



inTASC

TECHNOLOGY AND ASSESSMENT
STUDY COLLABORATIVE

Assessing Effects of Technology on Learning: Limitations of Today's Standardized Tests

Michael Russell & Jennifer Higgins
Technology and Assessment Study Collaborative
Boston College
332 Champion Hall
Chestnut Hill, MA 02467

www.intasc.org

Assessing Effects of Technology on Learning: Limitations of Today's Standardized Tests

Michael Russell & Jennifer Higgins
Technology and Assessment Study Collaborative
Boston College
Released August 2003

Michael K. Russell, Project Director/Boston College

Copyright © 2003 Technology and Assessment Study Collaborative, Boston College

Supported under the Field Initiated Study Grant Program, PR/Award Number R305T010065, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the Office of Educational Research and Improvement, or the U.S. Department of Education.



Assessing Effects of Technology on Learning: Limitations of Today's Standardized Tests

Michael Russell & Jennifer Higgins
Technology and Assessment Study Collaborative
Boston College

Over the past decade, students' use of computers has increased sharply, particularly for writing and research (Becker, 1999; Russell, O'Brien, Bebell, & O'Dwyer, 2003). At the same time, the use of large-scale tests to make decisions about schools and students has exploded. But, in addition to making decisions about students and their schools, results from large-scale tests are also used to assess the impact of computer-based technology on student learning. As one example, Wenglinsky (1998) used test scores from the National Assessment of Educational Progress (NAEP) mathematics test to examine the relationship between computer use during mathematics instruction and mathematics achievement.

Clearly, the use of standardized test scores to assess the effects of technology is attractive for at least two reasons. First, the public and political leaders tend to accept standardized tests as valid and reliable measures of student achievement. Second, these tests provide a common measure across large numbers of students. Yet, despite these advantages, using test scores from standardized tests to assess the impact of technology on student learning can be problematic.

Student Computer Use, Writing, and State Testing Programs

Despite regular use of computers by students to produce writing, state-testing programs currently require students to produce responses to open-ended questions using paper-and-pencil. However, research indicates that tests that require students to produce written responses on paper underestimate the performance of students who are accustomed to writing with computers.

In one randomized experiment, students accustomed to writing on computer were forced to use paper-and-pencil and only 30 percent performed at

a “passing” level; when they wrote on computer, 67 percent “passed” (Russell & Haney, 1997). In a second study, the difference in performance on paper versus on computer, for students who could keyboard approximately 20 words a minute, was larger than the amount students’ scores typically change between seventh and eighth grade on standardized tests. However, for students who were not accustomed to writing on computer, taking the tests on computer diminished performance (Russell, 1999). Finally, a third study demonstrated that removing the mode of administration effect for writing items would have a dramatic impact on the study district’s results. As Figure 1 indicates, based on 1999 Massachusetts Comprehensive Assessment System (MCAS) results, 19 percent of the fourth graders classified as “Needs Improvement” would move up to the “Proficient” performance level. An additional 5 percent of students who were classified as “Proficient” would be deemed “Advanced” (Russell & Plati, 2001, 2002).

Figure 1: Mode of Administration Effect on Grade 4 MCAS Results

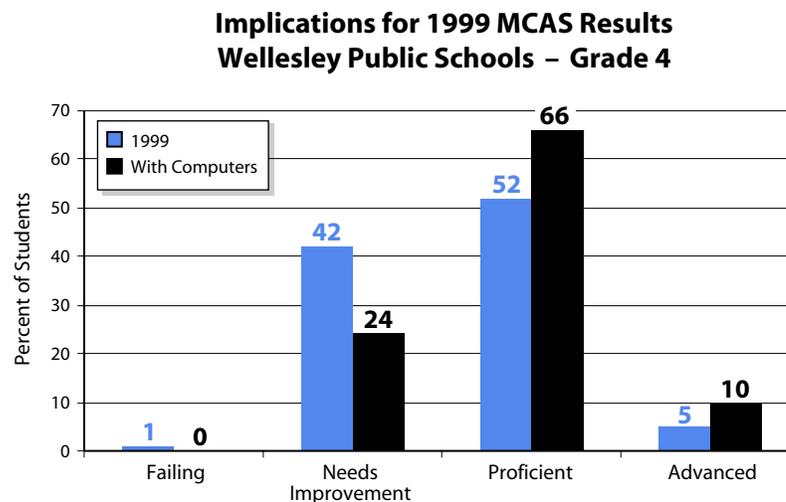


Figure from Russell & Plati (2001)

This body of research provides evidence that paper-based tests severely under-estimate the achievement of students who are accustomed to writing with computers. Given this mis-measurement, using paper-based tests to assess the effect of word processors on student writing skills is misleading.

State-Tests and Instructional Use of Computers

Beyond mismeasuring the performance of students and effects of computers on writing skills, recent evidence indicates that the mode of administration effect may also have negative consequences on instructional uses of computers. A national survey of teachers conducted by the National Board on Educa-

tional Testing and Public Policy (Pedulla et al., 2003) provides insight into the ways in which teachers believe they are changing their instructional practices in response to state-level testing programs. While many states have reported increases in students' scores on state tests (e.g., Texas, Massachusetts, and California among others have all celebrated gains over the past few years), for some students these gains come at the expense of opportunities in school to develop skills in using computers, particularly for writing (Russell & Abrams, in press). Although the majority of teachers report that they do not believe that the use of computers for writing has been affected by the testing program, 30.2 percent of teachers across the nation do believe that they are not using computers for writing because the state-mandated test is handwritten. Across the nation, a higher percentage of teachers in urban locations and in lower-performing schools, as compared to suburban and high-performing schools, believe they have decreased instructional use of computers for writing because of the format of the state test. Despite rising test scores, teachers and newly enacted school policies that decrease the use of computers, coupled with limited access to computers at home, under-prepare many students in urban and poorly performing schools for the workplace.

Beyond Traditional Assessment Measures

Writing is just one of several types of learning that computers can help students develop. Other areas of learning include: problem solving, research, non-linear thinking, understanding scientific concepts, spatial reasoning, statistics, and modeling complex relationships. Among all these areas, only one is extensively measured by current state-mandated testing programs, namely writing.

Arguably, some tests measure problem-solving skills, scientific understanding, statistics, and spatial relations. However, the number and types of items used to test students' achievement in these areas is insufficient for assessing the impact of computer use on these skills. As just one example, many third- and fourth-grade teachers use computers as a part of their mathematics instruction to help students develop spatial reasoning skills. However, on the fourth grade MCAS test, only 2 of the 39 items relate to spatial reasoning. It would be tenuous to use changes in MCAS scores to examine the impact computer use has on students' spatial reasoning.

Similarly, most mathematics tests include items that test students' mathematical problem-solving skills. Typically, these items take the form of word problems for which students must define a function that represents the relationship described, plug in the appropriate numbers, and perform accurate computations. While it is important for students to develop these mathematical problem-solving skills, these skills are not what advocates of computer use envision when they discuss the potential impacts of computers on students' problem-solving skills.

Problem solving with computers is more than just decoding text to define functions. As Dwyer (1996, p. 18) describes, when developing problem-solving skills with computers, “students are encouraged to critically assess data, to discover relationships and patterns, to compare and contrast, to transform information into something new.” To help students assimilate, organize, and present their learning, some teachers have students use HyperCard and other multimedia tools.

After studying HyperCard use in a small set of Apple Classrooms of Tomorrow (ACOT) classrooms, Tierney (1996) concluded: “Technology appears to have increased the likelihood of students’ being able to pursue multiple lines of thought and entertain different perspectives. Ideas were no longer treated as unidimensional and sequential; the technology allowed students to embed ideas within other ideas, as well as pursue other forms of multilayering and interconnecting ideas” (p. 176).

Despite the skill development enabled by multimedia authoring tools, students who develop complex products using these tools do not necessarily perform well on current tests. While studying the impact of computers on student learning, Baker, Herman, and Gearhart (1996) found that “...a sizeable portion of students who used HyperCard well to express their understanding of principles, themes, facts, and relationships were so-so or worse performers judged by more traditional forms of testing” (p. 198). Over the past decade these and similar findings have led proponents of computer use in schools to conclude that technology enables students to develop new competencies, “some of which were not being captured by traditional assessment measures” (Fisher, Dwyer, & Yocam, 1996, p. 5). While we support this conclusion, critics of computers in schools are beginning to see this argument as a well-worn cover for “lukewarm results” (Jane Healy as quoted by Westreich, 2000).

Time for New Assessments

It is time that testing and accountability programs develop and apply testing procedures that capture the types of learning impacted by computer use. To make this happen, educators and parents must demand that the way students are assessed matches the medium in which they typically work. Advocates for disabled students have long argued that state and local assessment programs should “allow students the same assistance in the assessment process as they have in the learning process...” and reason that “it is only fair that the assessment of what they have learned should allow for them to demonstrate their knowledge and skills in the way most appropriate to them” (Hehir, 2000, p. 50). The same argument applies to all students. Students who are accustomed to writing with computers should be allowed to write with computers while being tested.

In addition, educators and advocates of computer use in schools must insist that testing programs develop tests that measure students’ technology

skills. Despite the large investments schools have made in computer-related technologies, only two states collect information about students' technology skills. And, until recently, paper-based multiple-choice tests were employed in both states. A thorough examination of the impacts of computers on student learning must include measures of students' computer skills.

Instruments that measure the "other types of learning" possible with computers must also be developed. But, before these instruments can be developed, educators must be clearer about what these new types of learning are. It is not enough to say that computers allow students to develop problem solving, simulation, or modeling skills. Test development begins by defining the domain and constructs to be measured. Catch-phrases like problem solving, simulating, and modeling do not provide clear descriptions of a domain or construct. Descriptions of these "new types of learning" must become more precise.

Finally, we must anticipate that computer-related technology will continue to evolve faster than the technology of testing. We must narrow the gap between emerging computer-related technologies and new testing methods. Researchers must work more closely with teachers and students to predict how these new technologies might affect student learning and then work with test developers to develop instruments that measure these constructs before these new technologies have permeated large numbers of schools.

McNabb, Hawkes, and Rouk (1999) are correct: "Standardized test scores have become the accepted measure with which policymakers and the public gauge the benefits of educational investments" (p. 4). Acknowledging this as reality, educators must be proactive in establishing administration procedures and instruments that provide more accurate measures of the types of learning technology is believed to impact. Until these instruments and procedures are developed, testing programs will continue to mis-measure the impact of computers on student learning.



References

- Baker, E., Herman, J., & Gearhart, M. (1996). Does technology work in school? Why evaluation cannot tell the full story. In C. Fisher, D. Dwyer, and K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 185–202). San Francisco: Apple Press.
- Becker, H. J. (1999). *Internet use by teachers: Conditions of professional use and teacher-directed student use. Teaching, learning, and computing: 1998 national survey. Report #1*. Irvine, CA: Center for Research on Information Technology and Organizations.
- Dwyer, D. (1996). Learning in the age of technology. In C. Fisher, D. Dwyer, & K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 15–34). San Francisco: Apple Press.
- Fisher, C., Dwyer, D. and Yocam, K. (1996). *Education and technology: Reflections on computing in classrooms*. San Francisco: Apple Press.
- Hehir, T. (2000). Some assessments treat learning-disabled students unfairly. In D. Gordon (Ed.), *The Digital Classroom* (pp. 49–50). Cambridge, MA: Harvard Education Letter.
- McNabb, M., Hawkes, M. & Rouk, U. (1999). *Critical issues in evaluating the effectiveness of technology*. Report prepared for the Secretary's Conference on Educational Technology: Evaluating the Effectiveness of Technology, Washington, DC. Retrieved December 15, 2001, from <http://www.ed.gov/Technology/TechCong/1999/confsum.html>
- Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Boston, MA: Boston College.
- Russell, M. & Abrams, L. (in press). Instructional uses of computers for writing: How some teachers alter instructional practices in response to state testing. *Teachers College Record*.
- Russell, M. & Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, 5(3). Retrieved February 1, 2003 from <http://epaa.asu.edu/epaa/v5n3.html>
- Russell, M. & Plati, T. (2001). Mode of administration effects on MCAS composition performance for grades eight and ten. *Teachers College Record*. Retrieved February 1, 2003 from <http://www.tcrecord.org/Content.asp?ContentID=10709>
- Russell, M. (1999). Testing on computers: A Follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, 7(20). Retrieved February 1, 2003 from <http://epaa.asu.edu/epaa/v7n20/>

- Russell, M., & Plati, T. (2002). Does it matter with what I write? Comparing performance on paper, computer and portable writing devices. *Current Issues in Education*, 5(4). Retrieved February 1, 2003 from <http://cie.ed.asu.edu/volume5/number4/>
- Russell, M., O'Brien, E., Bebell, D., & O'Dwyer, L.(2003). Students' beliefs, access, and use of computers in school and at home. Boston, MA: Boston College, Technology and Assessment Study Collaborative. Retrieved March 10, 2003 from http://www.intasc.org/PDF/useit_r2.pdf
- Tierney, R. (1996). Redefining computer appropriation: A five-year study of ACOT students. In C. Fisher, D. Dwyer, & K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 169–184). San Francisco: Apple Press.
- Wenglinsky, H. (1998). Does it compute? The relationship between educational technology and student achievement in mathematics. Princeton, NJ: Policy Information Center, Educational Testing Service.
- Westreich, J. (2000). High-tech kids: Trailblazers or guinea pigs? In D. Gordon (Ed.), *The Digital Classroom* (pp. 19–28). Cambridge, MA: Harvard Education Letter.

