

New Assessment Beliefs for a New School Mission

We have inherited an assessment legacy that has actually prevented us from tapping the full power of assessment for school improvement, Mr. Stiggins maintains. He offers a new vision of assessment that has the potential of bringing about remarkable gains in student achievement.

BY RICK STIGGINS

IN RECENT years, we have achieved major breakthroughs in our understanding of the effective use of assessment to benefit — not merely check for — student learning. We have gained new insights into cognitive processes and have succeeded in connecting them to new assessment strategies that promise unprecedented achievement gains for students. Yet in districts, schools, and classrooms across the nation, educators still assess student learning the way their predecessors did 60 years ago because they have not been given the opportunity to learn about these new insights and practices.

The time has come to take advantage of this new understanding of the potential of assessment and to fundamentally rethink the relationship between assessment practices and effective schools in the United States. For decades, beginning with districtwide testing in the 1960s and subsequently expanding to statewide, national, and

international testing, we have believed that the path to school improvement is paved with more and better standardized tests. The mistake we have made at all levels is to believe that once-a-year standardized assessments alone can provide sufficient information and motivation

RICK STIGGINS is the founder of Assessment Training Institute, Inc., Portland, Ore. An earlier version of this article was presented at the annual meeting of the National Council on Measurement in Education, San Diego, April 2004. ©2004, Richard J. Stiggins.

to increase student learning.

In fact, this belief in the power of standardized testing has blinded public officials and school leaders to a completely different application of assessment — day-to-day classroom assessment — that has been shown to trigger remarkable gains in student achievement. Before discussing the evidence of the power of classroom assessment, it is useful to examine the specific reasons why standardized tests are insufficient as the foundation for assessment's role in our school improvement efforts.

A NAIVE AND COUNTERPRODUCTIVE ASSESSMENT LEGACY

Let me be clear about my mission here. The arguments I advance do not arise from a desire to end accountability-oriented standardized testing. Such tests do provide opportunities for educators to reflect on what is and is not being achieved. If educators don't take advantage of these opportunities, it is not the fault of the tests. I will suggest specific ways for users to take far greater advantage of standardized tests in the future. But for assessment to become truly useful, politicians, school leaders, and society in general must come to understand the gross *insufficiency* of these tests as a basis for assessment for school improvement.

My argument is not with the idea of accountability per se. As public institutions under contract with their communities to help students learn, schools should be compelled to present evidence that they are doing their job. If standardized tests can provide part of that evidence, we should use them. Besides, the demand for accountability is helping educators clarify achievement expectations. This has already produced dividends in the form of focused standards — a solid foundation for greater student success — and the development of standards-referenced tests. When carefully developed, such tests provide educators with the assurance that good instruction will result in higher scores.

My argument is with those who believe that standardized testing for public accountability harnesses the full power of assessment in the service of better schools. I can find little evidence that this is the case. My quest for research on the effects of high-stakes tests per se on student achievement has yielded just one study that directly addresses this question. Margaret Raymond and Eric Hanushek report tiny test score gains attributable to the presence of high-stakes tests.¹ But at the same time, Audrey Amrein and David Berliner, among others, report that such tests are often accompanied by such negative outcomes as reduced achievement, increased dropout rates, and reduced graduation rates — especially for minority students.²

In fact, Robert Linn, for decades an international leader in the development of large-scale tests for all levels, laments the inability of these tests to improve school quality and student well-being:

As someone who has spent his entire career doing research, writing and thinking about educational testing and assessment issues, I would like to conclude by summarizing a compelling case showing that the major uses of tests for student and school accountability during the past 50 years have improved education and student learning in dynamic ways. Unfortunately, that is not my conclusion.³

In a similar vein, Lorrie Shepard, an assessment expert whose international stature parallels Linn's, issues a more stinging indictment:

The negative effects of high-stakes testing on teaching and learning are well known. Under intense political pressure, testing scores are likely to go up without a corresponding improvement in student learning. In fact, distortions in what and how students are taught may actually decrease students' conceptual understanding.⁴

So our investment of billions of dollars over six decades in district, state, national, and international testing for accountability has produced scant evidence that these tests have increased student achievement or provided the motivation to learn. At the same time, we have seen mounting evidence of great harm for some segments of our student population.

I believe this lack of demonstrably positive impact arises from the fact that our assessment systems have been built on a fundamentally flawed set of beliefs about how to use assessment for educational improvement. These mistaken beliefs have forced educators to approach standardized testing far more as a matter of compliance with political demands for test scores than as a matter of pedagogy. While this may not have been the intent, it has become the reality.

A LEGACY OF MISTAKEN BELIEFS

I will cite four commonly held beliefs about the use of assessment as a school improvement tool, all of which I will argue are wrong. I will state why and then suggest remedies that promise to bring the full force of assessment to bear on school improvement.

Mistaken belief 1. High-stakes standardized tests are good for all students because they motivate them to learn.

In recent years, we have witnessed increasing reliance on standardized tests of ever-more rigorous academic standards that are connected to high-stakes promotion and graduation decisions. Admittedly, the primary intent of these accountability-oriented assessments is to pressure educators to teach more effectively. But clearly, this pressure is being passed on to students as well. This represents the educational expression of a deep-seated societal value — when the going gets tough, the tough get going. By intensifying the pressure to succeed, we strive to cause students to try harder and learn more.

The impact. Raising the bar to world-class standards assessed by high-stakes tests will cause students to try harder *only* if they believe that increased effort will lead to success. Typically, students who believe this already have a demonstrated record of learning success. They have developed into confident learners who believe they can become even more successful in the future. At the very least, they believe the chances that they will succeed are high enough to justify taking the risk of trying.

Now consider those students whose academic record reveals a chronic history of failure. Their reality is different. For them, the realization that the bar is going even higher — that now it will be even more difficult to succeed in school — is neither invigorating nor motivating. On the contrary, it is deflating, discouraging, and defeating. These students will regard the entire movement to embrace high standards and high-stakes testing and the intimidation-driven school improvement process as representing yet another occasion when they lose. They will see the new higher standards as unattainable and will give up in hopelessness. This is why dropout rates increase and graduation rates decline as the stakes go higher.

When a child goes through such an experience, the high-stakes test of world-class standards has exactly the opposite effect of the one that society and politicians desire or expect. Instead, the testing program leads to greater failure for many of our students, particularly our minority students. I believe the fact that this type of assessment has destroyed the motivation of as many students as it has encouraged has contributed to the lack of demonstrable positive impact of high-stakes, standardized testing over the decades.

A more productive belief. High-stakes tests without supportive classroom assessment environments harm struggling students. The crucial question is, What exactly is our responsibility regarding these failing students? Shall we merely write them off as collateral damage in the school improvement wars? Please realize that the schools from which most of today's adults graduated did exactly that.

Our schools were designed to sort us from the lowest to the highest achievers in order to channel us into the various segments of our social and economic system. Our grades were used to rank us: there are the winners, and the devil take the hindmost. Back then, if students gave up in hopelessness, it was a good thing, for they would learn even less and easily fill the very lowest ranks. The greater the distance between the top and bottom of the grade distribution, the more dependable would be the rank order.

High-stakes tests without supportive classroom assessment environments harm struggling students.

But in recent years, our society has changed to a school mission that places educators clearly in the service of the success of all students. We began to realize that if all schools do is sort students, then the bottom third of the rank order plus all those who drop out before being ranked will fail to develop the essential reading, writing, and math proficiencies needed to survive in an increasingly complex society. So over the past few decades, the mission of sorting has evolved into a mission of ensuring certain minimal competencies. Now schools are to “leave no child behind”; that is, they are to help *all* students meet state standards and become competent readers, writers, and problem solvers as demonstrated by appropriately high scores on state assessments. Given this new mission, if some students regard those standards as unattainable, feel hopeless, and stop trying, it is a very bad thing. Those who stop trying stop learning. Those who stop learning fail to meet the standards that reflect the skills and knowledge needed by our society.

What, if anything, can schools do to prevent this hopelessness and loss? The answer is not to eliminate high-stakes tests. Rather, it is to build learning environments that help *all* students believe that they can succeed at hitting the target if they keep trying. We understand how to use classroom assessment to keep students confident that the achievement target is within reach. We know that high-stakes tests help only when accompanied by learning environments that consciously set students up for high-stakes success. I provide details below.

Mistaken belief 2. It is the instructional decisions of adults that contribute the most to student learning and school effectiveness. We have built our assessment traditions and systems — indeed, the entire accountability movement and the more recent data-driven decision-making

frameworks — on the belief that it is the instructional decisions of the adults in our classrooms, schools, districts, and states that determine school effectiveness. We have invested literally all of our assessment resources — billions of dollars across all grade levels and over decades — to provide these adults with the assessment results they need to make sound instructional decisions, utterly and completely ignoring the students as users.

The impact. Our collective assessment history includes no acknowledgment whatsoever of students as assessment users or instructional decision makers. Without question, maximum learning comes from productive interactions between teachers and students. In fact, both must share the responsibility for making schools effective. Clearly, adults make major contributions to the process. But in a normative sense, those adult decisions are not nearly as important in terms of their impact on learning as are the decisions students themselves make.

Students decide whether they are smart enough to meet standards, whether they have any reason to hope for success if they try. They decide whether meeting the standards is worth the required effort. They decide whether it is safe to try in the face of uncertainty — whether they are likely to succeed or be embarrassed by public failure. And they base these decisions on their own view of their personal history of academic success or failure.

The time has come for us adults to deepen our understanding of the relationship between assessment and student success *from the student's perspective*. For us, a test score becomes an entry in the gradebook or a report of scores received by a test-scoring service. We see scores as comparable and, therefore, as something we can average across students in a given classroom, school, district, or state. We see scores as a means to comply with reporting requirements — grades to parents, score reports to the state department of education, reports of student performance to the media, and so on.

But for students, the score or grade represents something far more important — far more personal. The score or grade provides the information by which students decide whether or how they fit into the world of writers, readers, or math-problem solvers. Students read the score as evidence of whether success is even within reach for them. And their decisions often have long-term implications.

A more productive belief. Students are crucial instructional decision makers whose information needs must be met. We must stop being so adult-centered in our thinking about assessment. We must build classroom environments in which students use assessments to understand what success looks like and how to do better the next time.

In effect, we must help students use ongoing classroom assessment to take responsibility for their own academic success.

Mistaken belief 3. The instructional decisions that have the greatest impact on student learning are those made once a year. If we regard the manner in which we have spent virtually all of our assessment dollars over the past 60 years as evidence of our beliefs about what will have the greatest impact on student learning, then we must conclude that once-a-year decisions informed by once-a-year standardized tests are the only ones that we have believed would improve school quality. The *only* assessments we have invested in are multilayered standardized tests, which speaks volumes about our unwavering belief that the once-a-year decisions made centrally by program planners based on once-a-year test scores drive school quality. If we had believed otherwise, we would have invested otherwise.

The impact. Obviously, no one (including me) actually believes that school quality turns only on once-a-year instructional decision making. We make instructional decisions in a variety of contexts with varying frequency — some based on standardized tests and others on classroom assessments — all of which can assist student learning. But given our tunnel-vision investment in standardized tests, as outlined here, if there are other assessment users at other levels of instructional decision making who need access to different forms of evidence with differing frequency, those information needs have been completely ignored. We have not invested in ensuring the accuracy of classroom assessments. Thus the chances of inaccurate assessment and therefore ineffective decision making at all other levels clearly increase. The negative impact of this process on student learning is obvious. And better, faster, cheaper, more precise once-a-year assessments cannot rectify the resulting problems. I submit that this has contributed to the lack of a demonstrable relationship between testing per se and school improvement.

A more productive belief. The instructional decisions that have the greatest impact are made day to day in the classroom. Many of the most crucial instructional decisions are made by students and teachers not once a year but every few minutes. Students decide if success is within reach and how to go about attaining it. Teachers diagnose student needs, allocate time, design and implement instructional interventions, judge student work, and assign grades. Without question, both need continuous access to evidence of student learning arising from high-quality classroom assessment. Yet we cannot provide it because our assessment beliefs and traditions have included no attention to the accuracy or effective use of day-to-day classroom as-

essment. Indeed, the evolution of assessment in this country reveals no awareness or acknowledgment of the primacy of assessment at this level.

The new belief is to acknowledge the critical importance of classroom assessment and provide teachers the tools they need to build classroom environments that promote learning through continuous student-involved assessment, record-keeping, and communication.

Mistaken belief 4. Teachers and administrators don't need to know about and understand the principles of sound assessment practice — the professional testing people will take care of that for us. If we had believed that it is important for practitioners to assess accurately at all times, including each day in the classroom, and to use assessments to inform important instructional decisions, then we would have provided them with the opportunity to learn to do so. Instead, we have invested in making sure that others — not the teachers — do the testing. In fact, our collective assessment actions over the past 60 years reveal a fundamental lack of trust in teachers and school leaders to accurately assess the achievement of their students. Society has demanded objective, third-party evidence of learning so that professional educators can't manipulate the data in their own favor. The record shows that we assumed that this use of assessment for accountability would be sufficient to meet our evidentiary needs. Obviously, it has not sufficed.

Decades ago, we separated assessment from instruction, assigned the tasks to different people, and built a wall between them. We told teachers to teach and not to worry about assessment: someone else will cover that. And we likewise told assessment people: you test and you don't need to know anything about teaching.

The impact. As a result of this apparent lack of understanding of the connection between assessment and instruction, teacher licensing laws have failed to require competence in assessment as a condition of licensure to teach. Thus teacher preparation programs have failed to weave assessment training into their curriculum. The same pattern has evolved in our preparation of school leaders. So in 2004 we remain a national faculty unschooled in the principles of sound assessment practice. We have those who teach and those who assess, and never the twain shall meet.

In addition, rather than providing teachers with the professional development they need to manage the assessment process effectively, some districts try to circumvent the problem by providing teachers with the tests they need. It's just that those assessments are often developed in the absence of quality control and so can be inaccurate. And on top of this, untrained teachers may develop their own inaccurate assessments. In either case, the evidence used

to inform day-to-day instructional decisions may frequently be invalid. The consequent incorrect decisions are apt to lead to counterproductive actions taken on behalf of student learning. And once-a-year tests and corresponding annual instructional decision making cannot overcome the negative consequences for students.

A more productive belief. Teachers must possess and be ready to apply knowledge of sound classroom assessment practices. The typical teacher will spend one-quarter to one-third of his or her professional time involved in assessment-related activities. If teachers assess accurately and use the results effectively, then students prosper. If they do it poorly, student learning suffers. And it has. Therefore, the new belief must be that, without question, teachers need to know and understand the principles of sound assessment. The evidence of student learning they gather each day influences the most crucial instructional decisions. The remedy to our current situation is to offer targeted, productive professional development to put the available classroom assessment wisdom into the hands of practitioners.

BUILDING OUR ASSESSMENT FUTURE AROUND NEW BELIEFS

As we look to our future, if we wish to create a different reality and tap the full potential of assessment as our ally in improving student learning, we must refocus our efforts around a new overarching assessment belief: we must strike a balance between standardized tests *of* learning and classroom assessment *for* learning.

Assessment systems that balance these purposes make use of an array of assessments and differentiate among the information needs of all assessment users. When systems are in balance, assessments at all levels are derived from the same set of achievement standards, but they treat those standards differently. For instance, classroom assessments provide a continuous flow of evidence of student mastery of the classroom-level learning targets that lead over time to attainment of the desired achievement standards. As assessments *for* learning, they inform instructional decisions along the way to success. Standardized tests, on the other hand, provide periodic evidence of student mastery of the standards themselves. As assessments *of* learning, they verify arrival at success.

In a balanced system, all assessments provide dependable information about student achievement, regardless of who developed the tests or where they are used. That is, we can count on standardized tests to accurately evaluate performance toward our achievement standards because we pay to have them developed by professionals who are

expert in developing high-quality assessments. And, in such a system, we have the same confidence in classroom assessments because we invest in professional development to ensure that teachers possess the wisdom needed to create high-quality day-to-day assessments.

In systems where we take full advantage of assessment for learning, we also see balance among the participants in the assessment process. Teachers involve their students in classroom assessment, record-keeping, and communication during learning. But when it's time for students to be accountable for what they have learned, the teacher takes the lead in conducting assessments of learning.

Finally, in balanced systems, care is taken to ensure that reporting procedures deliver assessment results into the hands of the various intended users of the information in a timely and understandable manner, regardless of the purpose for the assessment (of or for learning) or the origin of the results (classroom or standardized assessment).

HOW WOULD BALANCE AFFECT ACHIEVEMENT?

Extensive research on the impact of effective classroom assessment on student achievement has demonstrated remarkable gains of a full standard deviation or more in student scores on subsequent assessments of learning. Studies have demonstrated that assessment for learning rivals one-on-one tutoring in its effectiveness and that the use of assessment particularly benefits the achievement of low-performing students.⁵ The latter finding has direct implications for districts seeking to reduce achievement gaps between minorities and other students.

And the evidence continues to accumulate.⁶ Achievement gains of the magnitude seen in the research on balanced assessment are unprecedented in the literature of school improvement.

BALANCE IS WITHIN REACH

As a direct result of lessons learned over the past two decades, we understand how to blend standardized and classroom assessment into a synergistic system — how to help them work in harmony. We have standardized assessments of learning already in place. If school leaders were given the opportunity to learn more about how to use this type of assessment, they would be prepared to take advantage of standardized test results in making crucial instructional decisions. It is paradoxical that we have invested so heavily in the creation of these tests yet have invested nothing to ensure their proper use.

We also understand how to use classroom assessment

to trigger large achievement gains. We can accomplish this through deep student involvement in day-to-day classroom assessment, record-keeping, and communication. Through student involvement in classroom assessment, we can focus students on a clear path to ultimate success. If we engage students in continuous self-assessment over time, we can keep them believing that success is within reach if they keep striving. And if we provide them with the opportunity to use this evidence to tell the story of their success, such as in student-led parent/teacher conferences, we can tap a wellspring of confidence and motivation to learn that resides naturally within each student.

All that we lack now are the political will and the professional development resources needed to place these very powerful assessment tools into the hands of teachers and administrators. We have the ability to implement an exciting new vision of excellence in assessment that manifests four new beliefs.

- Our job is to set students up for high-stakes success by helping them believe in themselves as learners.
- We must acknowledge that students are assessment users and use assessment to help them discover gifts they didn't know they had.
- Crucial instructional decisions are made in the classroom every day, not just once a year, and must be based on accurate evidence of learning for the sake of student success.
- All educators absolutely must understand and use sound assessment practices.

In short, we know how to use classroom assessment to make success a driving force in the learning life of every student. We no longer need to accept the assessment legacy of our past. We know better.

1. Margaret E. Raymond and Eric A. Hanushek, "High-Stakes Research," *Education Next*, Spring 2003, www.educationnext.org/20033/48.html.

2. Audrey L. Amrein and David C. Berliner, "High-Stakes Testing, Uncertainty, and Student Learning," *Education Policy Analysis Archives*, vol. 10, no. 18, 2002, <http://epaa.asu.edu/epaa/v10n18/>.

3. Robert Linn, "Assessments and Accountability," *Educational Researcher*, February 2000, p. 14.

4. Lorrie A. Shepard, "The Role of Assessment in a Learning Culture," *Educational Researcher*, October 2000, p. 9.

5. See, for example, Benjamin S. Bloom, "The Search for Methods of Group Instruction as Effective as One-to-One Tutoring," *Educational Leadership*, May 1984, pp. 4-17; Paul Black and Dylan Wiliam, "Assessment and Classroom Learning," *Educational Assessment: Principles, Policy and Practice*, vol. 5, 1998, pp. 7-74; and idem, "Inside the Black Box: Raising Standards Through Classroom Assessment," *Phi Delta Kappan*, October 1998, pp.139-48.

6. Samuel Meisels et al., "Creating a System of Accountability: The Impact of Instructional Assessment on Elementary Children's Achievement Scores," *Education Policy Analysis Archives*, vol. 11, no. 9, 2003, <http://epaa.asu.edu/eapp/v11n9/>; and Michael C. Rodriguez, "The Role of Classroom Assessment in Student Performance on TIMSS," *Applied Measurement in Education*, vol. 17, no. 1, 2004, pp.1-24. 

File Name and Bibliographic Information

k0400sti.pdf

**Rick Stiggins, "New Assessment Beliefs for a New School Mission,"
Phi Delta Kappan, Vol. 86, No. 1, September 2004, pp. 22-27.**

Copyright Notice

The author holds copyright to this article and has given Phi Delta Kappa International, Inc., permission to distribute it. All further use must comply with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Note that photographs, artwork, advertising, and other elements to which Phi Delta Kappa does not hold copyright may have been removed from these pages.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.