Developing Assessments of Deeper Learning: The Costs and Benefits of Using Tests that Help Students Learn

By Linda Darling-Hammond and Frank Adamson

Stanford Center for Opportunity Policy in Education
Executive Summary

Despite a growing consensus that students must acquire higher-order thinking and performance skills to succeed in today’s world, current U.S. tests, which rely heavily on multiple-choice items, measure primarily low-level knowledge and skills. A recent RAND Corporation study found, for example, that fewer than 2% of mathematics items and only about 20% of English language arts (ELA) items on state tests ask students to analyze, synthesize, compare, critique, investigate, prove, or explain their ideas—the kinds of higher-order skills that students most need to become college- and career-ready.

Such skills are incorporated in the new Common Core State Standards (CCSS), which most states have adopted to ensure that all students graduate from high school prepared to succeed in the knowledge-based world they are entering. High-achieving nations have moved to place these skills at the center of their curriculum plans and assessments, using essays, open-ended problems, and performance tasks to evaluate students’ abilities to think critically and solve problems.

The United States could take a major step in improving the direction of curriculum and assessments for deeper learning by drawing from the work of the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (SBAC)—two multi-state consortia that were formed to develop next generation assessments of Common Core state standards. If the consortia are able to live up to their plans, the quality of assessment could improve substantially. An analysis of the content specifications for SBAC, for example, found that more than two-thirds of the assessment targets in ELA and mathematics intend to tap higher-level skills that are largely ignored in today’s tests.

But high-quality assessments like those used in other countries tend to cost more than lower-quality assessments, in part because performance tasks and essays often require human scoring, whereas low-level skills can be measured with multiple-choice items that are cheap to score. Many states have traditionally budgeted only about $20 per pupil for tests in math and reading, placing severe limits on the quality of learning their tests can measure. This represents less than two-tenths of 1% (.002) of average per-pupil spending on K-12 education. Given that most of us spend at least $500 a year to assess the health of our automobiles, it is clear that this tiny investment carries a disproportionately large burden for the health of the education system.

However, states and districts together spend much more than this on all they do to increase scores on the end-of-year tests. Recent estimates put average state spending on ELA and math tests at $25 to $27 per pupil (with a range from about $13 to $105 per pupil), and spending on interim and benchmark testing at an additional $17 to $18 per pupil, not counting the costs of test preparation materials, personnel for test administration and analysis, or teacher time for scoring or professional development associated
with this testing. The combined costs of state and local testing in ELA and mathematics alone exceed $50 per pupil on average.

This level of spending could support higher quality assessments that include the kinds of open-ended items and performance tasks that can measure more complex learning, scored both by teachers and by the evolution of more sophisticated artificial intelligence (AI) engines. Ironically, though, because these billions of dollars are largely pointed at boosting performance on narrow tests that do not measure or encourage the acquisition of higher-order skills, they do not result in the improvements to learning that would be possible if the same funds were spent differently.

From a cost-benefit perspective, this approach is penny wise and pound foolish. Although they may appear low in costs, today’s testing programs are generally not organized to produce the benefits of deeper student learning found in high-performing countries. Instead we have a set of fragmented, disjointed efforts, unable to measure the most important learning goals, and not useful to teachers’ efforts to understand how their students think and what could be done to support their success.

Current investments, which still total less than half of 1% of overall per-pupil spending, could support much higher-quality assessments, including performance tasks that tap critical thinking and problem solving skills, if they were refocused to do so. A wise use of resources would make it possible to develop a much more coherent system that not only provides assessments of deeper learning, but also offers formative supports for instruction and interim tools teachers could use to see how students are doing on tasks that reveal both how they think and what they know.

In order to realize the benefits of an instructionally helpful system, it will be important to make high-quality assessments both affordable and feasible to implement while strengthening teaching and learning at the same time. States that seek such a system can achieve their goals by:

- Understanding how the state and local components of a high-quality assessment system can operate together to strengthen learning.
- Taking advantage of the cost savings associated with multi-state consortia and productive uses of technology for online delivery and efficient scoring and reporting.
- Involving teachers in developing and scoring assessments in ways that also support teachers’ professional learning and improved instruction, making these investments doubly beneficial.
- Combining state and local resources strategically to make sound, coherent investments in higher quality assessments.

The question for policymakers has shifted from, “Can we afford assessments of deeper learning?” to, “Can the United States afford not to have such high-quality assessments?” The clear answer is that curriculum and instruction systems that include assessments of deeper learning are essential if students are to develop the skills they need for a knowledge society—a prerequisite for their success, and that of the nation.