3-1994

What You Assess May Not Be What You Get

Thomas R. Guskey

University of Kentucky, GUSKEY@UKY.EDU

Click here to let us know how access to this document benefits you.

Follow this and additional works at: https://uknowledge.uky.edu/edp_facpub

Part of the Curriculum and Instruction Commons, Educational Assessment, Evaluation, and Research Commons, and the Teacher Education and Professional Development Commons

Repository Citation


https://uknowledge.uky.edu/edp_facpub/18

This Article is brought to you for free and open access by the Educational, School, and Counseling Psychology at UKnowledge. It has been accepted for inclusion in Educational, School, and Counseling Psychology Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
Performance-based assessments may not bring significant change in instructional practice unless teachers are provided requisite time and training.

Few innovations in education have caught on as quickly as performance-based assessment. Nearly every current reform initiative includes a provision to assess students' performance on complex learning tasks.

The tasks specified include essays, demonstrations, computer simulations, performance events, portfolios of students' work, and open-ended questions and problems. Collectively, these measures are referred to as authentic assessments because they are valuable activities in themselves and involve the performance of tasks that are directly related to real-world problems (Linn et al. 1991).

Two major factors have intensified the interest in performance-based assessment. First, advances in cognitive science (see, for example, Resnick 1985 and 1987, or Shuell 1986) have compelled educators to acknowledge how complex learning is and how diverse are the means needed to assess learning fully and fairly.

Second, many educators have recognized the limitations of assessment systems that have relied on multiple-choice, standardized achievement tests. Researchers have found that such systems, especially those used to ensure accountability, encourage teachers to skew their instruction to the basic skills assessed in the tests (Haladyna et al. 1991, Shepard 1990). As a result, the curriculum narrows and the validity of information gathered from the tests diminishes (Mehrens and Kaminski 1989, Shepard 1989). These two effects are commonly amplified in schools serving at-risk and disadvantaged students because such schools are under great pressure to show improvement in test scores (Herman 1992).

Advocates of new performance-based assessments believe that if teachers are going to teach to tests, the tests (or other forms of assessment) should be worth teaching to. Then, reformers hope, first-rate tests will call forth first-rate instruction. For example, assessment devices that tap higher-order thinking skills will elicit instructional practices that emphasize and develop higher-order thinking skills. An added benefit is that the performance-based assessments are likely to become an integral part of the instructional process, rather than a separate, after-the-fact check on student learning (Wiggins 1989a). The distinction between instruction and assessment would thus become "seamless."

Some educators have carried this vision a step farther, suggesting that authentic, performance-based assessments could actually drive instructional improvements (McLaughlin 1991). This approach is called measurement-driven instruction, or MDL (Popham 1987, Popham et al. 1985).
Reform in Kentucky

Kentucky recently enacted reform legislation that takes the measurement-driven instruction approach. The new law, the Kentucky Education Reform Act (KERA), is one of the most comprehensive pieces of educational reform legislation ever enacted in the United States. It addresses administration, governance and finance, school organization, professional development, curriculum, assessment, and accountability.

A significant component of KERA is a student assessment program called the Kentucky Instructional Results Information System (KIRIS). KIRIS is a multifaceted program that evaluates:

- portfolios of students' work in writing and mathematics;
- students' achievement on "performance events" in the areas of mathematics, science, social studies, arts and humanities, and vocational education/practical living; and
- students' scores on "transitional tests," which include both multiple-choice and open-ended items similar to those in National Assessment of Educational Progress tests. Transitional tests assess performance in reading, writing, mathematics, science, social studies, arts and humanities, and vocational education/practical living.

KIRIS is a high-stakes assessment program. That means that results from the assessments will be used to grant financial rewards to schools that improve significantly and to levy sanctions against schools that fail to show progress (Foster 1991). The high-stakes nature of the assessment program is what makes KERA a measurement-driven reform effort (Guskey 1994).

KERA and KIRIS are particularly interesting to educators and policymakers for two important reasons. First, although KERA is not the first reform effort to include a comprehensive assessment system, it is the first to be driven by assessments that are primarily performance-based. Second, KERA is the first statewide reform effort with high-stakes performance-based assessments.

High-stakes assessment itself is not new to Kentucky educators. During the 1980s, results from the administration of a statewide test known as the Kentucky Essential Skills Test (KEST) were used to rank school districts throughout the Commonwealth and to dispense rewards and sanctions (Guskey and Kifer 1990). KIRIS is a sharp departure from KEST, however, in that KEST was composed entirely of multiple-choice items designed to assess basic skills.

The Vitali Study

The implementation of these two conceptually different high-stakes statewide assessment programs, both used within a 10-year period, presented an excellent opportunity to compare the impact of each program. One of my doctoral students, Gary Vitali, recently set out to determine the impact of such assessment systems on teachers' instructional practices.

Vitali's study involved extensive teacher interviews, several teacher questionnaires, and classroom observations (Vitali 1993). The findings offer new insights into the complexities of measurement-driven reform and also challenge the notion that what you test is invariably what you get in the classroom.

Vitali's findings support those of other researchers who have found that multiple-choice, standardized achievement measures employed for accountability purposes do focus instruction on the content of the tests. Most teachers narrow the curriculum and focus instruction on basic skills. They do so, Vitali found, because they want their students to do well on the tests, whether or not the teachers believe that the content and skills being measured are important. He also discovered that most teachers think teaching to standardized tests is fairly easy to do. After all, narrowing instruction is easier than broadening it, and most teachers reported that their instructional materials are generally aligned with a basic-skills orientation.

The performance-based assessment program, on the other hand, resulted in only modest changes in teachers' instructional practices. A few teachers who recognized that most of the performance tasks and portfolio entries required students to do some writing did respond by incorporating additional writing activities in their daily lessons and in classroom assessments. For the vast majority of teachers, though, lesson plans, classroom activities, and evaluations of
student learning remained unchanged.

This finding was all the more surprising in view of teachers’ positive attitudes toward performance-based assessments. Teachers regarded these more broadly based assessments as better measures of student learning than multiple-choice, standardized achievement tests.

In interviews and questionnaires, Vitali sought to determine why teachers did not make more significant adaptations, especially considering the high-stakes nature of the Kentucky assessment program. He discovered that, simply put, teachers did not know how to teach to the performance-based assessments, nor did they believe that they could do so within their current time constraints.

Although most teachers said they felt “under the gun” to adapt instructional practices to the performance-based assessments, the transition seemed insurmountable because it required professional training and time the teachers did not have. The need for training seemed especially critical since the realignment would involve an expansion both of what is taught (curriculum) and how it is taught (methods).

The teachers’ perceptions were borne out. Although Vitali found the vast majority of teachers to be dedicated, hard-working individuals who want their students to do well, he also discovered that, in general, teachers were ill-prepared to adapt their instructional practices to the new demands of a more authentic, performance-based assessment program. Most teachers had scant knowledge, personal background, experiences, or formal training with the various types of performance-based assessments or ways to use them as instructional tools. The only training that most teachers had received was scattered, one-day staff development workshops.

The lack of personal experience and professional training in instructional techniques that might help students prepare for performance-based assessments was a widespread problem that seemed to affect both elementary and secondary teachers. Many respondents also stressed that the lack of appropriate teaching materials was a problem.

Vitali sought to determine why, simply put, teachers did not know high-stake nature of the Kentucky assessment program. He discovered that, simply put, teachers did not know how to teach to the performance-based assessments, nor did they believe that they could do so within their current time constraints.

Although most teachers said they felt “under the gun” to adapt instructional practices to the performance-based assessments, the transition seemed insurmountable because it required professional training and time the teachers did not have. The need for training seemed especially critical since the realignment would involve an expansion both of what is taught (curriculum) and how it is taught (methods).

The teachers’ perceptions were borne out. Although Vitali found the vast majority of teachers to be dedicated, hard-working individuals who want their students to do well, he also discovered that, in general, teachers were ill-prepared to adapt their instructional practices to the new demands of a more authentic, performance-based assessment program. Most teachers had scant knowledge, personal background, experiences, or formal training with the various types of performance-based assessments or ways to use them as instructional tools. The only training that most teachers had received was scattered, one-day staff development workshops.

The lack of personal experience and professional training in instructional techniques that might help students prepare for performance-based assessments was a widespread problem that seemed to affect both elementary and secondary teachers. Many respondents also stressed that the lack of appropriate teaching materials was a problem.

Accountability Won’t Ensure Success

Although the Vitali study has limitations and its findings will require confirmation and elaboration by other researchers, the results clearly indicate that instituting a high-stakes, performance-based assessment program, even one as thoughtfully designed and as carefully implemented as Kentucky's KIRIS program, is not enough to bring about significant change in the instructional practices of most teachers. Adapting instructional practices to performance-based assessments, the study shows, is a much more complex process than many advocates of measurement-driven instruction assume. Bridging the chasm between authentic assessment and authentic classroom practice will demand a substantial amount of additional time, resources, and training opportunities.

Teachers perceived two general types of time pressures. First, teachers reported that they were being required to do more and teach more, without any increase in the amount of time allowed for planning or instruction. (Secondary teachers mentioned this obstacle more often, possibly because they have traditionally been more content-oriented and thus see performance-based assessments as a more drastic change than do skills-oriented elementary teachers.) Second, most teachers believed that performance-based assessments would require a lot more time to administer and score.

These perceptions of little time and lots of extra work, combined with inadequate experience, training, and materials, appeared to keep most teachers frozen in virtually the same instructional patterns that they had before the new assessment system. Thus Vitali concluded that “what you test may not be what you get” when performance-based assessments are the primary testing tool and teachers have neither adequate time nor sufficient training to teach to the test.

Bridging the chasm between authentic assessment and authentic classroom practice will demand a substantial amount of additional time, resources, and training opportunities.
processes, and recommend classroom assessment designs that are more performance-based (Stiggins 1987). Adequate treatment of these topics will certainly require more extensive time commitments than a one-day inservice program. Further, because the challenge involves the expansion of teachers’ expertise and instructional repertoires, regular follow-up and continuous support will also be important factors (Guskey 1991).

Thus the lesson from Vitali’s study is clear. Performance-based assessments, by themselves, appear to be insufficient to bring about significant change in the instructional practices of most classroom teachers, and without change in instructional practice, improvement in student learning cannot be expected. On the other hand, combining authentic, performance-based assessments with high-quality professional development opportunities to help teachers align instruction with improved assessments will make significant advances in student learning much more likely.

Some critics, like Airasian (1988) and Worthen (1993), have expressed reservations about MDI, and Cizek (1991, 1993) and Mathison (1990) are among those who view the approach as unethical. Even so, proponents of performance-based assessments argue that they are more likely to engage students in complex intellectual challenges than are the uninspired teaching practices—so common today—that promote only memorization of unrelated bits of information (Wiggins 1989a).

References


*Educational Researcher* 22, 4: 4-9.


Thomas R. Guskey is a Professor, Educational Policy Studies and Evaluation, 131 Taylor Education Building, College of Education, University of Kentucky, Lexington, KY 40506.