ED365312

Alternative Assessment and Technology [assess2]

ERIC Digest.
ERIC Clearinghouse on Information and Technology, Syracuse, NY.
Sponsor: Office of Educational Research and Improvement (ED), Washington, DC. 1993. 4 p.

Adapted from an article by Dorothy Bennett and Jan Hawkins in "News from the Center for Children and Technology and the Center for Technology in Education," v1 n3 Mar 1992.
EDRS Price - MF01/PC01 Plus Postage.

Descriptors: Academic-Achievement; Computer-Simulation; Educational-Cooperation; Educational-Objectives; Educational-Testing; High-Schools; High-School-Students; Multiple-Choice-Tests; Portfolios-Background-Materials; Technology-Education; Thinking-Skills; Video-Equipment *Competency-Based-Education; *Educational-Technology; *Evaluation-Methods; *Student-Evaluation

ABSTRACT

This digest focuses on going beyond multiple choice testing to the development of methods for assessing complex knowledge and performances. Newly designed assessment systems must accurately measure and promote the complex thinking and learning goals that are known to be critical to students' academic success and their eventual sustained achievement and contribution to their communities. The digest discusses performance-based assessment and portfolio assessment as alternative approaches. The contribution that technology can make to the creation of workable and meaningful forms of alternative assessment is also addressed. The digest overviews some of the approaches to alternative assessment that the Center for Technology in Education (CTE) has been investigating. CTE is working in collaborative projects with a variety of schools. Within these projects, CTE has experimented with a number of tasks in the development of technology-based performance assessment records in high school science and mathematics, including computer simulations, oral presentations, paired explanations, progress interviews, and videotaped demonstrations. CTE evaluates student performance on two levels: the quality of the oral presentation and the quality of the device. A list of eight additional readings is provided. (TMK)

INTRODUCTION

Considerable attention is now being paid to the reform of testing in this country--going beyond multiple choice testing that emphasizes facts and small procedures, to the development of methods for assessing complex knowledge and performances. This is because goals for education have substantially changed during the last decade, and because changes in assessment are believed to directly influence changes in the classroom. Altering assessment practices is likely to affect curriculum, teaching methods, and students' understanding of the meaning of their work. A newly designed assessment system must accurately measure and promote the complex thinking and learning goals that are known to be critical to students' academic success and to their eventual sustained achievement and contribution to their communities.

Two approaches that have shown considerable promise are performance-based assessment and portfolio assessment. In these approaches, judgments about students' achievement are based on their performances of complex tasks and selections of work over time.

The success of a new approach to assessment carries with it a deep change in how we think about the measurement of cognitive abilities. The view of assessment carried over from the last century is that there are underlying mental traits and that a test is a sample behavior which provides an imperfect measure of the underlying characteristic the test was meant to measure. We are attempting to develop a different paradigm of assessment. The new paradigm requires methods like performance assessment or portfolio assessment. Instead of giving a test that consists of a number of varied items believed to constitute a sample of some underlying knowledge or skill, the new approach attempts to record a complex performance that represents a rich array of a student's abilities. Rather than a representative sample, it is meant to be a measure of "demonstrated capability."

A key part of assessment research is developing tasks that will enable students to use and demonstrate a broad range of abilities. Successful tasks will be complex enough to engage students in real thinking and performances, open-ended enough to encourage different approaches, but sufficiently constrained to permit reliable scoring; they will allow for easy collection of records, and they will exemplify "authentic" work in the disciplines.

THE ROLE OF TECHNOLOGY

How does technology figure in this process of reconfiguring the way students are assessed? Technology has certain unique capabilities that can make crucial contributions to the creation of workable and meaningful forms of alternative assessment. Paper and pencil, video, and computers can give three very different views of what students can do. It's like three different camera angles on the complete picture of a student. You can't reconstruct a total person from just one angle, but with three different views you can triangulate, and discover a much richer portrait of students' abilities.

Well-designed educational technologies can support these new approaches to assessment, and consequently lend themselves to integration into curricula that stress alternative assessment. Computers and video records offer expanded potential for collecting--easily and permanently--different kinds of records of students' work. For example, final products in a variety of media (text, graphics, video, multimedia), students' oral presentations or explanations, interviews that capture students' development and justifications for their work, and in-progress traces of thinking and problem solving processes are now collectible using video and computer technologies. Decisions about what records to collect is a key part of the CTE research. Essential to success is discovering what kind of records are most efficient for scoring yet capture the most important aspects of the different target abilities.

An effort has been underway at the Center for Technology in Education (CTE) to investigate two approaches to assessment; both are based on students' work on complex tasks. They explore the potential that technology holds for facilitating innovative assessment techniques by using videotape and computers. The remainder of this digest describes some of the performance based alternative assessment projects that CTE is working with in collaborative projects with a variety of schools.

PERFORMANCE ASSESSMENT

Performance assessment refers to the process of evaluating a student's skills by asking the student to perform tasks that require those skills. Performances in science might examine the ability to design a device to perform a particular function or to mount an argument supported by experimental evidence. In contrast, answering questions by selecting from among several possible choices, as in multiple choice tests, is not considered a performance, or at least not a performance that is of primary interest to scientists or science educators.

If you ask scientists what qualities make a good scientist, they might come up with a list like the following: the ability to explain ideas and procedures in written and oral form, to formulate and test hypotheses, to work with colleagues in a productive manner, to ask penetrating questions and make helpful comments when you listen, to choose interesting problems to work on, to design good experiments, and to have a deep understanding of theories and questions in the field. Excellence in other school subjects, such as math, English, and history require similar abilities.

The current testing system only taps a small part of what it means to know and carry out work in science or math or English or history, and consequently it drives the system to emphasize a small range of those abilities. In science, the paper and pencil testing system has driven education to emphasize just two abilities: recall of facts and concepts, and ability to solve short, well-defined problems. These two abilities do not, in any sense, represent the range of abilities required to be a good scientist.

With the help of collaborating teachers at partnership school sites, the Center for Technology in Education has been conducting research studies to develop and understand how technology (both video and computers) can best be deployed in new assessment systems. In a study of this approach to assessment, CTE collects sample performances, or records, for a specific set of tasks, and design and test criteria for scoring those performances. Thus far, CTE has experimented with a number of tasks in the development of technology-based performance assessment records in high school science/mathematics. The tasks and criteria for scoring them are described below.

COMPUTER SIMULATIONS.

In one science project, CTE has collected data using a computer program called Physics Explorer. Physics Explorer provides students with a simulation environment in which there is a variety of different models, each with a large set of associated variables that can be manipulated. Students conduct experiments to determine how different variables affect each other within a physical system. For example, one task duplicates Galileo's pendulum experiments, where the problem is to figure out what variables affect the period of motion. In a second task, the student must determine what variables affect the friction acting on a body moving through a liquid. Printouts of students' work can be collected and evaluated in terms of the following traits: (1) how systematically they consider each possible independent variable, (2) whether they systematically control other variables while they test a hypothesis, and (3) whether they can formulate quantitative relationships between the independent variables and the dependent variables.

ORAL PRESENTATIONS.

This task asks students to present the results of their work on projects to the teacher. These interviews include both a presentation portion, where clarification questions are permitted, and a questioning period, where the students are challenged to defend their beliefs. Students' presentations can be judged in terms of: (1) depth of understanding, (2) clarity, (3) coherence, (4) responsiveness to questions, and (5) monitoring of their listeners' understanding.

PAIRED EXPLANATIONS.

This tasks makes it possible to evaluate students' ability to listen as well as to explain ideas. First, one student presents to another student an explanation of a project he or she has completed or a concept (e.g. gravity) he or she has been working on. Then the two students reverse roles. The students use the blackboard or visual aids wherever appropriate. The explainers can be evaluated using the same criteria as for oral presentations. The listeners can be evaluated in terms of: (1) the quality of their questions, (2) their ability to summarize what the explainer has said, (3) their helpfulness in making the ideas clear, and (4) the appropriateness of their interruptions.

PROGRESS INTERVIEWS.

This is a task in which students are interviewed on videotape about the stages of their project development and asked to reflect upon the different facets of their project work. The task was developed as a means for documenting the degree of progress students make in their understanding of key concepts. Preliminary scoring criteria that have been developed to evaluate these records are: (1) depth of understanding, (2) clarity of explanations, (3) justification of decisions/degree of reflectiveness, (4) use of good examples and explanations, (5) degree of progress made relative to where the student started, and (6) understanding of the bigger picture of the project.

VIDEOTAPED DEMONSTRATIONS.

CTE is collecting data on a task that has been developed by a high school teacher in charge of a mechanical engineering program for 11th and 12th graders at Brooklyn Technical High School. Working together on design teams, students design and construct mechanical devices according to a design brief that describes technical specifications. The students must "demonstrate" their work and explain before a panel of judges from the field of engineering how their devices work and why they made certain design decisions. Students are then required to subject the devices to a functional test. For example, one project required students to design a device which can lift and lower "heavy" objects and place them at specified locations. The functional test required students to demonstrate that the devices they constructed could successfully lift and deliver three weights to a specified location in less than four minutes.

The students' performances on this task are evaluated on two levels: the quality of the oral presentation, and the quality of the device. The oral presentation can be evaluated in terms of: (1) depth of understanding of the principles and mechanisms, and (2) clarity and completeness of the presentation. The device can be evaluated in terms of: (1) the economy of design (the degree to which there was an economical use of materials); (2) craftsmanship (degree of care in fabrication and assembly of device), (3) aesthetics, (4) creativity (interesting or novel ways of accomplishing the design), and (5) controllability (stability of the device). These tasks provide interesting windows into students' abilities in the physical sciences. To complete the picture of students' performances, however, this evidence should become part of a larger portfolio of records of their work on a project, such as written descriptions, analyses, and journals.

ADDITIONAL READING

Bruder, Isabelle. (1993, Jan.). Alternative assessment: Putting technology to the test. ELECTRONIC LEARNING, 12(4), 22-23, 26-28. EJ 457 876.

Clyde, Anne. (1992, Jan./Feb.). New technology, information access and educational outcomes. EMERGENCY LIBRARIAN, 19(3), 8-14, 16-18. EJ 441 739.

Gray, Bob A. (1991). Using instructional technology with at-risk youth: A primer. TECHTRENDS, 36(5), 61-63. EJ 441 780.

Magnussun, Kris, & Osborne, John. (1990, April). The rise of competency-based education: A deconstructionist analysis. JOURNAL OF EDUCATIONAL THOUGHT/REVUE DE LA PENSEE EDUCATIVE, 24(1), 5-13. EJ 407 351.

McClure, Robert M., & others. (1992, April). ALTERNATIVE FORMS OF STUDENT ASSESSMENT. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, April 20-24, 1992. 46pp. ED 347 209.

Seels, Barbara. (1993, Jan.). THE KNOWLEDGE BASE OF THE EVALUATION DOMAIN. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, New Orleans, LA, January 13-17, 1993. 12pp. ED 355 919.

Shavelson, Richard J., & others. (1991). Performance assessment in science. APPLIED MEASUREMENT IN EDUCATION, 4(4), 347-62. EJ 446 660.

U.S. Congress. (1992, Feb.). TESTING IN AMERICAN SCHOOLS: ASKING THE RIGHT QUESTIONS. [Full Report]. Washington, DC: Office of Technology Assessment. Rpt. No. OTA-SET-519. ED 340 770. 314pp. (Also available from U.S. Government Printing Office: S/N 052-003-01275-8.)

This digest was adapted from an article by Dorothy Bennett and Jan Hawkins which appeared in News from the Center for Children and Technology and the Center for Technology in Education, Vol. 1, No. 3, March 1992, Bank Street College of Education, 610 West 112th St., New York, NY 10025. As of January 1994, the Center for Technology in Education will be affiliated with the Education Development Center, 69 Morton St., New York, NY 10014. ----- ERIC

This publication was prepared with funding from the Office of Educational Research and Improvement, U.S. Department of Education, under contract no. RR93002009. The opinions expressed in this report do not necessarily reflect the positions or policies of OERI or ED.