ED327612 1990-06-00 Can Performance-Based Assessments Improve Urban Schooling? ERIC Digest Number 56.

ERIC Development Team

www.eric.ed.gov

Table of Contents

If you're viewing this document online, you can click any of the topics below to link directly to that section.

ERIC Identifier: ED327612
Publication Date: 1990-06-00
Author: Ascher, Carol
Source: ERIC Clearinghouse on Urban Education New York NY.

Can Performance-Based Assessments Improve Urban Schooling? ERIC Digest Number 56.

THIS DIGEST WAS CREATED BY ERIC, THE EDUCATIONAL RESOURCES INFORMATION CENTER. FOR MORE INFORMATION ABOUT ERIC, CONTACT

ACCESS ERIC 1-800-LET-ERIC
Nationally, the use of standardized short-answer and multiple-choice tests has grown dramatically in recent years. While both the form and content of these tests have increasingly driven curriculum, students' scores have also become a major influence on promotion and placement decisions, on the professional advancement of teachers, and even on school funding (Bracey, 1989). In urban schools serving low-income, linguistic and cultural minority students, the reliance on these tests has been especially heavy--narrowing curriculum and exacerbating tracking and failure rates (FairTest & NYPIRC, 1990).

In an effort to enrich curriculum to the fullest extent possible, and to evaluate students more fairly, 28 states have begun to use essays to assess writing, and there is a move to replace standardized tests with "performance-based assessments" in some subjects. Such assessments, it is claimed, stress the higher order skills that schools should be teaching, make good diagnostic instruments, and are much better at eliciting the potential of disadvantaged students. However, the urgency to initiate an alternative to short-answer and multiple-choice tests may be leading educators to underestimate some of the problems of performance-based assessments, particularly when used in high-stakes situations.

BENEFITS OF PERFORMANCE-BASED ASSESSMENT

The best way to discover how students think, or to diagnose where they are having difficulties in learning--which, aside from accountability and placement, is the main reason for testing--is to give them as much range as possible to express themselves fully (Archbald & Newmann, 1988), and to assess their learning in its natural context, as they make active use of the skill (Gardner, in press).
While traditional standardized tests assume that a multiple choice question about possible grammatical solutions to a sentence can indicate a student's writing skills, a performance-based assessment simply samples the student's writing itself. While traditional tests are scored on the basis of "objective" notions of right and wrong answers, performance-based assessments entail clear human judgments, and can even include dialogues that enable the students tested to ask for clarification and to explain their answers (Wiggins, 1989). Whereas most standardized tests measure only discrete linguistic and number skills, assessments in context can assess a far wider range of competencies (Gardner, in press). Like life, where most of the important problems faced are open-ended and complex, performance-based assessments require each student to demonstrate mastery in a personal and more integrated way (Archbald & Newmann, 1988). Finally, in contrast to standardized tests, which have "predictive validity," assessments in context have "ecological validity"--that is, students perform as they will have to in life.

TYPES OF PERFORMANCE-BASED ASSESSMENTS

Several variations on performance-based assessment have been studied by researchers, experimented with by private testing companies, and instituted by public school systems.
Station Activities. Students proceed through a series of discrete tasks, either individually or in teams, in a given amount of time, much as in a science laboratory. They might be asked to measure electrical currents, sort seeds, compare the absorbency of paper products, or infer the characteristics of objects sealed in boxes. The questions asked are open-ended to elicit students' thinking strategies.

Domain Projects. Students conduct a rich set of exercises designed to explore an idea, concept, or practice central to a particular academic or artistic domain. For example, students are asked to test which paper towels are best as judged by a variety of criteria. They must solve a wide range of science and math problems to set up the criteria and make their judgments (Raizen & Kaser, 1989).

Portfolios. An extension of domain projects, portfolios consist of several projects completed in a sequence to show students' progress with a subject. Portfolios can include initial plans, drafts, self-evaluations, feedback from peers and teachers, plans for subsequent projects, etc.

Videotaping. Although this technology is reliable and inexpensive, its use is still relatively experimental as an assessment technique. However, one project used videotaped interviews to assess the mathematics understanding of primary school students (Resnick & Resnick, 1989).

WIDE USE OF PERFORMANCE-BASED ASSESSMENTS

The most obvious difficulties with using performance-based assessments on a large scale stem from the reliability and cost of scoring. Supporters of this new assessment method generally point to the fact that performance-based assessments are already used widely in art and athletics, where, for example, pooled ratings are created on artists' portfolios or athletes' Olympic dives. Moreover, portfolios have been used in England and Wales, where teachers and external examiners are trained to score them with a high degree of agreement (Goldstein & Wolf, in press). In 1988 the National Assessment of Educational Progress (NAEP) processed some 18,000 writing assignments as part of its battery of tests (NAEP, 1988), and its scoring of these assignments suggests that reliable, publicly believable, quantitative measures can be derived from judgments of these assignments (Resnick & Resnick, 1989).

On the other hand, such assessments are expensive to score compared to machine-scorable tests. Most of the current performance-based assessments are being scored by groups of teachers. Needless to say, they cannot compete with the 10,000 mechanically scored tests per hour that the Educational Testing Service now achieves (Frederiksen, 1984). Still, there is no reason to test as often, or as many students, as is currently done. Sampling methods are now sophisticated enough that states could easily attain quite accurate information on their schools' effectiveness through assessing only a portion of their students (Resnick & Resnick, 1989).

Although it has been said that performance-based assessments can "protect students from teacher bias" (FairTest & NYPIRC, 1990), the problems of bias that have always plagued standardized testing remain unsolved in the new state-wide assessment programs. In fact, there are indications that, without significant changes in teaching, essays and other open-ended assessments may result in decreased scores for low-income whites and minorities (DePalma, 1990).

Finally, although states have mandated performance-based tests in order to enrich teaching and learning, the very pressure exerted by mandated testing programs is likely to work against these state goals. Although the Irish essay exams are often used as beacons for this country's path, the Irish experience contains examples of rote preparation of students with essays that can be shifted slightly, depending on the topic called for (Madaus, 1990, personal communication). In the United States, performance-based assessments have already opened out the ways in which subjects like English or mathematics are taught in some school districts; however, it is also already clear that instruction can be (and sometimes is being) geared quite narrowly to the format, or even the specific questions, on the assessments (Madaus, 1988).

CONCLUSIONS

The hope remains that, despite some obstacles, performance-based assessments will support a richer, more open ended curriculum and more accurately assess the skills of low-income minority students whose gifts and needs are diverse. However, as long as performance-based assessments are used as part of high stakes testing situations, pressure to generate good and improving test scores means that there is no sure safeguard against a new trivialization of learning.

REFERENCES

Archbald, D.A., & Newmann, F.M. (1988). Beyond standardized testing: Assessing authentic academic achievement in the secondary school. Reston, VA: National Association of Secondary School Principals.

Bracey, G.W. (1989, April). Standardized tests: Their nature, their problems, their replacements. Prepared for the Association of Colorado Educational Evaluators. Englewood, CO: Cherry Creek Schools.

DePalma, A. (1990, November 1). Revisions adopted in college entrance exams. The New York Times, A1, A24.

FairTest, & NYPIRC. (1990). Standardized tests and our children: A guide to testing reform in New York. Cambridge and New York: National Center for Fair and Open Testing and New York Public Interest Research Center.

Frederiksen, N. (1984, March). The real test bias: influences of testing on teaching and learning. American Psychologist, 3 (3), 193-202.

Gardner, H. (in press). Assessment in context: The alternative to standardized testing. In B. Gifford & M.C. O'Conner (Eds.), Future assessments: Changing views of aptitude, achievement, and instruction. Boston: Kluwer Academic Publishers.

Goldstein, H., & Wolf, A. (In press). Recent trends in assessment: England and Wales. In L.C. Wing & B.R. Gifford (Eds.), Trends in educational testing and assessments. Boston: Kluwer Academic Publishers.

Madaus, G. (1988). Testing and the curriculum. In L.N. Tanner, Cultural issues in curriculum. 87th Yearbook, Part I, National Society for the Study of Education. Chicago: University of Chicago Press, pp. 83-121.

National Assessment of Educational Progress. (1988). Writing report card. Princeton, NJ: Author.

Raizen, S.A., & Kaser, J.S. (1989, May). Assessing science learning in elementary school: Why, what, and how? Phi Delta Kappan, 70 (9), 718-722.

Resnick, L.B. & Resnick, D.P. (1989). Assessing the thinking curriculum: New tolls for educational reform. Prepared for the National Commission on Testing and Public Policy. Unpublished manuscript. Learning Research and Development Center, University of Pittsburgh, & Department of History, Carnegie-Mellon University, Pittsburgh, PA.

Wiggins, G. (1989, May). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70 (9), 703-713.

Wolf, D.P. (1987, December/1988, January). Opening up assessment. Educational Leadership, 45 (4), 24-29.

The comprehensive monograph on which this digest is based, "School Progams for African American Male Students," by Carol Ascher, is available for $5 from the ERIC Clearinghouse on Urban Education. In addition to providing an expanded discussion of the issues capsulized here, it includes descriptions of 17 programs around the country,

and a comprehensive bibliography.

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under OERI contract no. RI88062013. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or the Department of Education.

__

Title: Can Performance-Based Assessments Improve Urban Schooling? ERIC Digest Number 56.
Document Type: Information Analyses---ERIC Information Analysis Products (IAPs) (071); Information Analyses---ERIC Digests (Selected) in Full Text (073);
Descriptors: Disadvantaged Youth, Educational Innovation, Educational Testing, Elementary Secondary Education, Evaluation Methods, Evaluation Problems, Literature Reviews, Standardized Tests, Student Evaluation, Test Bias, Testing, Testing Problems, Urban Education
Identifiers: ERIC Digests, Performance Based Evaluation
###

__

[Return to ERIC Digest Search Page]