MICHAEL E. MARTINEZ AND JOSEPH I. LIPSON

Assessment for Learning

A new generation of tests promises to benefit teachers and students—and not just institutions.

Tests have long been used to classify people and to evaluate educational programs. The benefits of testing for these purposes, though, have accrued mainly to institutions and have generally not served learners directly. Therefore, critics have rightly questioned the pedagogical relevance of educational assessment in its present form (McLean 1986, Frederiksen 1984).

Recently, though, an exciting new vision of testing is emerging, in part from a sense of what is now technologically feasible. For example, advances in cognitive science have made it possible to understand in some detail how learning takes place in subject matter areas, such as science and mathematics. Likewise, the development of item-response theory in psychometrics has enabled us to describe growth in expertise more precisely than ever before. And, of course, great strides in computer and related technologies have been made during the past decade. In short, converging forces portend a new generation of tests—tests that better serve the interests of teachers and students in promoting learning.

Here we describe the essential elements of a new generation of tests, point out the ways in which tests could promote learning, and indicate what it will take to create them.

New Directions

The Mastery Assessment Project at Educational Testing Service was born out of a desire to make testing serve learners and teachers. The word *mastery* denotes progress along a scale from a novice state to expert status—the kind of progress we want to assist. Assessment supported by advanced technology can be better used to serve the interests of learners and teachers in various ways.

ASCD Photo by Jamieson McKenzie

Using computer simulations to present assessment tasks promises exciting use of still and moving images that will enable students to interact with atomic or galactic structures or with dangerous or unobtainable materials.

Assessment for Learners

We and other researchers at Educational Testing Service envision a new generation of assessment products that are instructionally useful—products in which testing and teaching complement and reinforce one another. Let us consider three ways in which this might be accomplished.

First, mastery systems would provide immediate, elaborated feedback to learners, a feature that separates them from most tests, in which feedback is delayed by days, weeks, or months—or never provided at all. Instead of receiving merely "correct" or "incorrect" signals, a student might be told *why* an answer was correct or incorrect or perhaps asked to justify an answer.

A second characteristic of mastery systems is that progress toward expertise would be expressed via "learning progress maps" (Bunderson 1988). These are "maps" in the sense that they would display "the big picture" of what students have learned, what knowledge and skills they are trying to master, and what remains to be learned. Such maps would give students a bird's-eye view of their intellectual odyssey. Using learning progress maps, students would be able to gauge their progress accurately and quickly at any stage in their intellectual growth. Using these maps, learners could judge their progress within a specific topic that might take a week to learn. Alternatively, they could monitor their growth across a spectrum of proficiency that might require several years of study.

A third way assessment can serve learners is by presenting questions and posing tasks that relate subject matter to the complex world we live in. "Real-world" tasks require planning, problem solving, troubleshooting, and other skills, which are frequently performed under ambiguous or changing conditions. Such tasks, though often difficult to construct and score, offer greater validity than most tests as we know them, and measure broad abilities, such as problem solving, that transcend particular topics or facts.

Researchers in education and psychology have pointed out that experts in their fields have acquired organized assemblages of knowledge and experience called "conceptual frameworks" or "schemata" (Glaser 1984). Such frameworks help experts to recall facts and principles and, like intellectual scaffolds, support further learning. Conceptual frameworks also allow experts to grasp quickly a problem or situation—much like a skilled quarterback reading a defensive alignment—and then initiate appropriate action sequences. Advanced assessment projects, including the Mastery Assessment Project, will develop ways of assessing recognition of complex patterns in the context of various subject matters.

Still another way in which assessment can tap knowledge and skills in a real-world context is to make greater use of images. Images used for assessment can take many forms, from line drawings to real objects. The value of using images and, of testing for the images in a student's head, is especially significant in science education. A deep understanding of scientific phenomena is often visual in nature, and expert scientists commonly consult "pictures in the mind" to solve problems (Miller 1986). Mastery assessment, as planned, will make great use of still and moving images, including computer simulations. Through images, students can interact with "materials" and "objects" that are dangerous, expensive, or simply unobtainable. Students will be able to take perspectives that are atomic or galactic or that are at places and times otherwise far removed from the child's experience.

We envision a new generation of assessment products that are instructionally useful—products in which testing and teaching complement and reinforce one another.

Assessment for Teachers

Assessment designed to serve the interests of individuals could provide teachers with greater information about pupil learning. From our conversations with teachers, we have learned that two kinds of information about pupils are valuable in making instructional decisions. The first is knowledge about the performance of a class as a whole. If this information were provided quickly, teachers could check to see if a large proportion of students had fundamental misconceptions that would warrant review of a concept. Alternatively, a particular misconception might be held by only a few students. A teacher with access to this information could provide tailored instruction or guide students to resources that would clarify their understanding.

Obviously, an assessment system like the one we describe could produce a great deal of data, which might overwhelm teachers who are already overloaded with information. A challenge, therefore, is to present teachers with the information they want in a *usable* form. One possible application is to combine assessment data with information about instructional materials available to the teacher. A sophisticated computer program might create an "editable" lesson plan that would specify learning objectives and materials needed. It could also suggest, for example, possible demonstrations, student assignments, questions,

and problems—all taking into account the resources and materials available in the school. A teacher would then accept or modify the lesson plan according to his or her professional judgment.

Planning Better Assessment

Assessment for learning won't happen automatically. To make assessment more pedagogically relevant, extensive research and development are needed. One area where more knowledge is needed is in the psychology of the school subjects.

During the last 10 years, a great deal of study has been devoted to how students learn scientific concepts. One line of research has shown that students sometimes progress through a series of understandings of a particular concept (Carpenter and Moser 1984, Vosniadou and Brewer 1987). If teachers know these levels for a given concept, they can trace how far along a path of mastery a student has traveled. Students often develop their own ideas, or "naive theories," to account for their observations. Spontaneously, or with the "help" of others, they pick up erroneous notions about why objects float or how big a star is. Knowledge of typical naive theories can help teachers organize instruction to deal with these misconceptions. If teachers know what is in the minds of students, they can optimize the fit of instruction to the beliefs and experiences of the students.

Another challenge is to understand better the nature of teaching and tutoring, especially within particular subject matter areas. If assessment is to have pedagogical value, then student responses, whether correct or incorrect, must be addressed skillfully. The ideal model might be a teacher who is working with students individually or in small groups. This teacher, if very skilled, can construct questions that are precisely targeted to a student's level of knowledge and experience. Skillfully posed questions are neither too hard nor too easy, so that with enough mental effort and persistence, the student has a good chance of answering correctly. The student's response might prompt a clue or follow-up question that will probe understanding more deeply, expand the concept, or illustrate the concept's various applications.

If assessment is to have pedagogical value, then student responses, whether correct or incorrect, must be addressed skillfully.

A third area of great concern is how innovation in testing can fit into the real world of schools. Our technical aspirations are difficult in themselves, but to try to make them work in a social and organizational sense is an even greater challenge. How, for example, will a larger amount of assessment time fit into the school schedule? Surely it will not if it has no instructional value. But even if we can help students learn through assessment systems, will they accept them? Will teachers see a new set of instructional tools as useful or as time-wasting gadgets? It is easy to be glib about how the computer age will transform our educational system, but many educators are justifiably skeptical about the role of technology in learning. We want to be sensitive to the realities of schooling and exercise great care to ensure respect for the central role of the teacher, the richness of good education, and the social forces that operate in schools.

A New Era of Testing

Advances in cognitive science, psychometrics, and computer technology have made possible a new generation of tests that better serve the interests of learners and teachers. That commitment to serving individuals is what drives the concepts described here, some of which have taken shape as prototypes. Full development and implementation of mastery assessment systems are still at least a few years away. As planned, they will present students with challenging questions and problems that are interconnected and related to the complex real world. These assessment systems will in turn serve as tools for teachers in selecting and modifying instruction tailored to students' levels of understanding. Immediate elaborated feedback, learning progress maps, and complex tasks will be combined by advanced technologies to offer students assessment with value for learning.□

References

Bunderson, C.V. (1988). *Measurement Science and Training*. Paper presented at the Invitational Theory-to-Practice Research Symposium, University of Minnesota.

Carpenter, T.P., and J.M. Moser. (1984). "The Acquisition of Addition and Subtraction Concepts in Grades One through Three." *Journal for Research in Mathematics Education* 15: 179-202.

Frederiksen, N. (1984). "The Real Test Bias: Influences of Testing on Teaching and Learning." *American Psychologist* 39, 3: 193-202.

Glaser, R.G. (1984). "Education and Thinking: The Role of Knowledge." *American Psychologist* 39, 2: 93-104.

McLean, L. (June 1986). *The Search for Pedagogical Relevance in Large-Scale Assessment*. Paper presented at the Invitational Seminar, International Approaches to Student Performance, London University Institute of Education.

Miller, A. (1986). *Imagery in Scientific Thought: Creating Twentieth Century Physics*. Cambridge, Mass.: MIT Press.

Vosniadou, S., and Brewer, W.F. (1987). "Theories of Knowledge Restructuring in Development." *Review of Educational Research* 57, 1: 51-67.

Michael E. Martinez is Associate Research Scientist, Division of Cognitive and Assessment Research, Educational Testing Service, Princeton, NJ 08541. **Joseph I. Lipson** is Principal Research Scientist at ETS and Professor of Communication, California State University, Chico, CA 95926.