Assessment Matters:

Constructing Model State Systems to Replace Testing Overkill

FairTest
National Center for Fair & Open Testing

Assessment Matters:

Constructing Model State Systems to Replace Testing Overkill

By Monty Neill
Executive Director, FairTest

A Report by the National Center for Fair & Open Testing

October 2016

FairTest
P.O. Box 300204
Boston, MA 02130
www.fairtest.org
fairtest@fairtest.org
617-477-9792

Assessment Matters:

Constructing Model State Systems to Replace Testing Overkill

Table of Contents

Note: The full report that this section is excerpted from is available online at:
http://www.fairtest.org/assessment-matters-constructing-model-state-system

Assessment Matters:
Constructing Model State Systems to Replace Testing Overkill

Part IV

Principles and Uses of Assessment:
"To Assess" Means "To Sit Beside"

The word "assess" derives from the Latin term meaning "to sit beside." Assessing implies a direct and active relationship between or among people. Assessing could involve an observation by a teacher, a conversation between teacher and student, or looking at a student's academic work. It typically includes interaction to provide feedback or find out more about the student's thinking or depth of knowledge.

In this document, we often use the term "assessing" to mean the process rather than the instrument or system ("assessment"). In schools, assessment is a relationship that revolves around teaching and learning. Thus, it must be rooted in the content and skills students should learn.

The primary purpose of assessing should be to improve the depth and breadth of student learning, including their ability to learn. Other applications can be built on that foundation. These could include ascertaining whether students have met a threshold of knowledge and skills needed for high school graduation, or evaluating a school or district. These secondary decisions can be derived from the primary assessment evidence gathered by educators. Other forms of evidence, such as external or large-scale exams and measures of school or district quality, might be considered.

"Test" and "exam" are largely interchangeable, but they do not mean the same thing as "assessment." Tests are tools that are used in a direct relationship, as when a teacher administers a test to her students. Or districts or states could place tests between teachers and student. In the latter case, the *Standards for Educational and Psychological Testing* say that such tests "standardize the process by which test takers' responses to test materials are evaluated and scored" (AERA, APA, NCME, 2014, p. 2).

Tests are commonly used by themselves or in conjunction with other kinds of evidence to judge students, educators or schools. Using them on a stand-alone basis – as a mandatory hurdle – to make high-stakes judgments, generally violates the *Standards* (e.g., Standard 12.10, p. 198, for students; or 13.9, p. 213, for programs or schools). Evaluation of students, educators and

schools can and should include more than academic achievement – as ESSA now requires, at least minimally, of schools.

This chapter first considers three primary assessment purposes, then proposes a set of principles to guide assessing.

Assessing for, of and as learning

Assessing is a relationship and a process that employs various tools to support and evaluate student learning. Experts increasingly describe assessment as *for learning*, *of learning*, and *as learning*.

Assessing for learning, or "formative assessment," is used to provide feedback to students or be used by the teacher to modify instruction to improve student learning (Third International Conference, 2009). Formative assessment can range from a teacher observation or conversation with a student to a multiple-choice quiz to probing questions when a student is engaged in an extended project. It must provide meaningful, useful feedback. In order to fit the curriculum and provide actionable information, it should be primarily controlled by the teacher. Students also can assess one another and self-assess. Practices such as the "descriptive review of the child" also serve formative purposes as teachers closely appraise a child as a learner and a whole person in order to better understand and serve her (Himley & Carini, 2000).

Teachers often administer mini-tests created outside the classroom. Commonly, these are mandated by local administrators. Many are now taken on a computer. Their purpose is to determine how well a student has advanced in a pre-set curriculum or toward higher scores on standardized tests. Use of such instruments does not represent good formative assessment. Those are interim tests (also called benchmark, periodic or predictive).

Assessment of learning, or summative assessment, evaluates a student's learning or attainments. Grades and most standardized tests serve this purpose. Standardized tests have often been misused to determine graduation, promotion, or program placement. Culminating projects or evaluations based on portfolios serve this purpose as well. Such assessments can also play a formative role in improving teaching year to year, strengthening school curriculum, and improving how districts use resources. Summative assessments should be rooted in ongoing student work. External exams should serve only as a check on the system. They should not be a primary measure of student progress or a requirement for placement, advancement, graduation, or college admissions.

Assessment as learning emerged as a concept in the 1990s as educators sought to craft projects, tasks or exams in which the effort required to do well provided instructional benefits. Well-designed projects are a much more powerful practice than exams because they provide depth of learning and student engagement. For example, a student could develop a project,

with teachers providing feedback along the way. The final product is evaluated, and the students reflect on the process and their learning. Such evidence of learning can also become part of evaluating schools.

Big Picture Learning student presenting. Photo from Big Picture Learning.

Clearly there can be many forms of assessment "for, of and as" learning. Indeed, one criterion for a good assessment system is the use of "multiple measures." That means students have the opportunity to demonstrate their learning in varied formats over time. This, in turn, creates different ways in which teachers can assess to help students learn. At times, the term has been reduced to "several standardized tests" or "mostly tests with a bit of something else." Those misuses undermine educational quality and damage students (Neill, Guisbond & Schaeffer, 2001).

The related term "classroom-based" refers to assessing that is integrated with the particular curriculum and work students do. It includes "for, of and as." Classroom-based assessing is under the control of teachers and primarily constructed by them – "practitioner developed," in the words of the New York Performance Standards Consortium (NY Consortium, N.D.).

As with all such definitions, lines blur. Teachers may share assessments or borrow from a library of tasks assembled from many classrooms. The key point is that teachers have control. Students deserve a strong voice in the work they do and should also learn to self-assess. The NY Consortium (N.D.) often refers to their Performance-Based Assessment Tasks as "school-based," since the shaping of these tasks and the scoring guides used are set by schools as collectives of teachers/practitioners.

Principles for Assessment

States and districts are rethinking assessment in light of the opportunities provided in ESSA, such as its innovative assessment program. As they do so, they should be guided by these core principles. In drafting them, FairTest has relied on previous work, especially the *Principles and Indicators for Student Assessment Systems* (National Forum on Assessment, 1995). The primary differences are that we now more strongly emphasize student and teacher agency, and we have compiled even more evidence of the damage centralized control, especially via standardized tests, can wreak on school quality.

We are mindful of the strong advice in the report of the Forum on Educational Accountability (FEA, 2007). It states that the foundational requirements for high-quality assessment are that "all students have equitable access to the resources, tools, and information they need to succeed and by building capacity to improve teaching and learning" (p 2). With that in mind, FairTest offers these Principles:

1. ***The primary purpose of assessing is to enhance learning****.* Formative assessing, including assessment as learning, and classroom-based practices should have top priority. It involves documenting student progress, allowing students multiple methods to demonstrate their learning over time, and providing usable feedback to students. Using assessment evidence for summative decision-making or reporting about students or systems is secondary and must not undermine the primary purpose.

2. ***Assessing is rooted in significant learning outcomes and is therefore integrated with standards and curricula.*** As explained by the National Forum on Assessment (1995), "Learning goals or content standards describe broad, important intellectual competencies – knowledge, skills, understandings and habits of mind – that students should acquire and be able to demonstrate. These include important learning in and across subject areas, with a focus on thoughtful application and meaningful use of knowledge" (p. 5). Assessing focuses heavily on critical thinking, problem solving, research, writing, public speaking, presentations, initiative, self-development and group collaboration. In doing so, it pays attention to building blocks of declarative and procedural knowledge.

3. ***Assessments are primarily practitioner developed and controlled.*** Classroom educators must have the primary authority and responsibility to ensure high-quality assessing as they determine curriculum and engage directly with students. Professionals must know and use strong assessment practices.

4. ***Assessing is student focused so that they exercise significant control where appropriate,*** such as choice of books to read, research topics and projects. James Coleman (1966) found that while family background was the primary predictor of student outcomes, the second most powerful predictor was students' sense of control over their learning. In college admissions, this means students should decide whether to include SAT, ACT or similar test results in their applications.

5. ***Assessing should be fair, unbiased and culturally responsive for use in an increasingly diverse society.*** This means ensuring teachers are culturally competent and can recognize and counter biases they may have individually or as a group. Instruments such as standardized tests or scoring guides used for projects and portfolios are carefully vetted to eliminate cultural bias. In addition, assessing by teachers and instruments they use must be well constructed to best instruct students with disabilities and English language learners (ELL) and allow them to demonstrate their progress. Universal design is one key to that process. (Oregon, 2015, principle 2; FEA, 2007, Principle III.)

6. ***An assessment system uses tasks, projects, portfolios and learning records as core tools.*** Teachers may find a test is useful as a supplementary tool. On a large scale, state tests could be used as a check on the system but should not themselves carry decision-making weight.

7. ***An assessment system is as decentralized as possible.*** States build toward "systems of systems" comprised of local assessments. School and system evaluations should primarily use classroom/school-based evidence. State- or district-mandated exams are used sparingly, if at all. As required under the federal Every Student Succeeds Act, a state must establish adequate comparability across local systems. States are also responsible for ensuring local systems are valid, reliable and used fairly.

8. ***Professional collaboration and development are necessary to ensure high-quality assessing.*** A good deal of research finds that teachers often have too narrow a repertoire of assessment practices (Stiggins, 2014). This problem became much worse under NCLB's focus on standardized tests with the ensuing proliferation of interim exams and falsely labeled formative instruments. However, much evidence demonstrates teachers prefer to use a rich array of practices, and they can collaboratively learn and improve (Gallagher, 2007).

9. ***Communities participate in developing assessments.*** As schools, districts and states overhaul their practices, they must engage communities. This starts with a shared process of defining the purposes of education, its desired goals and outcomes, then considering how to evaluate the system. It includes participating in system reviews. As assessments should support teachers and students in reaching the goals, communities should understand and contribute to how assessment can do so. For example, parents can share their knowledge about how their child learns, her interests and needs. Students and parents can share their views on the benefits and drawbacks of particular assessment approaches. (See NFA, Principle 5.)

10. ***Systems incorporate feedback loops to ensure continuing improvement.*** This includes reviewing professional learning programs, examining goals as well as tools such as scoring guides, and periodically revisiting how assessment can best strengthen teaching and learning.

Taken together, these principles direct states to build systems of local assessments that are practitioner-controlled and student-focused. The new systems rely primarily on performance assessments that are classroom- and school-based. They must be unbiased and culturally responsive. They also must gather diverse forms of evidence of student learning over time. Such systems minimize statewide exams and the use of multiple-choice and short answer questions, including computer-based tests. They require professional learning and collaboration and well as meaningful community participation. And they are rooted in expectations of rich student learning in a strong, deep, culturally responsive curriculum in well-supported schools.

Resources and References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: 2014.

Coleman, J., et al. 1966. *Equality of educational opportunity*. Washington, DC: United States Department of Health, Education, and Welfare, Office of Education. U.S. Government Printing Office.

Forum on Educational Accountability. 2007, August. *Assessment and Accountability for Improving Schools and Learning.*
http://www.edaccountability.org/AssessmentFullReportJUNE07.pdf

Gallagher, C. 2007. *Reclaiming Assessment*. Portsmouth, NH: Heinemann.

Himley, M., and Carini, P. 2000. *From Another Angle*. New York: Teachers College Press.

National Forum on Assessment. 1995. *Principles and Indicators for Student Assessment Systems.*
http://www.fairtest.org/principles-and-indicators-student-assessment-syste

Neill, M., Guisbond, L., and Schaeffer, B. 2001. *Failing Our Children*. FairTest.
http://www.fairtest.org/node/1778

New York Performance Standards Consortium. N.D. *Educating for the 21st Century.*
http://performanceassessment.org/articles/DataReport_NY_PSC.pdf

Oregon Education Investment Board, Oregon Education Association, Oregon Department of Education. 2015, May. *A New Path for Oregon: System of Assessment to Empower Meaningful Student Learning.*
https://www.oregoned.org/images/uploads/blog/FINAL_July_2015_Assessment_Document_a.pdf

Stiggins, R. 2014. *Revolutionize Assessment.* Thousand Oaks, CA: Corwin Press.

Third International Conference on Assessment for Learning, Dunedin, New Zealand. 2009. *Position Paper on Assessment for Learning.* http://www.fairtest.org/position-paper-assessment-learning