are especially important when assessing growth.

- Create a system to help you keep track of the assessment information. Some teachers use a 3-ring notebook with a section for each student, others use a computerized system, and others use a combination of work folders/ portfolios and teacher records. Whatever system you use, be sure to keep samples of student work and to document your assessments. These samples will help you communicate with both parents and students, providing the evidentiary trail of students' learning. If collections of work are set up collaboratively with students, they provide an excellent vehicle for conversation about students' strengths, needs, and future goals. In addition, by looking across students' folders, you will be able to analyze your instruction. It will become obvious, for example, the kinds of activities on which students are spending most of their time and areas in which they need more support.
- Use classroom assessments to help with grading. You do not need to grade every piece of work or every assessment. The evidence you collect will provide the basis for the grades you assign. Some of the more formal assessments such as tests, performance activities, and projects are easier to grade. Other assessments such as oral discussions, response journals and rough drafts of writing are more difficult to grade but still provide useful information. Together, these graded and ungraded artifacts provide strong evidence for your grading decisions.
- Begin classroom assessment slowly. Make good use of assessments that come with your instructional program or assessments you already have in place. You don't need to develop everything from scratch. Begin with several important outcomes, take time to review assessment results, and then use those results to shape your instruction. Classroom assessment is the piece of the puzzle that will most influence your teaching and student achievement. It takes time, but it is time well spent.

## Final Thoughts

The goal of assessment is to improve learning and teaching. In truth, we cannot be good teachers nor can students be effective learners unless we use evidence to guide instruction and learning. By understanding the different types of assessments and using them wisely, we are more likely to make those good decisions. When the pieces of the puzzle are fit together, they provide a clear picture of learning and a road map to success.

## References

- Andrade, H.G. (2000). Using rubrics to promote thinking and learning. Educational Leadership, 57, 13-18.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80, 139-148.
- Blair, J. & Archer, J. (2001). NEA members denounce high-stakes testing. Available: [http://www.edweek.org/ew/ewstory.cfm?slug=42neatest\\_web.h20&keywords=National%20Education%20Association](http://www.edweek.org/ew/ewstory.cfm?slug=42neatest_web.h20&keywords=National%20Education%20Association)
- Boser, U. (2001). Pressure without support. A better balance: Standards, tests, and the tools to succeed. Education Week, 20(17), 68-71, Jan. 11, 2001.
- Brennan, R.T., Kim, J., Wenz-Gross, M., & Siperstein, G.N. (2001). The relative equitability of high-stakes testing versus teacher-assigned grades: An analysis of the Massachusetts Comprehensive Assessment System (MCAS). Harvard Educational Review, 71, 173-216.
- Education Week (2001). A Better Balance: Standards, Tests, and the Tools to Succeed. 20(17), Jan 11, 2001.
- Gonzalez, K. (2001). GRE General test to change format in October 2002. Press release, Educational Testing Service. June 27, 2001, Available: [gonzalez@ets.org](mailto:gonzalez@ets.org)
- Graue, M. E. (1993). Integrating theory and practice through instructional assessment. Educational Assessment, 1(4), 283-310.
- Haney, W. (1991). We must take care: Fitting assessments to functions. In V. Perrone (Ed.), Expanding student assessment. Alexandria, VA: Association for Supervision and Curriculum Development.
- Herman, J.L. (2001). Accountability bottom up. The CRESST Line. Newsletter of the National Center for Research on Evaluation, Standards, and Student Testing. Los Angeles: UCLA Center for the Study of Evaluation.
- Herman, J. L. & Golan, S. (1993). The effects of standardized testing on teaching and schools. Educational Measurement: Issues and Practice, 12(4), 20-25, 41-42.
- Hiebert, E. H., & Calfee, R. C. (1989). Advancing academic literacy through teachers' assessments. Educational Leadership, 46(7), 50-54.
- Hiebert, E. H., Valencia, S. W., & Afflerbach, P. P. (1994). Understand authentic reading assessment: Definitions and perspectives. In S. W. Valencia, E. H. Hiebert, & P. P. Afflerbach (Eds.), Authentic reading assessment: Practices and possibilities, (pp. 6-21). Newark, DE: International Reading Association.
- International Reading Association. (1999). High-stakes assessments in reading. A position statement. Available [http://www.reading.org/positions/high\\_stakes.html](http://www.reading.org/positions/high_stakes.html)
- International Reading Association. (1982). Misuse of grade equivalents. The Reading Teacher, 35(4) 464.
- Jenkins, J.R., Johnson, E., & Hileman, J. (2000). Sources of individual differences on the new reading performance assessments. unpublished manuscript.
- Johnston, P.H. & Winograd, P.N. (1985). Passive failure in reading. Journal of Reading Behavior, 17(4), 279-301.
- Linn, R. L. (2000). Assessments and accountability. Educational Researcher, 29(2), 4-16.
- Meier, D. (1995). The power of their ideas : Lessons for America from a small school in Harlem. Boston : Beacon Press.
- National Council on Education Standards and Testing. (1992). Raising standards for American education. Washington, DC: U.S. Government Printing Office.

- National Research Council (1999). Testing, teaching, and learning: A guide for states and school districts. Committee on Title I Testing and Assessment, R.F. Elmore & R. Rothstein (Eds.). Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Office of Technology Assessment. (1992). Testing in American schools: Asking the right questions. Government printing office, Washington DC.
- Orlofsky, G.F. & Olson, L. (2001). The state of the states. Education Week, 20(17), 86-108, Jan 11, 2001.
- Pearson, P. D., & Valencia, S. W. (1987). Assessment, accountability, and professional prerogative. In J. E. Readence & R. S. Baldwin (Eds.), Research in literacy: Merging perspectives, (pp. 3-16). Rochester, NY: National Reading Conference.
- Popham, J.W. (1999). Why standardized test scores don't measure educational quality. Educational Leadership, 56(6), 8-15.
- Rief, L. (1990). Finding the value in evaluation: Self-evaluation in a middle school classroom. Educational Leadership, 47(6), 24-29.
- Shavelson, R. J. & Stern, P. (1981). Research on teachers pedagogical thoughts, judgments, decisions and behavior. Review of Educational Research, 51(4), 455-498.
- Shepard, L.A. (1999). The role of assessment in a learning culture. Educational Researcher, 29(7), 4-14.
- Smith, M.L. (1991). Put to the test: The effects of external testing on teachers. Educational Researcher, 20(5), 8-11.
- Spandel, V. (1996). Seeing with new eyes : A guide-book on teaching & assessing beginning writers. Portland, OR: Northwest Regional Educational Laboratory.
- Stiggins, R. J. (1997). Student-centered classroom assessment. 2nd edition. Upper Saddle River, N.J.: Merrill.
- Stiggins, R.J. & Conklin, N.F. (1992). In teachers' hands. Investigating the practices of classroom assessment. Albany, NY: State University of New York Press.
- Valencia, S.W. (1998). Literacy portfolios in action. Belmont, CA: Wadsworth.
- Valencia, S. W. & Place, N. (1994). Literacy portfolios for teaching, learning, and accountability: The Bellevue literacy assessment project. In S. W. Valencia, E. H. Hiebert, & P. P. Afflerbach (Eds.), Authentic reading assessment; Practices and possibilities, . Newark, DE: International Reading Association.
- Wiggins, G. (1998). Educative assessment : Designing assessments to inform and improve student performance. San Francisco, CA: Jossey-Bass.
- Wiggins, G. (1993). Assessment: Authenticity, context, and validity. Phi Delta Kappan, 200-214.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 79(7), 703-713.
- Wolf, D. P. (1989). Portfolio assessment: Sampling student work. Educational Leadership, 46(7), 35-39.

# Author

## Sheila W. Valencia



Sheila W. Valencia is Professor of Curriculum & Instruction at the University of Washington, Seattle where she conducts research and teaches in the area of Language, Literacy, & Culture. A nationally recognized expert in the field of literacy assessment and professional development, Dr.

Valencia has published widely in journals such as *Reading Research Quarterly*, *Journal of Literacy Research*, *The Reading Teacher*, and *Language Arts*. She is author of *Literacy Portfolios in Action* and contributing author to ***Houghton Mifflin Reading: A Legacy of Literacy***.