

COMPARISON BETWEEN TRADITIONAL
& COMPUTERIZED STANDARDIZED
TESTING IN SCHOOLS

Issues Paper for EDF 6432

William Cooper

April 9, 2009

Introduction

Which method is better at measuring the actual knowledge that a student has attained in a given subject? Is it through traditional paper and pencil testing? Or will it come from a test done on the computer? This paper examines the issue of testing and technology in America's schools. I will present findings from recent research studies which describes the advantages and disadvantages of each method.

This issue has grown in importance with each year as the use of technology in classroom learning activities has steadily increased. (Russell & Haney, 2000) stated that

The widening gap between the increased use of technology in schools and the absence of computers in state-level testing programs raises important implications for policies related to the use of both technology and testing in schools.

The primary purpose for ensuring that all schools have computers and utilize them is that it will improve both teaching and learning. Thus students will be prepared as early as possible to succeed in an increasingly technological workplace. In addition, the ever-increasing accountability demanded of schools through the No Child Left Behind law (NCLB) has resulted in all states implementing high-stakes standardized testing programs. Then, using the results of those tests as one of the primary factors both in incentive decisions for staff and also education decisions made about each student.

Between 1984 and 1993, the percentage of people using computers in the workplace has jumped from 24.6% to 45.8%. (Russell & Haney, 2000) During this period, schools went from one computer for every 125 students to one computer for every nine students during this same time frame. (Russell & Haney, 2000) With the influx of computers in the classrooms, schools have increased students' accessibility to CD-ROMs and the Internet. The ramifications are that schools are responding to the new demands of businesses and technology and are making a concerted effort to produce students who will be able to step into the jobs of the future.

Despite these facts, when the day comes to administer standardized tests, school districts eschew technology and revert to students using a pencil. They fill out their answers on a Scantron form. Thousands of instruction sheets and test forms are distributed and collected each year in each school district.

Advantages of Traditional testing format

School districts, especially at this present time with deep budget cuts that must be made, see that the cost to maintain the status quo is lower. The cost savings from not having to receive, distribute and collect the test booklets would not be offset by the initial added cost to increase the computing infrastructure to support computerized testing. This consists of more computer terminals, additional wiring and connectivity.

As for the test form itself, when a diagram, map or any type of graphic is needed for an individual question or series of questions, the paper form of the test allows for proper test construction. The graphic can be placed above the question text. (Bennett, Braswell, Oranje, Sandene, Kaplan, & Yan, 2008) Meanwhile, on the computer, the graphic is often placed to the side of the applicable text. A study done on results from a mathematics test comparing paper and pencil to computer format showed that three of four items requiring considerable change for computer rendering appeared to be more difficult than the paper counterpart. (Bennett, Braswell, Oranje, Sandene, Kaplan, & Yan, 2008) The paper test is better served for mathematics testing because the use of mathematic symbols is easier. Plus, the test taker does not have to take any additional steps to express the desired answer in the proper units. Units can be indicated in the question or the multiple choice options. A third advantage for traditional testing in mathematics concerns the use of the calculator. As long as all students use the same model of calculator, the comfort level with using one's own calculator is greater than utilizing the calculator application on a computer. Not all students may be familiar with the computer version of the calculator. Finally, when students need to perform the necessary calculations to a problem on scratch paper or the side of the form, it is easier for the student to look at the test booklet and/or answer sheet and fill in their response. The student is not constantly shifting back and forth between writing on paper and keying on a keyboard-while looking at a computer screen.

In assessing reading comprehension, a study by Choi and Tinkler in 2002 found that 3rd grade students had an easier time scanning the text for key words or cued phrases on the traditional form versus scanning the same passage on the computer screen. But, as students got older, the differences between the two methods shrank, as students gained more experience using the computer.

In summary, constructed-response test items for mathematics testing need to be adapted more than multiple-choice items in order to be rendered on the computer screen. Thus, it may be sometimes harder to hold the difficulty level constant for constructed-response than for multiple-choice questions. (Bennett, Braswell, Oranje, Sandene, Kaplan, & Yan, 2008)

Disadvantages of Traditional testing

From the school district, state or national perspective, the accountability measures that are in place call for increased measurement and measurement broken down into many layers. The standardized test has to be scheduled in a time window that maximizes student performance. Then the test results are needed as quickly as possible in order to grade the performance of the principals and teachers for incentive decisions. Finally, the results are used to make the needed educational decisions of promotion or graduation of the student. The traditional testing mechanism does not provide the levels of reporting by demographics as quickly and efficiently as a computerized test would.

As schools have pushed for testing to measure critical thinking than mere memorization, the test design itself is an outmoded cognitive-scientific basis. (Bennett, 2001) Bennett continues in his article on how the Internet will help large-scale assessment that there is a mismatch with the content and format of curriculum. This mismatch consists of forced-choice problems dominating many tests versus the increased attention paid to test preparation. In addition, there are some skills that cannot be measured through traditional means, but address logical reasoning, problem solving and critical thinking wants. One example is spatial memory. (Wainer, 1990) Finally, all students are asked the same questions. A computerized approach would allow questions that a student receives to be based on whether they got the initial questions on a concept correct. The degree of difficulty of the test would be a closer match to the student's level of mastery and knowledge. Thus, provide the needed challenge and diagnostic of their knowledge and abilities.

From the student point of view, as computer usage in the classroom continues to increase, conducting the standardized test with pencil and paper becomes more inconsistent with how students learn. During the test administration, students must fill in a bubble on an answer sheet. This takes more time to do than circling a response or pressing a key on a keyboard. Error is also involved in stray or incomplete markings will affect the questions incorrectly.

Finally, students whose native language is not English are at a disadvantage when it comes to standardized testing. Because tests have strict time limits, they will spend more time trying to read and understand the question.

Advantages to Computerized testing

From the school district perspective, a computerized standardized testing format is more efficient. Results are available almost immediately. Computer-based testing is viewed by many policymakers as a way to meet NCLB requirements. (Thompson, Thurlow, & Moore, 2003) Specifically, the need to produce itemized score analyses, disaggregation by school by demographic variables such as gender, race, ethnicity, income, disability or other disadvantaged status. The economic savings will not be immediate as there are conversion expenses. But over the long run, computerized testing will save money.

For the student, assessments that require an open-ended written response could generate better responses than those written longhand for students who are accustomed to writing on the computer. Russell & Haney (1997) reported this in a randomized experiment of 114 students who took math, science and language arts test. The finding occurred across all subjects on both short answer and extended answer items. A critical moderating variable is the student's typing speed. Those students with what are deemed to be moderate to fast typing speeds will score higher on the computer in a writing test. But students with low typing speeds will score worse on the computer. (Russell, 1999)

The majority of students prefer the computerized testing to traditional testing, regardless of their actual computer abilities. Students look upon any opportunity to be on the computer as golden, perhaps as a reward. If the student will feel more engaged and/or motivated, this should lead to better scores in the majority of cases.

Legislation over the past decade has required testing of disabled students. The computerized format opens up avenues for accommodating these students. These include large print and magnification of the type on the screen, audio presentation of instructions and test items and use of speech recognition software to dictate responses to the teacher. (Thompson, Thurlow, & Moore, 2003) In addition, Braille keyboards can be utilized for the visually impaired.

Technology is well-suited to supporting many of the data collection, analysis and feedback features needed for the formative use of assessment. (Quellmalz & Pellegrino, 2009) For instance, multiple forms of the test can be created easier on the computer, with the same degree of difficulty. Not all students would get the exact same questions. This would reduce testing time and the burden placed on the student. But NCLB regulations prohibit adopting a testing mechanism that adapts to the student's ability level. (Quellmalz & Pellegrino, 2009)

Disadvantages of Computerized testing

Not all schools have the technology infrastructure needed to effectively prepare students for the world that they will face, let alone for a standardized test. The differences in computer usage, speed, quality and quantity are significant within individual schools in the districts. For instance, low versus high socioeconomic schools. In addition, the current software needed to do a standardized test is complicated and cumbersome. As technology advances, the software will become less expensive and more user-friendly. (Honey, McMillan Culp, & Spielvogel, 1999)

Student performance will be greatly influenced by their level of comfort and familiarity with the computer. The activities done with computers during the regular school activities differ by socioeconomic status. Teaching in low income schools with the computer focuses on skills reinforcement and remediation. But in high income schools, the computer time is concentrated on analyzing information and presenting it to an audience. (Honey, McMillan Culp, & Spielvogel, 1999) Teachers face a couple of challenges when it comes to computerized testing. The ability to monitor students during the test is severely reduced. Also, the degree of

familiarity and usage of the computer in general during the regular school day and the software specifically will impact the teacher's comfort level with utilizing the computers on test day. Especially should there be a network or connectivity problem with some or all students. Additional staff would be needed to quickly resolve the problem. Finally, there is the issue of test security. Multiple layers of security are needed on test day to access the test and secure the tests when completed.

When it comes to administering the test, technology problems can be problematic for all parties involved. Students will likely get frustrated as not try as hard. Thus, they would perform more poorly. The reliability of scores may be compromised due to the deviations from standard procedures, along with temporary loss of the Internet connection, a hardware issue or a problem with the software. Schools must ensure that the item presentation and test scoring software works impeccably. (Wainer, 1990)

Computerized testing places a greater emphasis on certain skills that are not being measured by the standardized assessment. Specifically, typing, using the mouse, maneuvering on the keyboard. (Thompson, Thurlow, & Moore, 2003) Jones (2000) offered the following steps to be taken to address these concerns in his paper:

The user interface must be designed to minimize reliance on computer skills that are not part of the construct being measured by the assessment. Test sponsors can also reduce test taker anxiety by providing tutorials and CBT sample tests to candidates in advance of the testing session. Finally, the test delivery procedures should include adequate on-site support in the event that test takers need assistance during the testing session.

Test anxiety will be a significant factor in test performance should this be the first time that a school uses computerized testing. Thus, comparing performance to the previous year will be very difficult. As for the disabled student population, web developers do not typically design web pages such that they would be accessible. (Thompson, Thurlow, & Moore, 2003) The web pages hinder disabled students because the font is not large enough. Also, the student is not able to see everything on the screen when graphics and text are to be included. The web page would be easily adjustable.

Conclusion

Increasingly, schools are using computers to improve student learning. But, states continue to measure these learning outcomes using traditional paper and pencil testing methodology. There is incongruence between the learning and assessment modes used. The question becomes are we really measuring a student's true achievement level.

This paper detailed findings from several research studies. The fundamental conclusion to be gained is that the subject matter and the grade level of the students are the greatest determining factors in which mode to use. Mathematics and reading comprehension lends themselves more to paper and pencil tests due to concerns with the graphical interface. Essay exams are better suited for computerized tests. Computerized testing should be introduced at the middle school level for best results.

High-stakes testing continues to drive the agenda and the allocation of educational time and resources. Politicians and school boards are beginning to acknowledge the criticisms from educators that the current testing system fails everyone. Teachers do not get valuable or timely feedback to improve how to teach their students.

Technology cannot be fully implemented without an improved technology infrastructure throughout the entire school district. The state of the economy will dictate if and when districts implement computerized testing for part or their entire standardized testing program.

Skill development with the use of the computer must be done months in advance in concert with test preparation. An alternative approach to be studied is a mixed testing approach, where some subjects are traditional and some are paper. Should a mixed approach be adopted, then only one mode per day should be utilized.

Finally, when assessing students with a disability, school districts must take great care in the specifications needed for tests should they be administered via computer. Then, school districts must select a vendor with the necessary expertise to build in the needed accommodations for administering the test.

REFERENCES

- Bennett, R. E. (2001, February 14). How the Internet Will Help Large-Scale Assessment Reinvent Itself. *Educational Policy Analysis Archives* , 9 (5). Retrieved March 28, 2009, from ERIC database
- Bennett, R. E., Braswell, J., Oranje, A., Sandene, B., Kaplan, B., & Yan, F. (2008). Does It Matter if I Take My Mathematics test on Computer? *Journal of Technology, Learning and Assessment* , 6 (9). Retrieved March 28, 2009, from ERIC database
- Choi, S. W., & Tinkler, T. (2002). Evaluating comparability of paper-and-pencil and computer-based assessment in a K-12 setting. *National Council on Measurement in Education*. New Orleans. Retrieved March 28, 2009, from Google search.
- Honey, M., McMillan Culp, K., & Spielvogel, R. (1999). *Critical Issue: Using Technology to Improve Student Achievement*. North Central Regional Educational Laboratory. Retrieved March 28, 2009, from Google search.
- Quellmalz, E. S., & Pellegrino, J. W. (2009, January). Technology and Testing. *Science* , 323 (5910), pp. 75-79. Retrieved March 28, 2009, from Google search.
- Russell, M., & Haney, W. (2000). Bridging the Gap between Testing and Technology in Schools. *Education Policy Analysis Archives*, 8(19). (ERIC Document Reproduction Service No. EJ612353) Retrieved April 3, 2009, from ERIC database.
- Thompson, S., Thurlow, M., & Moore, M. (2003). *Using Computer-based Tests with Students with Disabilities*. University of Minnesota. Minneapolis: National Center on Educational Outcomes. Retrieved March 28, 2009, from ERIC database.
- Wainer, H. (1990). *Computerized Adaptive Testing: A Primer*. Hillsdale, NJ: Lawrence Erlbaum Associates.