CCSSO
Council of Chief State School Officers

Performance Counts: Assessment Systems that Support High-Quality Learning

By Linda Darling-Hammond

*This paper was made possible by a grant from the **Sandler Foundation** and the **Ford Foundation.***

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

PERFORMANCE COUNTS: ASSESSMENT SYSTEMS THAT SUPPORT HIGH-QUALITY LEARNING

Linda Darling-Hammond, Author

Council of Chief State School Officers
One Massachusetts Avenue, NW, Suite 700
Washington, DC 20001-1431
Phone (202) 336-7000
Fax (202) 408-8072
www.ccsso.org

Foreword

This white paper describes what a student assessment system could look like if built from the principles and best practices found in current educational research and effective educational systems in the U.S. and high-achieving nations around the world.

With this paper, the Council of Chief State School Officers seeks to illuminate and enrich the discussion around comprehensive systems of student assessment and to help lead the development of more effective ways to assess student learning. It describes existing successful assessment systems that use a variety of ways to ensure student achievement in programs in our country and in high-achieving countries around the world.

We are indebted to Linda Darling-Hammond, for developing and presenting this paper. Her experience and expertise are highly regarded. The Council is engaged in a number of conversations with the states about the nature, substance, and design of assessment consortia, and while this paper does not represent the official position of the Council, it will serve as a catalyst and resource for our ongoing conversations and planning.

We believe this paper can be a tremendous resource to states as they design new programs that will engage educators and learners in new and powerful ways. But this will not be a quick fix. This new direction rewrites the rules about assessing students from a top-down concept to a more inclusive engagement of educators and learners in the process.

Gene Wilhoit, Executive Director

Council of Chief State School Officers

Assessment Systems that Support High-Quality Learning

By Linda Darling-Hammond

Over a number of years, CCSSO has been working with key stakeholders to develop a set of principles for student assessment systems. These principles suggest that the student assessment process should be considered as a ***system*** that supports a variety of purposes, such as informing learning and instruction, determining progress, measuring achievement, and providing partial accountability information.

The ***Student Assessment System*** should

- address the depth and breadth of standards as well as all areas of the curriculum, not just those that are easy to measure
- consider and include all students as an integral part of the design process, anticipating their particular needs and encouraging all students to demonstrate what they know and can do
- honor the research indicating that students learn best when given challenging content and provided with assistance, guidance, and feedback on a regular basis
- employ a variety of appropriate measures, instruments, and processes at the classroom, school, and district levels, as well as the state level. These include multiple forms of assessment and incorporate formative as well as summative measures
- engage teachers in scoring student work based on shared targets

In this system:

All students have a clear idea of how learning progresses and what they can do to improve. Next generation learners are encouraged to demonstrate their learning as a continuous process.

Parents understand the expectations for their children's learning as well as the information they receive from school, district, and state assessments. They can work with educators to support their children's growth and progress.

Teachers are skilled at developing and using a range of assessments based on standards, learners' needs, and their professional judgment. Scoring student work based on shared learning targets is common classroom practice for teachers. Teachers are well educated and supported in these new expectations.

Supportive educators, including school principals, administrative staff, and leaders at the school and district levels, understand the standards and assessment elements and create conditions for successful learning.

Student achievement information generated at all levels of the assessment system becomes part of the longitudinal state data system and contributes to a rich profile of accomplishment for every student.

In the light of these principles and the new federal commitment to rethink assessment, this paper was commissioned to examine the lessons from research and experience—both international and from the United States—about assessment systems that support students in acquiring higher order thinking and performance skills and that support learning for students, educators, and schools, and states.

This paper outlines three areas: lessons learned from best practices in the United States and from other countries, potential responsibilities for different actors in the system, and possible steps for consortia of states to create a new system.

Creating Internationally Benchmarked Standards and Assessments

Contemporary efforts to create a set of common core standards in the United States have been grounded in a desire to create more internationally competitive expectations by benchmarking learning objectives to those in high-performing nations abroad. The Common Core State Standards Initiative aims for standards that are "fewer, higher, and deeper," based on analyses revealing that higher achieving countries teach fewer topics more deeply each year, focus more on reasoning skills and applications of knowledge, and have a well worked out sequence of expectations grounded in developmental learning progressions within domains.[i]

Implementing standards also requires an understanding of how these topics are taught and assessed and, as a consequence, what students actually learn and can do. European and Asian nations that have steeply improved student learning have created curriculum guidance and assessments that focus on defining how knowledge is to be used to solve problems and develop cognitive skills: the abilities to find and organize information to solve problems, frame and conduct investigations, analyze and synthesize data, self-monitor and improve one's own performance, communicate well in multiple forms, work in teams, and learn independently.

The assessment systems in high-achieving jurisdictions like Australia, Finland, Hong Kong, the Netherlands, and Singapore have long relied largely on open-ended items—essays and problem solutions—that require students to analyze, apply knowledge, and write extensively. Furthermore, a growing emphasis on inquiry-oriented learning has led to an increasing prominence for school-based tasks, which include research projects, science investigations, development of products, and presentations about these efforts.

Because these assessments are embedded in the curriculum, they influence the day-to-day work of teaching and learning, focusing it on the use of knowledge to solve problems. Standardized performance tasks are incorporated into examination scores in systems as wide-ranging as the General Certificate of Secondary Education in Britain, the Singapore examinations system, the certification systems in Victoria and Queensland, Australia, and the International Baccalaureate program, which operates in more than 100 countries around the world.

According to the Hong Kong Education Examinations Authority, which is increasing the use of school-based assessments (SBA) in its examination system,

> The primary rationale for SBA is to enhance the validity of the assessment, by including the assessment of outcomes that cannot be readily assessed within the context of a one-off public examination, which may not always provide the most reliable indication of the actual abilities of candidates SBA typically involves students in activities such as making oral presentations, developing a portfolio of work, undertaking fieldwork, carrying out an investigation, doing practical laboratory work or completing a design project, help students to acquire important skills, knowledge and work habits that cannot readily be assessed or promoted through paper-and-pencil testing. Not only are they outcomes that are essential to learning within the disciplines, they are also outcomes that are valued by tertiary institutions and by employers.[ii]

High-achieving systems seek to implement their standards with assessments that measure performance in authentic ways and with intensive teacher engagement throughout the assessment process, as teachers work with others to develop, review, score, and use the results of assessments. Comparability in scoring is achieved through the use of standardized rubrics, as well as training and moderation systems that enable scorers to use the same standards in

consistent ways that result in reliable scores. Systems for inter-rater moderation, statistical calibration, and auditing of scores allow the use of tasks eliciting ambitious intellectual work and extensive teacher involvement.[iii]

These strategies resemble those that have been used in leading-edge assessment systems in the United States, such as those in Connecticut, Kentucky, Maine, Maryland, and Vermont, as well as programs like the Advanced Placement program. Research has shown that these systems have supported improvements in instruction and student learning, as well as more authentic evaluations of performance.[iv]

A Vision Built on Educational Research and Successful Practices

An analysis of common elements of effective assessment systems in the United States and abroad reveals several key themes:

1) ***The student assessment process is guided by common standards and grounded in a thoughtful, standards-based curriculum. It is managed as part of a tightly integrated system of standards, curriculum, assessment, instruction, and teacher development.***

Large nations like Australia, Canada, and China manage curriculum and assessment at the state or provincial level, while small nations like England and Singapore—which have school populations about the size of California and Kentucky, respectively—have national systems managed by a ministry of education. Each of these jurisdictions has undertaken a careful process of developing standards (generally described as curriculum expectations) and curriculum guidance, often in the form of syllabi, to guide teachers' instruction in the classroom, as well as professional development that is organized around the curriculum.

- Curriculum guidance is lean but clear and focused on what students should know and be able to do as a result of their learning experiences. Assessment expectations are described in the curriculum.
- Curriculum and assessment are organized around a well defined set of learning progressions along multiple dimensions within subject areas. These guide teaching decisions, classroom-based assessment, and external assessment.
- Teachers and other curriculum experts are involved in an extensively vetted curriculum development process and in the process of developing assessment measures grounded in the curriculum standards. These guide professional learning about curriculum, teaching, and assessment. Thus, everything that comes to schools is well aligned and pulling in the same direction.

2) ***A balance of assessment measures that includes evidence of actual student performance on challenging tasks that evaluate applications of knowledge and skills.***

The curriculum and student assessment process seek to teach and evaluate knowledge and skills in authentic ways that examine a broad array of skills and competencies and generalize to higher education and multiple work domains. They emphasize deep knowledge of core concepts within and across the disciplines, problem solving, collaboration, analysis, synthesis, and critical thinking. As a large and increasing part of their examination systems, high-achieving nations use open-ended performance tasks and school-based, curriculum-embedded assessment to give students opportunities to develop and demonstrate higher order thinking skills such as the abilities to find and

organize information to solve problems, frame and conduct investigations, analyze and synthesize data, and apply learning to new situations. The curriculum and assessment systems evaluate students' abilities in a variety of tasks such as projects, group work, open-ended tasks, and oral presentations. The system would also employ summative measures such as examinations that include essays and open-ended tasks and problems, along with tests using selected-response (multiple-choice) items, usually given at the end of a course or year.

3) Teachers are integrally involved in the development of curriculum and the development and scoring of assessment measures for both the on-demand portion of state or national examinations and local tasks that feed into examination scores and course grades.

Most successful systems in the U.S. and other high-achieving nations invest in extensive moderation to ensure an accurate, reliable, and consistent scoring process and enable teachers to deeply understand the standards and develop stronger curriculum and instruction. The moderated scoring process is a strong professional learning experience, and as teachers become more skilled at using new assessment practices and developing curriculum, they become more effective at teaching the standards. The assessment systems are designed to increase the capacity of teachers to prepare students for the demands of college and careers in this new century and global society.

4) Assessment measures are structured to continuously improve teaching and learning.

Assessment ***as***, ***of***, and ***for*** learning is enabled by several features of successful assessment systems:

- The use of school-based, curriculum-embedded assessment (more complex assessment exercises that all students are expected to complete over an extended timeframe) provides teachers with models of good curriculum and assessment practice, enhances curriculum equity within and across schools, and allows teachers to see and evaluate student learning in ways that can feed back into instructional and curriculum decisions.
- Close examination of student work and moderated teacher scoring of both school-based components and externally developed open-ended portions of examinations are sources of ongoing professional development that improve teaching.
- Developing both school-based and external assessment measures around learning progressions allows teachers to see where students are on multiple dimensions of learning and to strategically support their progress.
- School-based, curriculum-embedded assessment engages students in their own learning process and builds their capacity to assess their own learning.

5) Assessment and accountability systems are designed to improve the quality of learning and schooling.

The student assessment process produces evidence of learning. That evidence is critical information for informing the learning process for both the student and teacher and for informing decision makers about the quality of the educational program and the accountability of the personnel who are responsible. But there must be a balance in the system between these two uses of the evidence. The need for accountability using large-scale, high-stakes, summative assessments should not overshadow assessment's primary purpose of providing timely feedback to the teachers and learners engaged in the instructional process. The interval of time between when the evidence is produced and when it is used to alter the course of instruction is crucial to

improving the quality of the learning. A shorter time interval increases the value of the information used to modify the learning process.

High-achieving states and nations invest most of their resources in high-quality assessments that aim to drive the learning of ambitious intellectual skills in the classroom. In order to maintain investments in well vetted expert processes of development and scoring, most countries implement external tests for students only once or twice prior to high school (generally around grades three and six), with continuous school-based assessment throughout these years.

High school examinations in high-achieving nations are generally selected from an array of subjects by students to demonstrate their areas of competence for colleges and employers. These assessments also inform course grades, support individual student learning, and shape curriculum improvement. The tests are typically not used to determine student graduation from high school; they set a higher standard linked to college and career expectations.

High-achieving states and nations implement accountability systems that publicly report outcomes and take these into account, along with other indicators of school performance, in a well designed system focused on continual improvement for schools. Many nations combine assessment data with information from school inspections to design intensive professional development supports and interventions that improve school performance. Many of these inspectorate systems use experts to examine teaching, learning, and school operations up-close in order to diagnose school needs and guide more targeted improvement efforts.

6) Assessment and accountability systems use multiple measures to evaluate students and schools.

High-achieving countries use multiple measures (multiple sources of evidence of varying types) to evaluate skills and knowledge needed for the demands of this dynamic, technological era. Students engage in a variety of tasks and tests that are both curriculum embedded and on demand, providing many ways to demonstrate and evaluate their learning. These are combined in reporting systems at the school and beyond the school level. School reporting and accountability are also based on multiple measures, including student achievement as one indicator among many. Other indicators often include student participation in challenging curricula, progress through school, graduation rates, college attendance, citizenship, a safe and caring climate, and school success and improvement.

7) New technologies enable greater assessment quality and information systems that support accountability.

New technologies enhance and transform the way the assessment process is developed, delivered, and used, providing adaptive tools and access to information resources for students to demonstrate their learning, and appropriate, immediate feedback by supporting both teacher scoring and computer-based scoring.

Technology also organizes data about student learning, enhancing system accountability for instruction and reporting by providing more efficient, accurate, and timely information to teachers, parents, administrators, and policymakers. In the current U.S. context, technology can help to integrate information at all levels of the system as part of a longitudinal state data system, contributing to a rich profile of accomplishment for every student.

By applying these lessons as well as new knowledge from the leading edge of assessment development, we can imagine a systemic approach to transforming assessment of learning in the United States.

Responsibilities in This New System

States have the responsibility of providing education. and the oversight for teacher and leader education and certification. Therefore, states should assume a significant leadership role for defining the standards and assessment process. In a new system in the United States, states, districts and schools, and the federal government should function in the following ways.

States

- Create "common core" standards—mapped across the grade spans in a set of learning progressions around key dimensions of learning—to serve as the basis for state curriculum and assessment efforts.
- Adopt and augment the standards as appropriate to their context.
- Create and deploy curriculum frameworks that address the standards, drawing on exemplars and tested curriculum models.
- Build and manage an assessment system that includes both on-demand and curriculum-embedded components that evaluate the full range of standards and allow evaluation of student progress. Consortia of states might create joint assessment measures and an assessment bank of performance tasks linked to the standards that can be used as part of both on-demand and curriculum-embedded assessments.
- Develop rubrics that embody the standards and clear examples of good student work, benchmarked to performance standards.
- Provide resources, materials, and opportunities for teachers and school leaders to fully understand and incorporate standards and assessments into their practice.
- Create an oversight/moderation/audit system for ensuring the comparability of locally managed and scored assessment components.
- Ensure that teacher and leader education and development infuse knowledge of learning, curriculum, and assessment, along with opportunities to evaluate student work and develop responsive teaching strategies.
- Implement high-quality professional learning focused on examination of student work, curriculum and assessment development, and moderated scoring.
- Create robust information systems that collect, report, and use data to support teaching and inform policy.

Districts and Schools

Local districts and schools have the responsibility to develop and adapt the curriculum and instructional materials necessary for teaching the standards and assuring that the learning process is effective. Consequently, they would

- Examine the standards and evaluate current curriculum, assessment, and instructional practice in light of the standards.
- Evaluate state curriculum guidance and further develop and adapt the curriculum to support local student learning, select and augment curriculum materials, and continually evaluate and revise the curriculum in light of student learning outcomes.

- Design, select, and incorporate formative assessment into the curriculum, organized around the standards, curriculum, and learning progressions to inform teaching and student learning.
- Participate in administering and scoring relevant portions of the on-demand and curriculum-embedded components of the assessment system and examine student work and outcomes.
- Help design and engage in professional development that includes learning, teaching, curriculum, assessment, and the effective use of information systems.
- Engage in review and moderation processes to examine assessment measures and student work, within and beyond the school.

The Federal Government

In our system, the federal government plays an important supporting role providing guidance and research, legislation and funds to address overarching national issues and ensure that our educational systems and practice conform to prevailing federal laws. In a new assessment system, the federal government would

- Consider the traditional role of the National Assessment of Educational Progress and how it might use the new blueprints already established to reflect the standards and more intellectually ambitious assessment of knowledge and skills.
- Support research on the design, outcomes, and consequences of curriculum and assessment.
- Allow, encourage, and fund the use of performance assessment for state assessment systems under the Elementary and Secondary Education Act, as well as the use of open-ended diagnostic assessment that can evaluate student performance over time.
- Support and fund initiatives to infuse knowledge of assessment and learning into pre- and in-service professional development for teachers and leaders.

How a High-Quality Assessment System Might Operate

Drawing from successful practices in the U.S. and abroad, a new assessment system might be constructed as follows:

Develop curriculum frameworks

As the common core standards are released, vetted, and adopted, states or consortia of states would work with curriculum and assessment experts to develop (or adapt from previously successful work) curriculum frameworks mapped to standards and learning progressions. There has been enormous investment in the United States in high-quality curriculum, for example through the auspices of such groups as the National Science Foundation and other organizations at the national level, and in many states and districts. Other English-speaking nations have also developed high-quality curriculum materials linked to standards and learning progressions that should be evaluated in this process. This effort would cull and inventory existing efforts with a strong evidence base of success to build curriculum frameworks, deeper curriculum development at the local level, state and local assessment development, instructional supports, and professional development.

Create a digital curriculum and assessment library

The results of this curriculum development effort should ultimately be made available online in a digital platform that offers materials for curriculum building, model syllabi for specific courses linked to the standards, formative and summative assessment tasks and instruments, and materials for training teachers and school leaders in both strategies for teaching specific curriculum concepts or units and assessment development and scoring. Assessment tasks linked to specific standards could be accessed from an assessment task bank, such as one recently developed in Hong Kong, so that they are available both for formative classroom use. In addition, as described below, an electronic scoring platform should also be developed and made available across the states.

Develop state and local assessment measures

Initially, one or more state consortia would work to create *a* ***common reference examination,*** linked to the common core standards, such as the New England Common Assessment Program assessment recently developed by a set of New England states for grades 3–8. This ***summative assessment*** would be designed to incorporate more analytic selected-response and open-ended items than many U.S. tests currently include. It would also include strategically selected performance measures that are part of the final score and provide formative information at the classroom level.

A design much like this one was developed by the New Standards project in the 1990s and has been implemented in states like Maine, Kentucky, and Vermont, which tied results on a set of performance tasks to those of a reference examination in English language arts and mathematics. These comprehensive systems were found to measure the full range of standards while improving student learning.

In the new system, ***curriculum-embedded performance components***, counted as part of the overall examination score, would be developed around core concepts or major skills that are particularly salient in evaluating students' progress in English language arts and mathematics. Exemplars to evaluate and build upon are already available in many states and in nations like England that have developed a set of "tests and tasks" for use in classrooms

that help teachers evaluate students' learning in relation to well described learning progressions in reading, writing, mathematics, and other subjects.

The curriculum-embedded components would link to the skills evaluated in the on-demand test and would allow for more ambitious performances that take more time than can be allocated in a two- or three-hour test on a single day. They would evaluate skills in ways that require student-initiated planning, management of information and ideas, interaction with other materials and people, and production of extended responses that reveal additional abilities of students (e.g., oral presentations, exhibitions, and product development—as well as written responses).

In the context of summative assessments, curriculum-embedded tasks would be standardized and scored in moderated fashion, and scores would be aggregated to count as part of the external assessment. The design of curriculum-embedded assessments would also include marker tasks to be used formatively to check for essential understandings and to give teachers useful information and feedback as part of ongoing instruction. Thoughtful curriculum guidance would outline the scaffolding and formative assessment needed to prepare students to succeed on the summative assessments.

All components of the system would incorporate ***principles of universal design*** that seek to remove construct-irrelevant aspects of tasks, which could increase barriers for non-native English speakers and students with other specific learning needs. In other words, the tests would be designed to measure only the knowledge and skills of interest by ensuring that the meaning of questions is very clear to a wide range of learners. The ways students are asked to respond would not obscure what they know by conflating what is supposed to be assessed with other factors that are not relevant to the question at hand. In addition, designers who are skilled at developing linguistically supportive assessments and tests for students with learning disabilities would be engaged from the beginning in considering how to develop the assessments for maximum access, as well as how to design appropriate accommodations and modifications to enable as many students as possible to be validly assessed within the system.

The emphasis on evaluating ***student growth over time*** and on tying standards to a conception of learning progressions should encourage a growth-oriented frame for both the on-demand examination—that is, the two- or three-hour test that students sit down and take together—and the more extended classroom assessments. Ideally, the reference exam would incorporate computer-based adaptive testing that creates vertically scaled assessments based on the full range of learning progressions in English language arts and math. This would allow students to be evaluated in ways that give more accurate information about their abilities and their growth over time. This approach would not preclude any state or group of states from evaluating grade-level standards or from pursuing additional advanced work (e.g., Science, Technology, Engineering or Mathematics or career and tech certificates). Nor should it preclude a significant number of constructed-response, open-ended items, as the technology for machine-scoring structured open-ended items is now fairly well developed. As described later, strategic use of partial teacher scoring for these items would also be a desirable element of the system to support teachers' understanding of the standards and assessments and their planning for instruction.

The emphasis on evaluating student growth should also inform the development of the curriculum-embedded elements of the system, which should be strategically selected or developed to evaluate students' progress along the learning continuum. Centrally developed tasks administered and scored by teachers with moderation, using common rubrics, would be part of the set of reported examination scores. Existing tools like the Developmental Reading Assessment and the Primary Learning Record, which evaluates student progress along a learning continuum in ways that can inform both instruction and reporting, should be examined as well for their contribution to the classroom-embedded component of the assessment system.

In sophisticated state systems, it may be possible to begin to incorporate information about student learning that teachers develop from their own classroom evidence, linked to standards and learning progressions and guided by the curriculum frameworks. This is the primary approach to assessment before high school in countries like Australia, England, Finland, and New Zealand. This approach is likely to be most productive of more sophisticated and adaptive teaching and well supported student learning. This could be an optional aspect of the consortium's work for states and communities with interest and capacity.

At the ***high school level***, the consortium could develop summative assessments of college- and career-ready performance tied to the common core standards. In addition to assessments that create a link between high school expectations and college admissions expectations in English language arts and mathematics, a consortium might explore one or two options for assessment or combine elements of both for a third:

Option 1: Course- or syllabus-based systems like those in Alberta (Canada), Australia, England, Hong Kong, and Singapore, as well as the International Baccalaureate

Generally conceptualized as end-of-course-exams in this country, this approach should become a more comprehensive through-course assessment approach like that pursued in these other countries. Such an approach would include within-course performance assessments that count toward the examination score, as well as high-quality end-of-course components that feature extended constructed-response as well as selected-response items. Within-course performance assessments would tap central modes of inquiry in the disciplines, ensuring that students have the opportunity to engage in scientific investigations; literary analyses and other genres of writing, speaking, and listening; mathematical modeling and applications; and social scientific research. Such an approach might require an English language arts and math assessment at a key juncture that evaluates an appropriate benchmark level for high school standards, and then, as in high-achieving nations, allow for pursuit of other courses/ assessments that are selected by students according to their interests and expertise. These could serve as additional information on the diploma for colleges and employers.

Option 2: Standards-driven systems that include a more comprehensive benchmark assessment in ELA and mathematics complemented by collections of evidence that demonstrate students' abilities to meet certain standards within and across the disciplines

This set of assessments would allow more curriculum flexibility in how to meet the standards. Systems like these are used in some provinces in Australia and Canada; in states like Nebraska, New Hampshire, Rhode Island, and Wyoming; and in school organizations like Envision Public Schools, New Tech High, Asia Society schools, and the New York Performance Standards Consortium. Sometimes this evidence is organized into structured portfolios, such as the technology portfolio in New Hampshire and the broader graduation portfolios in these sets of schools that require specific tasks in each content area, scored with common rubrics and moderation. Electronic portfolios used for college admissions and placements are outcomes of many of these systems and are welcomed by universities.

Option 3: A mixed model combining elements of both course- and standards-driven models

This would allow for demonstrations of proficiency to occur in any one of a range of courses (rather than a single, predetermined course) or even outside the bounds of a course, like the efforts by some states to allow students to pass courses via demonstrations of competence rather than seat time (as legislated in New Hampshire and Ohio, among other states). Such a system could also include specific components intended to develop and display research and inquiry skills that might also be interdisciplinary, such as the project work requirements in England, Singapore, and the International Baccalaureate, and the senior project requirements in Ohio and Pennsylvania.

Develop moderation and auditing systems for teacher-scored work

A state consortium could develop protocols for managing moderation and auditing systems and training scorers to enable comparable, consistent scoring of performance assessments. In other nations' and states' systems that include these features routinely, procedures have been developed to ensure both widespread teacher involvement—often as part of professional development time—and to create common standards and high levels of reliability in evaluating student work. A range of models is possible, and the consortium would serve as a resource to individual states for developing and implementing strong, efficient approaches.

Provide time and training for teachers and school leaders

To implement an integrated system of curriculum, assessment, and instruction, time must be set aside for teacher development and participation in the system. Creative use of existing professional development days and incentives provided by recertification requirements (e.g., continuing education units) can be part of this commitment. In order to secure benefits for the quality of teaching and learning, states will need to establish concrete commitments to support teacher engagement in curriculum and assessment development, scoring, and analysis.

Use technology to support the assessment system

Technology can be used to enhance these assessments in three ways: by delivering the assessments; in online tasks of higher order abilities, allowing students to search for information or manipulate variables and tracking information about the students' problem-solving processes; and in some cases, scoring the results or delivering the responses to trained scorers/teachers to assess from an electronic platform. Such a platform may also support training and calibration of scorers and moderation of scores, as well as efficient aggregation of results in ways that support reporting and research about the responses. These technologies are already being used in the International Baccalaureate and Hong Kong assessment systems, which include both on-demand and classroom-based components.

In order to gain the efficiency and cost benefits of machine scoring and the teaching and learning benefits of teacher-moderated scoring, a mixed system would be developed. Computer-based scoring would be utilized for constructed-response tasks where useful (though teachers would score some of these tasks for anchoring and learning purposes) and teachers would be engaged to score tasks that require more nuanced professional judgments in order to support improvements in instruction.

Previously in the United States, attempts to integrate performance-based elements into testing systems have often been deemed too expensive to maintain during times of budget cuts. However, cost estimates by experienced developers based on industry standards have shown that it is possible to construct an affordable system of assessment that includes a significant number of constructed-response items along with reliably scored classroom-based performance tasks.[v] These analyses show that a high-quality system can be designed for no more than the costs paid by an average state for today's tests (generally about $20 per pupil), by making sound decisions that take advantage of the economies of scale state consortia can achieve, new uses of technology for distribution and scoring, and thoughtful approaches to teacher scoring of tasks. Great advances have been made over the last 15 years in methods and technologies for more efficiently designing and scoring assessments. Consortia should work proactively to cost out and design affordable and sustainable systems that offer intellectually challenging tasks to students, rich formative evidence to teachers, and reliable, timely information to parents, administrators, and policymakers.

Conclusion

As policymakers and citizens weigh the costs and benefits of new assessment systems, it is critically important to consider their influences on instruction. At the end of the day, the goal of new standards and assessments must be to improve the quality of learning, not just its measurement. And the goal of education must be to improve the ability of students to survive and succeed in a rapidly changing world that requires a deeper knowledge base and greater use of thinking, problem solving, and learning skills than ever before. Investments that achieve these goals will pay dividends for every member of our society for generations to come.

---

[i] See for example, W.H. Schmidt, H. C. Wang, and C. McKnight (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective, Journal of Curriculum Studies, 37 (5), 525–559; G.A. Valverde & W. H. Schmidt (2000). Greater expectations: Learning from other nations in the quest for 'world-class standards' in US school mathematics and science, Journal of Curriculum Studies, 32 (5): 651–687.

[ii] HKEAA (2009). School-based assessment: Changing the assessment culture. http://www.hkeaa.edu.hk/en/hkdse/School_based_Assessment/SBA/.

[iii] L. Darling-Hammond. (2010). Benchmarking Learning Systems: Student Performance Assessment in International Context. Stanford, CA: Stanford Center for Opportunity Policy in Education.

[iv] For a review see Linda Darling-Hammond & Elle Rustique-Forrester. (2005). The Consequences of Student Testing for Teaching and Teacher Quality. In Joan Herman and Edward Haertel (eds.) The Uses and Misuses of Data in Accountability Testing, pp. 289-319. Malden, MA: Blackwell Publishing, 2005.

[v] B. Topol, J. Olson, & E. Roeber (2010). The Cost of New Higher Quality Assessments: A Comprehensive Analysis of the Potential Costs for Future State Assessments. Assessment Solutions Group. Published by Stanford Center for Opportunity Policy in Education.