

Can **New Tests** Lead to Better Teaching and Deeper Learning?

Many states have adopted the Common Core State Standards (CCSS), which emphasize “deeper learning” skills, such as mastery of core academic content, critical thinking, problem-solving, and collaboration. But can the new tests being developed to align with the CCSS be used to promote deeper learning and high-quality classroom instruction?



Key Points

- Testing can have a wide variety of effects on teachers' activities in the classroom, including changes in what is taught, how teachers allocate time and resources, and how teachers interact with students.
- Tests of deeper learning are more likely to have an impact on classroom instruction if there are consequences for educators or students; however, attaching stakes that are too high can have negative effects on practice.
- The impact of new tests will be enhanced by policies that ensure that the tests mirror high-quality instruction, are part of a systemic change effort, and are accompanied by training and support to help teachers interpret and use test scores effectively.



How Does Testing Affect Teachers' Practice in the Classroom?

To understand how tests can be used as a lever for reform, RAND researchers explored whether new tests and associated assessment systems that are being developed to align with the Common Core State Standards (CCSS), which emphasize “deeper learning” skills, such as mastery of core academic content, critical thinking, problem-solving, and collaboration, might improve instructional quality. Researchers reviewed published research on assessment systems, including “high-stakes testing”—tests associated with important consequences, such as a student’s ability to advance to the next grade or to graduate. The study focused on the extent to which new tests might influence instructional practices and what changes in policies or context might make new tests—particularly tests of deeper learning—have a greater influence on teacher practice. The research team found considerable research on the effects of testing, with studies describing a wide variety of effects on teachers’ and students’ activities in the classroom.

Curriculum content and emphasis. Testing can lead to changes in what is taught in the classroom. In particular, a large body of research documents unanticipated and often undesirable changes in practice in response to some high-stakes tests, such as excessive emphasis on tested subjects or topics. However, studies also show that educators’ tendency to focus more on tested than nontested content can be beneficial when testing covers a broad range of skills and knowledge or encourages an increased focus on higher-order thinking skills.

Instructional activities. Testing not only can influence what teachers teach, but also in some cases affects *how* they teach. Evidence suggests that some teachers change their approach to instruction to emphasize the skills measured by a test. For example, testing has led some teachers to focus on strategies that students can use to perform well at basic skills or on certain types of test items. On the other hand, some tests can make teachers’ practices more student-centered or result in an expanded repertoire of teaching strategies and techniques.

Teachers’ interactions with students. Testing can also influence the ways in which teachers allocate their time, resources, and attention among their students. Research suggests that testing can encourage teachers to focus on meeting students’ needs by individualizing instruction for all students. But when accountability systems rely primarily on a “cut score” or proficiency level, testing can lead teachers to shift attention toward students who are “on the bubble”—i.e., on the threshold of passing the test—because doing so is most likely to increase the number of students who reach the target score or proficiency level.



What Factors Influence Teachers' Responses to Testing?

While studies suggest that teachers frequently alter their practices in response to testing, research also indicates that those changes are influenced, for better or worse, by a number of mediating factors, as shown below.

Attributes of tests. The purpose, technical quality, and format of a test can influence the ways in which it affects teacher practice. For example, tests that are explicitly intended to shape instructional practice may be more likely to promote changes in instruction than tests that are used for other purposes, such as placing students in programs. The format of the test also sends a message regarding the kinds of tasks in which students are expected to engage and therefore can influence teachers' choices regarding curriculum and instruction.

Accountability context. Much research shows that educators respond to high-stakes tests differently than to lower-stakes tests. Teachers' responses can also be influenced by specific features of the accountability system, including the grade and subjects tested, who is held accountable, and what types of metrics are used.

Educators' background, beliefs, and knowledge. The literature suggests that the characteristics of the teachers themselves—particularly the depth and breadth of teachers' domain knowledge, beliefs about curriculum and instruction, and “buy-in” of the assessment system—may affect whether and how testing influences their instructional practices.

Student and school characteristics. Some studies have found that testing can have a stronger effect in elementary school than in secondary school. Other characteristics that can influence testing effects include whether the school is located in an urban or rural area and whether it is a traditional public school, a charter school, or a private school.

District or school policy. The effects of tests are also mediated by school or district policies, including those related to the use of time, professional development and training, collaboration, and curriculum choices. Professional development was consistently highlighted as an enabling condition for testing to influence teaching practice.

The William and Flora Hewlett Foundation

The William and Flora Hewlett Foundation is committed to promoting “deeper learning” for America’s students, including mastery of core academic content, critical thinking, problem-solving, collaboration, communication, and learning how to learn. The foundation recognizes that the systems used to assess student learning can be a lever for change and asked RAND to explore whether new tests and associated assessment systems that are being developed to align with the CCSS might improve instructional quality. The CCSS, which have been adopted in more than 40 states, establish goals and expectations for student achievement in English language arts and mathematics in kindergarten through 12th grade.

Factors That Influence the Relationship Between Testing and Instructional Practices



How Can New Tests Have a Positive Impact?



Although the relationship between test-related policies and classroom practice is complex and is influenced by mediating factors, the literature does provide some guidance for thinking about the ways that new CCSS-aligned tests might affect practices. RAND researchers identified a set of conditions that would promote a positive impact of testing on instructional practices and, ultimately, deeper learning.

Test content and format should mirror high-quality instruction. For a test to have any chance of promoting deeper learning, it is critical that at least a portion of the test reflects learning activities that are consistent with the goals of deeper learning. Some sacrifice in test reliability (e.g., scoring consistency) may be appropriate to represent more demanding content and to signal its importance to teachers. However, the extent to which reliability and other aspects of technical quality can be compromised depends in large part on the stakes attached to the scores.

Score reporting should be optimized to foster instructional improvement. Tests should provide score reports that are tailored to the needs of educators. Important features of score reporting systems include rapid results, score reports that are clear and accessible to educators, and a reporting mechanism that can provide information about performance for individual students and relevant groups of students (e.g., English language learners).

Teachers should receive training and support to interpret and use test scores effectively. Teachers need ongoing guidance on how to interpret and respond to data from tests. If the tests assess skills that are unfamiliar to the teachers, then teachers will need support to improve both their own subject-matter knowledge in these areas and their skills for using the test-score data to impart this learning to students.

Accountability metrics should value growth in achievement, not just status, and should be sensitive to change at all levels of student performance, not just changes at a single cut point. Accountability indexes do not need to focus only on a “cut score” (i.e., the level of performance needed to demonstrate student proficiency) but can be constructed to focus on growth in achievement, taking into account performance all along the achievement scale. Designers of assessment systems need to understand what types of measures can be used to measure growth and need to select a growth modeling approach that is well suited to their assessments and to the purposes of the evaluation and accountability system.

The test scores should “matter,” but important consequences should not follow directly from test scores alone. Thoughtful planners build in mechanisms to ensure that tests focus on deeper learning. These mechanisms might include multiple measures that emphasize important outcomes or processes that are not measured by the test, to prevent the test from becoming the sole incentive or source of guidance.

Externally mandated, high-stakes tests should be part of an integrated assessment system. A comprehensive assessment system should provide timely and consistent information that can be used for multiple purposes, including instructional improvement, student self-reflection, mastery certification, and system monitoring. In such a system, different assessments would address different purposes but would all be implemented in support of each other.

Testing should be one component of a broader systemic reform effort. One way to reduce teachers’ tendency to overemphasize tests as a source of instructional guidance is to adopt a coherent system of reforms that starts with the standards and aligns other elements to those standards. These elements include curriculum and instructional materials, professional development and support for teachers, data systems, accountability policies, and strategies for community engagement.

Tests should be used only for purposes for which they were designed and validated. Those who design assessment policy or use test scores for decisionmaking should also monitor the uses and consequences of testing programs over time so that they can identify and address inappropriate uses or unintended consequences. Users should not make important decisions on the basis of tests until their validity for those specific purposes is demonstrated.

This brief describes work done for RAND Education and documented in Susannah Faxon-Mills, Laura S. Hamilton, Mollie Rudnick, and Brian M. Stecher, *New Assessments, Better Instruction? Designing Assessment Systems to Promote Instructional Improvement*, Santa Monica, Calif.: RAND Corporation, RR-354-WFHF (available at www.rand.org/t/RR354), 2013. The RAND Corporation is a nonprofit research institution that helps improve policy and decisionmaking through research and analysis. RAND’s publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

www.rand.org

HEADQUARTERS CAMPUS

1776 Main Street
P.O. Box 2138
Santa Monica, California
90407-2138
TEL 310.393.0411
FAX 310.393.4818



CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND
TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Support RAND

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore the [RAND Corporation](#)

View [document details](#)

Research Brief

This product is part of the RAND Corporation research brief series. RAND research briefs present policy-oriented summaries of individual published, peer-reviewed documents or of a body of published work.

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).