9

Although instructional methods are moving in ever greater number to a multimedia base, testing is not. What principles should be considered in correcting this misalignment?

Aligning Paper Tests with Multimedia Instruction

Scott L. Howell

Although the "click-and-point" virtual classrooms of today hardly resemble the brick-and-mortar classrooms of yesterday, one thing seems not to have changed: the prevalence of paper-based tests. Paper-based tests have been the staple of education for centuries and will most likely persist for many years to come. This article explores some of the issues surrounding the growing chasm between the way students are now taught and how they are still tested from three perspectives: researcher, student, and teacher.

Should educators be concerned with the widening gap in how teachers teach, using rich multimedia in a multidimensional context, and how teachers test, on paper in grayscale colors in a two-dimensional setting? Should a new generation of students who communicate, learn, and almost live on keyboards be subjected to paper-based tests that require them to use less familiar handwriting tools, such as pens and number two lead pencils in an unfamiliar context—not on the computer? I believe educators should be concerned with the lag that prevails in aligning and changing assessment practices to the changed, and changing, instructional practices that now depend so much on computer-based multimedia environments.

Reliability and Validity

In assessment parlance, two words prevail in almost every discussion and analysis: reliability and validity. A test's reliability is a measure of how consistent the test measures each time administered, and its validity is a measure of accuracy in representing the intent or construct of the test instrument.

A casual dart game at home where the goal is to see who can get the most darts closest to the bull's-eye may be illustrative of these rather abstract concepts of reliability and validity and helpful in our discussion about paper tests in a computer-based instructional context. The proximity of the darts to the bull's-eye is the goal of the game or the "construct." The participant who successfully clusters the darts tightly will have high reliability or consistency; the participant who not only clusters the darts but does so nearest to, or even on, the bull's-eye is also the most accurate or valid. A test or a dart game can have high reliability with low validity, but one with high validity will also have high reliability.

Whether in dart play or test taking, all participants expect—independent of high, average, or low scores—reliable and valid scores that predict within some acceptable range of variability the appropriate level of mastery. However, the stakes are much higher in test taking than in dart playing, especially with their legal implications, when the measurements from these tests determine either directly or indirectly placement, college admission, and even employment.

Just as a dart player anticipates after weeks of practicing to see the same kind of dartboard at the "big" contest, so does the student in either a physical or virtual classroom studying a subject expect to see the same kind of subject matter presentation on the exam. If at the casual dart contest the darts are made of lighter or heavier material, the board resized, and the bull's-eye situated in the corner rather than the center as found previously, would anyone predict the results or measurements to be either reliable or valid on the day of the contest?

Much like the dart game is advancing to the electronic age with its new darts and electronically sensitized boards, so is the instructional experience changing from using only pencils, pens, and blackboards to the ubiquitous use of keyboards, computer screens, and software applications. Everyone would agree that practicing the dart game the "new" way but then showing up at a contest where the game was played the "old" way just would not be fair, but is this not exactly what is happening in many of our classrooms (physical and virtual) as students taught one way take tests another way?

What Are the Issues?

What are some of the issues with teaching online and testing on paper from the perspectives of a researcher, a student, and a teacher?

Researcher Perspective. This specific question of comparability between paper- and computer-based test results—assuming rich, multimedia instruction—has not been thoroughly studied. Not only are there few studies on this topic extant, but it is expected that there will never be many in the future because of research methods concerns and the commonsensical need for alignment between teaching and testing methods. Most researchers hesitate to conduct this comparison study because of

well-founded criticism that has mounted over media comparison studies in recent years, especially in the field of education. One of the most compelling reasons either for not researching this comparison or for discrediting the results of any comparability studies conducted is the difficulty of randomly selecting student participants in control and treatment groups, especially with all of the practical and ethical concerns associated with inaccurate test results. Nonetheless, one practical researcher has suggested that this issue still merits evaluation. Frederick (2003) has said that "as the environment changes for the delivery of instruction, it is important to reevaluate the ways in which we assess the learning outcome for students using this new format and develop and apply assessment techniques that are more consistent with the learning environment" (p. 17).

Student Perspective. The students of today are technologically savvy and sophisticated; they also are not patient with what they call old-fashioned teaching or testing practices. Students are accustomed to keyboarding as they communicate with friends using e-mail and instant text messaging; a toggle stick is a mere extension of their hand, and they have developed motor skills to prove it in computer game competitions with self and others; and most students have access to a computer either at home or at school where many of them learn, play, and do their homework. At the time of this writing, I attended a junior high awards ceremony, and the assistant principal mentioned as an aside that the ninth-grade student who was receiving an outstanding academic award also keyboarded at a rate of 130 words per minute.

As accustomed as today's students are to keyboards, toggle sticks, and computers, they are also as unaccustomed to using number two pencils to fill in test bubble sheets using optical character recognition technology and pens of yesteryear to write long (or even short) papers and essays. Although many students have refined the motor skills of "click, point, and drag" with a mouse, keyboard, or toggle stick and can experience little fatigue doing so, it does not take long for cramps, physical fatigue, and psychological discomfort to impair a student who is unaccustomed to writing by hand for extended periods of time with pencils and pens. Two researchers from Boston College have concluded that "recent research shows that written tests taken on paper severely underestimate the performance of students accustomed to working on computers. The situation is analogous to testing the accounting skills of modern accountants, but restricting them to the use of an abacus for calculations" (Russell and Haney, 2000, p. 2).

Teacher Perspective. Even though teachers may be involved as much as a fourth of their time doing some form of assessment, most have never received formal instruction on good test-writing practices. Training in such practices would help teachers improve their tests, but even then, little of the training would emphasize the need to incorporate the same multimedia used throughout the course in their tests. Furthermore, tests are often the one neglected part of the instruction process because developing good tests (and even poor tests) is a difficult and time-consuming process.

In recent years, school administrators have told teachers that they will "jump aboard" in the classroom (physical or virtual) with any of a number of technological initiatives that have swept their schools, districts, and universities. Many publishers have packaged textbooks and instructors' manuals with CD-ROMS, videos, and supporting Web sites to help "sell" the textbooks and enrich the instruction; yet, the exam banks and tests stay the same two-dimensional, static, grayscale items that they have always been. Some teachers have embraced these new technological and instructional methods willingly and have even taught themselves new skills necessary to keep up with—not ahead of—their own students' technological expectations; other teachers would if they could but lack the necessary training, support, and time necessary to acquire new skills; and still others refuse to abandon the habits and practices of many years of instruction, knowing that retirement cannot be far away. The unevenness that exists among the teachers in this regard also exists in different regions of the country for a complexity of reasons, from cultural to economic to political to others.

Even though there may not be parity across teachers' integration of technology, multimedia, animation, simulation, color, audio, and video into the curriculum, there is no question that this integration into instruction has occurred in the classrooms of today much more quickly than has the integration of the same into assessment. That this alignment between how teachers teach and how they test needs to occur is clear. What is not as clear is what the barriers are to change and how best to remove them!

Conclusion

Whereas the multimedia-enhanced instruction and the virtual classrooms of today no longer resemble the classrooms of even ten years ago, regrettably, the one instructional element that has stayed the same is the tests. Chalkboards have given way to computer projection screens, keyboarding has replaced handwriting, and multimedia simulations and animations now enhance many learning activities, but still tests are categorically unchanged in their paper-based format.

I suggest that the very pillars and fundamentals of sound test construction and delivery—reliability and validity—are threatened by this misalignment between teaching and testing methods. Researchers, students, and instructors alike agree, although from different perspectives, that urgent changes are needed to update and align the antiquated paper-based tests of the past to the rich multimedia- and computer-based teaching environments of today and tomorrow.

References

Frederick, P. "The Need for Alternative Authentic Assessments in Online Learning Environments." *Journal of Instruction Delivery Systems,* 2003, *16*(1), 17–20.

Russell, M., and Haney, W. "Bridging the Gap Between Testing and Technology in Schools." *Education Policy Analysis,* 2000, *8*(19). Retrieved June 7, 2004, from http://epaa.asu.edu/epaa/v8n19.html.

SCOTT L. HOWELL is adjunct professor of instructional psychology and technology at Brigham Young University in Provo, Utah.