J·T·L·A

The Journal of Technology, Learning, and Assessment

Volume 1, Number 1 · June 2002

Inexorable and Inevitable: The Continuing Story of Technology and Assessment

Randy Elliot Bennett

www.jtla.org

A publication of the Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College

Volume 1, Number 1

Inexorable and Inevitable: The Continuing Story of Technology and Assessment

Randy Elliot Bennett

Editor: Michael Russell
russelmh@bc.edu
Technology and Assessment Study Collaborative
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

JTLA is a free on-line journal, published by the Technology and Assessment Study Collaborative, Caroline A. and Peter S. Lynch School of Education, Boston College.

Preferred citation:

Bennett, R.E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *Journal of Technology, Learning, and Assessment, 1*(1). Available from http://www.jtla.org.

Abstract:

This paper argues that the inexorable advance of technology will force fundamental changes in the format and content of assessment. Technology is infusing the workplace, leading to widespread requirements for workers skilled in the use of computers. Technology is also finding a key place in education. This is occurring not only because technology skill has become a workplace requirement. It is also happening because technology provides information resources central to the pursuit of knowledge and because the medium allows for the delivery of instruction to individuals who couldn't otherwise obtain it. As technology becomes more central to schooling, assessing students in a medium different from the one in which they typically learn will become increasingly untenable. Education leaders in several states and numerous school districts are acting on that implication, implementing technology-based tests for low- and high-stakes decisions in elementary and secondary schools and across all key content areas. While some of these examinations are already being administered statewide, others will take several years to bring to fully operational status. These groundbreaking efforts will undoubtedly encounter significant difficulties that may include cost, measurement, technological-dependability, and security issues. But most importantly, state efforts will need to go beyond the initial achievement of computerizing traditional multiple-choice tests to create assessments that facilitate learning and instruction in ways that paper measures cannot.

J·T·L·A

Inexorable and Inevitable: The Continuing Story of Technology and Assessment[1]

Randy Elliot Bennett
Educational Testing Service, Princeton, NJ

The Inexorable

In the business world, almost everywhere one looks new technology abounds—computers, printers, scanners, personal digital assistants, and mobile phones, plus networks for these devices to plug into. Why is this?

First, since 1995 the United States has experienced dramatic improvements in productivity. Although information technology's contribution continues to be debated, evidence suggests that it is responsible for at least some of the effect (Federal Reserve Bank of San Francisco, 2000; McKinsey & Company, 2001; Oliner & Sichel, 2000). Technology may help increase productivity because manipulating information electronically is often cheaper than physically manipulating things (Negroponte, 1995). For example:

- Southwest Airlines expends $10 to sell a ticket through a travel agent, while it pays $1 to make an online sale (Abate, 2001). Because Southwest gets 30% of its passenger revenue through online sales, the savings are enormous.
- GE purchases large amounts of supplies and raw materials through online exchanges, generating an estimated savings of $600 million through lower prices and cheaper transaction costs (De Meyer & Steinberg, 2001).
- Wal-Mart has extensive databases that give real-time information on individual store inventories to the chain's 10,000 suppliers, who can refill inventory as needed (Steinberg, 2001). Among other things, this system saves Wal-Mart the cost of buying and storing more inventory than it can quickly sell.
- Cisco Systems must train thousands of sales and technical support staff each time a new product is launched. Since those employees are deployed worldwide, the travel costs alone are huge. Through e-learning, Cisco has reported savings of 40–60% over instructor-led courses (Cisco Systems, 2001).

Besides aiding efficiency, technology can break down traditional barriers, enhancing the value of products and services. One such barrier was the age-old relationship between richness and reach (Evans & Wurster, 2000). It used to be that a business could either reach many customers with a relatively spare message—say, through broadcast advertising—or engage a few customers in a deeper interaction (e.g., via a showroom encounter). The Internet has fundamentally changed that traditional relationship by permitting companies to reach many individuals in a personalized way. Customers can use the Web to order a computer built to particular specifications (http://www.dell.com), design their own sneakers (http://www.customatix.com), configure office furniture to suit their needs (http://www.hermanmiller.com), or purchase pants to fit (http://www.landsend.com).

The value that technology provides to businesses and consumers is continually increasing. One illustration is Moore's Law: Since the late 1960s, computational capability has doubled every 1.5–2 years (Kurzweil, 1999; Negroponte, 1995). The impact has been to dramatically expand the tasks computers can perform and simultaneously drive down hardware prices. A second illustration is Metcalfe's Law, which states that the worth of a network increases by the square of the number of participants (Evans & Wurster, 2000). This means that as more organizations and individuals gain access to the Internet, the more useful that access becomes (e.g., consider e-mail).

So, what's inexorable? The relentless advance of technology, not only in terms of its ever-growing capability but, more importantly, its steady infusion into the world of work.

The Inevitable

The inexorable (and rapid) advance of technology in business has important implications for the workforce. For one, it suggests that the ability to use technology will become a standard job-entry requirement. In fact, it's estimated that by 2006 almost half of all U.S. workers will be employed by industries that are either major producers or intensive users of IT products and services (Henry et al., 1999).

But learning how to use new technology once is not enough. Businesses that currently use technology routinely upgrade to remain competitive. Those businesses that do not yet use technology eventually will. Thus, job entry requirements are likely to increase continually, as they have for the past several decades (Moe & Blodget, 2000).

In addition to this rise in entry qualifications, the knowledge required to maintain a job in many occupations is changing so fast that 50% of all employees' skills are estimated to become outdated within 3–5 years (Moe & Blodget, 2000). Therefore, those who are not conversant with technology and skilled at rapidly learning the next technology will be at a disadvantage in finding and keeping a job.

Of course, being skilled in a global economy is not simply about knowing how to use new technology. It also concerns knowing how to solve problems; learn new things; work on a team; communicate; and locate, evaluate, and act on information. But, ultimately, it may be most about *combining* those more traditional skills with technological ones to get work done.

How do we ensure that our population is able to develop and continually renew the competencies needed for success in a global economy? Several blue ribbon panels have recently studied this question and come to similar conclusions. Interestingly, their conclusions aren't restricted to workforce training but extend to education generally. The Commission on Technology and Adult Learning (2001), sponsored by the American Society for Training and Development and the National Governors Association, stated the following: "The Commission … encourages governors, CEOs and other leaders to make e-learning the cornerstone of a national effort to develop a skilled workforce for America's digital economy…. By embracing e-learning in our states, our communities and our organizations, we can improve our competitiveness and point the way to a new era of unprecedented growth and opportunity for all Americans." Similarly, the National Association of State Boards of Education (2001) concluded: "E-learning will improve American education in valuable ways and should be universally implemented as soon as possible." The President's Information Technology Advisory Committee Panel on Transforming Learning (2001) recommended that "…the federal government set as a national priority the effective integration of information technology with education and training." Finally, the bipartisan Web-based Education Commission's Report to the President and Congress (Kerrey & Isakson, 2000) concluded, "The question is no longer *if* the Internet can be used to transform learning in new and powerful ways. The Commission has found that it can. Nor is the question *should* we invest the time, the energy, and the money necessary to fulfill its promise in defining and shaping new learning opportunity. The Commission believes that we should."

The statements of these panels certainly have an air of what Alan Greenspan might term "irrational exuberance." Is there any indication whatsoever that the e-learning vision implied by these groups has the least chance of coming true?

Maybe there is. For example, with respect to postsecondary education, 35 states now have virtual universities (or other organizations) to deliver or promote Web-based distance learning (Young, 2001). Are students enrolling? Some online institutions, in fact, have already reached traditional big-campus levels: the University of Maryland's University College has 26,500 pupils (Shea, 2001), the University of Phoenix Online has 25,000 students (Konrad, 2001), and the SUNY Learning Network reports offering over 2,000 courses to an enrollment of 25,000 individuals (*State University's Premiere,* 2001).

E-learning's potential to bring postsecondary education to even more diverse audiences is also being explored. The Army recently awarded a $450 million contract to create eArmyU, a consortium of institutions designed to allow soldiers to get degrees online (Emery, 2001; Schwartz, 2001). EArmyU is now operational, enrolling 12,000 students in its first year (Arnone, 2001). Internationally, the People's Republic of China plans to have 5 million students in 50–100 online colleges by 2005 (Walfish, 2001). Also, the World Bank and Australia have announced plans to spend $750 million to bring Internet distance learning to developing countries (Maslen, 2001; World Bank Group, 2001).

Electronic resources are playing a growing role for students attending class on campus, too. As of September 2001, 96% of students between the ages of 18 and 24 reported using a computer, with 86% of students indicating that they did so in school (U.S. Department of Commerce, 2002). This pervasive use is not just a matter of personal preference: Web resources are a component of the syllabus in over 40% of all college courses (Campus Computing Project, 2000). Several major institutions are aiding the trend: MIT has begun a 10-year, $100 million effort to put all of its course material online, which includes some 2,000 courses (Schwartz, 2001). A 500-course "pilot" is already underway. The University of Phoenix, the nation's largest private postsecondary institution, is eliminating paper texts, with 60–80% of classes to be "bookless" by Spring 2002 (Blumenstyk, 2001). How can Phoenix do this? Simple: the "Big 5" publishers already offer over 850 textbook titles in digital format.

Given the activity at the postsecondary level, it should be no surprise that e-learning is migrating downward. At least 13 states have established virtual high schools to serve districts without qualified staff, home-schooled students, and children of migrant workers (Carr, 1999; Carr & Young, 1999; Kerrey & Isakson, 2000). Apex, a for-profit virtual high school that specializes in advanced placement courses, claims agreements with 29 states representing 15,000 high schools (*Apex Learning Instructor,* 2001).

Not surprisingly, the teaching force is catching on: Nationally representative data suggest that, in Spring 1999, 66% of public school teachers were using computers or the Internet for instruction during class time (National Center for Education Statistics, 2000). In what specific ways were they using them? Teachers said they assigned students work to a large or moderate extent that involved word processing, spreadsheets or other applications (41% of teachers), practice drills (31%), and research using the Internet (30%). In a Spring 2001 survey, Quality Education Data found that 90% of K–12 teachers indicated using the Internet as a professional resource (Quality Education Data, 2001). They employed it most frequently for research (89% of teachers), communication with colleagues (85%), and professional development (73%).

What about students? An Education Department survey indicates that, in the fall of 2000, 77% of public-school instructional rooms contained computers connected to the Internet and that the ratio of students to those computers was 7:1 (National Center for Education Statistics, 2001). Are students using those computers? The U.S. Department of Commerce (2002) reports that, as of September 2001, 75% of children aged 5–9 and 85% of those 10–17 used a computer at school.[2] For late elementary through high school, lower, but still considerable, percentages used the Internet—45% of those aged 10–13 and 55% of those 14–17.

What do students use computers for? In a nationally representative sample taking the 2001 NAEP U.S. History assessment, 82% of 8th grade and 86% of 12th grade students indicated using a word processor to write reports at least to some extent; 73% and 77%, of those respective groups reported using the computer for research (Lapp, Grigg, & Tay-Lim, 2002). Similar results appear in *Education Week's* 2001 technology survey, where an overwhelming majority of grade 7–12 students said they employed a computer for school research or for writing papers (*Technology Counts,* 2001). The significant use implied by these various data sources will be further encouraged by the *No Child Left Behind Act of 2001*. Among other things, this landmark legislation appropriates $700 million in fiscal year 2002 to improve student achievement through the use of technology in elementary and secondary schools (U.S. Department of Education, 2002).

To be sure, e-learning and, more generally, educational technology are not without critics and not without failures. Critics have argued that empirical support for the effectiveness of e-learning is inconclusive, that technology has not improved teaching or student achievement, and that continued investment is unjustified (e.g., Cuban, 2000, 2001).[3] Although this argument raises very important concerns, its conclusion does not necessarily follow. First, it may be true that computers have not changed how teachers instruct or how much students learn (at least in the ways we currently measure achievement). However, the *tools* students use to learn certainly are changing, as the data cited above suggest. There is no question that the need to learn these tools is, in part, driven by their pervasive position in the world of work.[4] But these tools are also becoming part of the equipment of 21st century scholarship and, consequently, necessary for college-bound students generally. Knowing how to do intellectual work *with* technology—to model a problem using a spreadsheet, create a presentation, use data analysis tools, find information using the Internet, or write and revise a paper with a word processor—is becoming a critical academic skill.

Second, even if e-learning is no better than traditional instructional methods it may still make for a wise investment in some circumstances. For instance, many individuals can't get the education they desire from local sources. These individuals include adults whose work or family responsibilities prevent them from physically attending class on a set schedule, as well as younger students whose school districts don't have staff qualified to teach certain specialized courses. For these students, it's either electronic access to a desired educational experience or no access at all.

Besides critics, e-learning has had its share of failures. The speculative dot.com bubble that burst in late 2000 hit online learning ventures, too: Temple University pulled the plug on its for-profit distance learning company; Caliber and Pensare went bankrupt; NYUonline and the U.S. Open University announced that they would shut down; UNext laid off staff; and Columbia's Fathom changed its approach (Arnone, 2002a, 2002b; *Caliber Files,* 2001; Carnevale, 2001; Mangan, 2001a).

Still, it is hard to dismiss the widespread indications that technology is beginning to play an important role in education. At the very least, to quote Bob Dylan, "Something is happening here...."

The Implications for Assessment

Increasingly, people have to know how to use technology to work and to learn. Thus, technology is becoming a *substantive* requirement in its own right. The emergence of technology as a necessary skill means there will be tests of it. For those who specialize in computing, such tests have become common for certifying job proficiency (Adelman, 2000). But as working and learning begin to require technology competence of almost everyone, assessing these skills will become routine.

Perhaps more important for assessment, however, is that technology is also becoming a *medium* for learning and work. The CEO Forum on Education and Technology (2001) suggests, "...as schools...integrate technology into the curriculum, the method of assessment should reflect the tools employed in teaching and learning." At the least, a mismatch between the modes of learning and assessment could cause achievement to be inaccurately estimated (Russell & Haney, 2000). Writing presents a good example: More and more, students are using the computer to complete composition assignments; however, research suggests that testing these students on paper underestimates their proficiency (Russell & Plati, 2001). The bottom line is that, as students come to do the majority of their learning with technology, asking them to express that learning in a medium different from the one in which they routinely work will become increasingly untenable, to the point that much of the paper testing we do today will be an anachronism (Bennett, 2001).

Acting on the Implications

Education leaders in several states and numerous school districts have recognized the inevitability of technology-based assessment. Table 1 (page 10) gives a summary of the state-level activity involving current or planned operational delivery of such tests. (Practice tests or pilot projects with no current commitment to operational delivery are not included.) Several points should be noted. First, states are implementing technology-based tests in elementary and secondary schools and in all of the key content areas (i.e., reading, math, science, English, social studies). For example, Georgia is creating a testing program at both levels, each program covering a range of subjects. Second, states plan to deliver both low- and high-stakes examinations through this medium. Some tests, like Virginia's Algebra Readiness Diagnostic Test, are intended for instructional purposes, whereas its eSOL (Standards-of-Learning) tests will be used for making graduation decisions. Third, some of these examinations are *already* being administered statewide, whereas others will take several years to bring to fully operational status. For instance, South Dakota, which has broadband access in *every classroom* and a 2:1 ratio of students to Internet-accessible computers, began administering the Dakota Assessment of Content Standards for all students in Grades 3, 6, and 10 exclusively via the Internet in April 2002; there is no paper version of the test. Other states, such as Georgia and Oregon, will offer assessments in both Web and paper format until electronic delivery can be made universal. Fourth, the tests generally use multiple-choice items exclusively, though the intention typically is to move in later versions to more complex tasks that better represent the instructional uses of technology that schools are moving toward (e.g., Internet research, writing on computer, data modeling). Finally, in some cases it is clear that electronic assessment is part of an integrated state plan to employ technology throughout the educational process. Virginia has as an explicit goal the online delivery of instructional, remedial, and testing services. South Dakota is well along toward realizing a similar vision, having intensively trained over 40% of its teachers to use technology in the curriculum and having connected its schools, libraries, and postsecondary institutions into a high-speed data/video network.

Table 1: State Activity in K–12 Technology-Based Assessment

State	Test	Grades and subjects	Stakes	Implementation date	Comments
South Dakota (DECA, 2002)	Dakota Assessment of Content Standards	Mandatory in Grades 3, 6, & 10 in reading and math	Low (for measuring student progress toward state standards)	Implemented statewide as of Spring 2002 after administering over 70,000 pilot exams	1. Web-based, adaptive, multiple-choice CRT 2. Student to Internet-computer ratio of 2:1 3. All classrooms have broadband access 4. No paper counterpart 5. Extensive teacher training in using technology for instruction
Oregon (ODE, 2002)	Technology Enhanced Student Assessment (TESA)	Grades 3, 5, 8, & 10 in reading, math, and language arts in current phase; science and social studies to follow	Low (for measuring progress toward state Certificate of Initial Mastery)	Available in 300 schools in 2001–2002; available statewide in 2003–2004	1. Web-based, adaptive, multiple-choice 2. Students may take TESA or a paper test 3. Available for students with IEPs
Virginia (Virginia Department of Education, n.d., 2001)	eSOL (Virginia Standards-of- Learning Web-based Assessments)	High school math, science, social studies, English	High (for graduation beginning with class of 2004)	In all high schools by 2004	1. Overall goals include 5:1 student to Internet-computer ratio & broadband delivery of instructional, remedial, and testing services 2. Web-based, multiple-choice test
	Algebra Readiness Diagnostic Test	Grades 6–9 in algebra readiness	Low (as pretest for instructional purposes and as post-test to document growth)	Statewide in 2002–2003	1. Adaptive, multiple-choice, diagnostic, Web-based test 2. Linked to Standards of Learning

Table 1 continues on page 11

Table 1 State Activity in K–12 Technology-Based Assessment *(continued)*

State	Test	Grades and subjects	Stakes	Implementation date	Comments
Georgia (Georgia Department of Education, n.d., 2001)	Criterion-Referenced Competency Tests	Grades 1–8 in reading, math, English, language arts; Grades 3–8 in science and social studies	Low (for student practice and classroom testing) High (for promotion in Grades 3 and 5–8)	Spring 2002 in paper and Web formats	1. Web-based, criterion-referenced, 3-level item bank for (a) student use (b) teacher use (c) state use 2. Multiple item types at Level 1 & 2 3. Multiple-choice initially at Level 3
	End-of-Course Tests	High school math, science, social studies, English	Low (for student and teacher use) High (for graduation)	Spring 2003 in paper and Web formats; complete phase-in of Web format by Spring 2005	1. Web-based, criterion-referenced, 2-level item bank for (a) student, teacher, and parent use (b) state use 2. Multiple item types at Level 1 3. Multiple-choice at Level 2
Idaho (Olson, 2002)	Achievement Level Tests	Grades 2–9 in math, reading, & language arts	Not yet determined	Spring 2003	1. Web-based, adaptive, multiple-choice
Utah (State of Utah, 2001)	Core Assessment Science Series	Science in Grades 4–8 and high school	Low (for measuring progress against state standards)	2003–2004	1. CRT, multiple-choice
North Carolina (State Board of Education, 2001)	North Carolina Computerized Adaptive Testing System	Grades 3–8 reading and math	High (for promotion)	2001–2002	1. Adaptive, Web-based, multiple-choice 2. Offered only as alternative to state paper tests for selected students with disabilities
Maryland (Maryland State Department of Education, 2001; Wise 1997)	Maryland Functional Testing Program	Reading and math; usually taken starting in Grades 6–7	High (for high school graduation)	Operational since early 1990s; revised version due Fall 2002	1. Adaptive, multiple-choice 2. Paper version also available

Practical Concerns

These leading-edge states are, of course, not alone in their attempts to move testing programs to computer. Over the past decade, many occupational and professional, as well as postsecondary-education tests, have done so. Examples include the College-Level Examination Program® (CLEP®), Graduate Record Examinations® General Test (GRE®), Graduate Management Admission Test® (GMAT®), National Council of State Boards of Nursing NCLEX® Examination, Test of English as a Foreign Language® (TOEFL®), and United States Medical Licensing Examination™ (USMLE™). The collective experience of these programs is that computer delivery presents considerable challenges (Wainer & Eignor, 2000).

Cost

The early entrants into computerized testing bore the cost of creating a computer-based test-center infrastructure, electronic tools for writing items, presentation software, and the large item pools needed to support continuous high-stakes testing. K–12 education will benefit enormously from the fact that (1) hardware is much cheaper than it was a decade ago, (2) integrated test authoring and delivery software is now readily available, (3) a universal electronic delivery network built with other people's money now exists (the Internet), and (4) because it has broad instructional use, the local-school space, software, hardware, and Internet connections used for testing can be expensed through other budgets. Still, the required investment and operational costs will, at least initially, be large relative to paper testing. Among other things, these costs will include vendor charges for testing software, central servers to house test content, training, and technical support. Although some analyses (e.g., Neuburger, 2001) suggest that the savings from eliminating such paper processes as printing and shipping will eventually outweigh these costs, state budget deficits from recession and the unplanned expense of homeland defense may cause financial support for computer testing efforts to waver.

How can these costs be met? Perhaps the least effective approach would be for each state to create it own system, duplicating the efforts of others with similar population characteristics and education needs. Instead, consortia, cooperative agreements, or buying pools for obtaining test questions, telecommunications equipment, computer hardware, testing software, and equipment maintenance should be considered.

Measurement and Fairness Issues

The experience of the early entrants also suggests that there will be nontrivial measurement and fairness issues. For those K–12 agencies that plan to offer the same tests on paper and computer, comparability will be a concern, especially for high-stakes decisions. Although comparability has often been supported (Bridgeman, 1998; Mead & Drasgow, 1993; Schaeffer, Bridgeman, Golub-Smith, Lewis, Potenza, & Steffen, 1998; Schaeffer, Steffen, Golub-Smith, Mills, & Durso, 1995), in some instances it does not hold, as when examinees are tested in a mode different from the one in which they routinely work (Russell & Plati, 2001). Further, while it is desirable from a fairness perspective, comparability may limit innovation by preventing the computer-based version from exploiting the technology to broaden measurement beyond what traditional methods allow. Where feasible, the wise choice may be to eliminate the problem entirely by offering only a computer-based instrument, as South Dakota has done, or by delivering in the two modes for as short a period as possible.

Regardless of whether the test is delivered solely on computer, there is a second comparability concern. This concern is for "platform" comparability. From one school to the next (and even within the same school), monitor size, screen resolution, keyboard layout, connection speed, and other technical characteristics may vary, causing items to appear differently or to take more time to display. (Display-time variations may occur for the same machine as a function of time of day.) Any of these variations may affect scores unfairly. For instance, Bridgeman, Lennon, and Jackenthal (2001) found that lowering the screen resolution, and thus increasing the need for scrolling, diminished test performance on reading comprehension items by a small (but nontrivial) amount. Similarly, Powers and Potenza (1996) presented evidence to suggest that essays written on laptops might not be comparable to those written on desktops having better keyboards and screen displays.

A third measurement issue is differential computer familiarity. Although physical access to computers at school differs little by income and racial group, home-access disparities are still substantial (U.S. Department of Commerce, 2002). Thus, there may be group (or individual) differences in computer experience that affect test performance in construct-irrelevant ways. For multiple-choice tests, the research to date suggests that differences in computer experience have little, if any, effect on test scores (e.g., Bridgeman, Bejar, & Friedman, 1999; Taylor, Jamieson, Eignor, & Kirsch, 1998). However, as electronic tests incorporate more performance tasks, the complexity of the mechanisms for responding—and the demands on computer facility—could well increase.

Technological Dependability

As we all know, and as the large, high-stakes testing programs have found, computers do not always work as intended (e.g., Mangan, 2001b). Such glitches are conceptually similar to printing and shipping errors in paper programs that cause examinees to receive no test, take an erroneous test, or complete a correctly printed test that never reaches its destination. Though any such event is unacceptable, for paper and electronic programs, these occurrences are, in fact, extremely rare. For K–12 assessment programs, the situation may be initially more troublesome because computer delivery is a new business for most vendors and because schools do not always have ready access to the onsite technical staff needed to fix problems quickly.

Security

High-stakes electronic testing entails security problems that are not very different from those of paper programs; in particular, items can be stolen or examinee data tampered with regardless of delivery mode. Rather, it is more the specific methods of wrongful access and prevention that differ. For example, electronic delivery depends on such preventatives as encryption, firewalls, and controlled software access. In paper delivery, security depends on tracking shipments and keeping materials under lock and key. (In either paper or digital form, once stolen, item content can be shot around the world at the touch of a button. For paper, there is only the extra step of first scanning the pilfered document.) The challenge for K–12 programs, then, is not that security threats will necessarily be greater, but that staff must manage new methods of detection and prevention.[5]

Conclusion

This paper has argued that the advance of technology is inexorable in at least two ways. First, technological capability is increasing exponentially. Second, new technology is pervading our work, and it is beginning to infuse learning.

The paper also argued that the incorporation of technology into assessment is inevitable because, as technology becomes intertwined with what and how students learn, the means we use to document achievement must keep pace. However, it is similarly inevitable that this incorporation will not be easy. There are still enough open issues, especially of cost and measurement, that at least some significant setbacks will occur. But even if all of the existing issues were resolved, the history of technology is one of unanticipated consequences that are not always positive.

Given the dangers, one can see why some states chose to begin the transition with low-stakes assessments. The decisions based on these tests can tolerate lower levels of measurement quality, technological dependability, and security; moving too quickly to high-stakes tests would maximize risk—political, financial, legal, and educational. Similarly, the use of multiple-choice questions is very sensible. They can be easily presented on-screen and require little computer skill for responding. Incorporating significant numbers of performance tasks at this stage would raise costs, demand more sophisticated presentation software, and increase the potential for construct-irrelevant variance in responding.

What the states are doing now, however, must be only a beginning. If all we do is put multiple-choice tests on computer, we will not have done enough to align assessment with how technology is coming to be used for classroom instruction. Sadly, our progress in using the computer to improve assessment has been limited. Almost a decade ago, we moved the first large educational tests to computer, fully intending to use technology to introduce new measurement approaches. These efforts got as far as adaptivity and then due to cost, technical complexity, the need to maintain scale, and the sufficiency of multiple-choice for summative decision-making, moved no farther. Fortunately, K–12 agencies have educational responsibilities that may force them to go beyond the initial achievement of computerization to create assessments that support learning and instruction in ways that paper tests cannot. Researchers can help them meet this challenge by discovering how to cost-effectively design coherent systems of assessment that have both summative and formative components (Pellegrino, Chudowsky, & Glaser, 2001). These systems might include simulations and other complex performances that not only indicate achievement level, but offer proficiency inferences with clear instructional implications. Creating such systems will be a difficult challenge, but it is aided by an emerging science of assessment design (Mislevy, Steinberg, Almond, Breyer, & Johnson, 2001; Pellegrino, Chudowsky, & Glaser, 2001).

To be perfectly clear, it is not at all inevitable that we will incorporate technology into assessment in ways that bring lasting educational benefit. The question is no longer *whether* assessment must incorporate technology. It is how to do it responsibly, not only to preserve the validity, fairness, utility, and credibility of the measurement enterprise but, even more so, to enhance it. In this pursuit, we must be nothing less than inexorable.

Notes

1 This paper was presented as part of the symposium, The Impact of Technology on Assessment: New Opportunities for Knowing What Students Know (J. Pellegrino, Chair), American Educational Research Association, New Orleans, April 2002. I thank Dan Eignor, Drew Gitomer, and Ellen Mandinach for their helpful comments on an earlier draft of this paper.

2 These figures combine those who use computers only at school and those who use computers at school and home. The values are calculated from Figure 5-4, Computer Use by Age and Location, 2001.

3 *No Child Left Behind* authorizes up to $15 million for an independent, long-term study of the effectiveness of educational technology, including the conditions under which it increases student achievement (Educational Testing Service, 2002). The study's final report is due by April 2006.

4 There is certainly a clear and influential economic dimension to the argument for using computers in schools. This argument has become critical only in the past few decades with the emergence of the global economy. The genesis of the current school reform movement is often linked to the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983), which opens with these words:

> "Our Nation is at risk. Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world."

Since the publication of that report, almost 20 years ago, the perceived threat has not diminished. In a report titled, *School Technology and Readiness,* the CEO Forum on Education and Technology (2001) puts it this way:

> "Student achievement must be improved in order to prepare students to succeed in the global economy. Many observers liken the need for a world class, high-quality educational system to a national security issue. The United States can only remain a leading power in the global economy if it continues to ensure students will be prepared to thrive in the future."

The concern of both these reports is in not having a workforce skilled and productive enough to compete with those of other nations. A skilled and productive workforce keeps jobs at home by encouraging both domestic and foreign businesses to invest here. That, in turn, helps us maintain a high standard of living and a tax base strong enough to support an effective national defense. And, as suggested, the ability to use technology in conjunction with other competencies helps make for a skilled and productive workforce.

5 Security will be more costly and complex if testing is done continuously. Continuous testing may call for larger item pools, methods of controlling item exposure, and the frequent rotation of pools in and out of service.

References

Abate, C. (2001, September 1). The smart business 50: Southwest Airlines. EWeek. Retrieved April 19, 2002, from http://www.eweek.com/article/0,3658,s=1884&a=10418,00.asp

Adelman, C. (2000). *A parallel postsecondary universe: The certification system in information technology.* Retrieved April 19, 2002, from the U.S. Department of Education Web site: http://www.ed.gov/pubs/ParallelUniverse

Apex Learning instructor earns national board certification. (2001, December 13). [Press release]. Retrieved April 19, 2002, from http://apexlearning.com/about/press_room/articles/pr_2001_dec13_certification.asp

Arnone, M. (2001, December 10). Army's distance-education project seeks more colleges as participants. *The Chronicle of Higher Education.* Retrieved April 19, 2002, from http://chronicle.com/free/2001/12/2001121002u.htm

Arnone, M. (2002a, February 5). United States Open University announces it will close in June. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2002/02/2002020501u.htm

Arnone, M. (2002b, February 8). Fathom adds corporate training to its distance-education offerings. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2002/02/2002020801u.htm

Bennett, R. E. (2001). How the Internet will help large-scale assessment reinvent itself. *Education Policy Analysis Archives,* 9(5). Retrieved April 19, 2002, from http://epaa.asu.edu/epaa/v9n5.html

Blumenstyk, B. (2001, May 18). Publishers promote e-textbooks, but many students and professors are skeptical. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/v47/i36/36a03502.htm

Bridgeman, B. (1998). Fairness in computer-based testing: What we know and what we need to know. In *New directions in assessment for higher education: Fairness, access, multiculturalism, and equity* (GRE FAME Report Series Vol. 2). Retrieved April 19, 2002, from the Educational Testing Service FTP site: ftp://ftp.ets.org/pub/gre/241343.pdf

Bridgeman, B., Bejar, I. I., & Friedman, D. (1999). Fairness issues in a computer-based architectural licensure examination. *Computers in Human Behavior,* 15, 419–440.

Bridgeman, B., Lennon, M. L., & Jackenthal, A. (2001). *Effects of screen size, screen resolution, and display rate on computer-based test performance* (RR-01-23). Retrieved April 19, 2002, from the Educational Testing Service FTP site: http://www.ets.org/research/dload/RR-01-23.pdf

Caliber files for Chapter 11. (2001, June 15). *Baltimore Business Journal*. Retrieved April 19, 2002, from http://baltimore.bizjournals.com/baltimore/stories/2001/06/11/daily41.html

Campus Computing Project. (2000). *The 2000 national survey of information technology in U.S. higher education: Struggling with IT staffing*. Retrieved April 19, 2002, from http://www.campuscomputing.net/summaries/2000/index.html

Carnevale, D. (2001, November 28). NYUonline will shut down. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2001/11/2001112801u.htm

Carr, S. (1999, December 10). 2 more universities start diploma-granting virtual high schools. *The Chronicle of Higher Education,* p. A49.

Carr, S., & Young, J. R. (1999, October 22). As distance learning boom spreads, colleges help set up virtual high schools. *The Chronicle of Higher Education,* p. A55.

CEO Forum on Education and Technology. (2001). *School technology and readiness—Key building blocks for student achievement in the 21st century: Assessment, alignment, accountability, access, analysis*. Retrieved April 19, 2002, from http://www.ceoforum.org/downloads/report4.pdf

Cisco Systems, Inc. (2001). *Listen, share, and deliver: 2001 annual report*. Retrieved April 19, 2002, from http://www.cisco.com/warp/public/749/ar2001/pdf/AR.pdf

Commission on Technology and Adult Learning. (2001). *A vision of e-learning for America's workforce* (p. 27). Retrieved April 19, 2002, from the National Governors Association Web site: http://www.nga.org/center/divisions/1,1188,C_ISSUE_BRIEF%5ED_2128,00.html

Cuban, L. (2000, February 23). Is spending money on technology worth it? *Education Week on the Web*. Retrieved May 30, 2002, from http://www.edweek.org/ew/ewstory.cfm?slug=24cuban.h19

Cuban, L. (2001). *Oversold and underused: Computers in classrooms*. Cambridge, MA: Harvard University Press.

De Meyer, D., & Steinberg, D. (2001, September 1). The smart business 50: General Electric. *EWeek*. Retrieved April 19, 2002, from http://www.eweek.com/article/0,3658,s=1884&a=10319,00.asp

Department of Education and Cultural Affairs (DECA). (2002). DACS (Dakota Assessment of Content Standards). Retrieved April 19, 2002, from http://www.state.sd.us/deca/dacs/

Educational Testing Service. (2002). *The No Child Left Behind Act: A special report*. Princeton, NJ: Author.

Emery, G. R. (2001, July 16). Army project boosts e-learning prospects. *Washington Technology*, 16(8). Retrieved April 19, 2002, from http://www.washingtontechnology.com/news/16_8/federal/16859-1.html

Evans, P., & Wurster, T. S. (2000). *Blown to bits: How the economics of information transforms strategy*. Boston, MA: Harvard Business School Press.

Federal Reserve Bank of San Francisco. (2000, November 10). Information technology and productivity. *FRBSF Economic Letter* (Number 2000-34). Retrieved April 19, 2002, from http://www.frbsf.org/econrsrch/wklyltr/2000/el2000-34.html

Georgia Department of Education. (n.d.). *Criterion-Referenced Competency Test (CRCT) frequently asked questions*. Retrieved April 19, 2002, from http://www.doe.k12.ga.us/sla/ret/crct_faq.html

Georgia Department of Education. (2001). *Georgia End-of-Course Test information brochure*. Retrieved March 12, 2002, from http://www.doe.k12.ga.us/sla/ret/eoct_tests/eoct-broc.pdf

Henry, D., Buckley, P., Gill, G., Cooke, S., Dumagan, J., Pastore, D., & LaPorte, S. (1999). *The emerging digital economy II*. Washington, DC: U.S. Department of Commerce.

Kerrey, B., & Isakson, J. (2000). *The power of the Internet for learning: Moving from promise to practice: Report of the Web-based Education Commission to the President and the Congress of the United States* (p. 134). Retrieved April 19, 2002, from http://interact.hpcnet.org/webcommission/index.htm

Konrad, R. (2001, June 26). E-learning companies making the grade. *CNET News.com*. Retrieved April 19, 2002, from http://news.cnet.com/news/0-1007-200-6386106.html?tag=owv

Kurzweil, R. (1999). *The age of spiritual machines: When computers exceed human intelligence*. New York: Viking.

Lapp, M. S., Grigg, W., & Tay-Lim, B. S.-H. (2002). *The nation's report card: U.S. history 2001*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.

Mangan, K. S. (2001a, September 13). UNext lays off 135 employees, citing need to spend conservatively. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2001/09/2001091301u.htm

Mangan, K. S. (2001b, September 25). Error message leads many who took GMAT to question their scores. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://www.uh.edu/admin/media/topstories/che092501gmat.htm

Maryland State Department of Education. (2001). *Computer adaptive testing software, support services and technical support* (RFP #R00R2200687). Baltimore, MD: Author.

Maslen, G. (2001, August 8). Australia will spend $100-million on distance programs for developing countries. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2001/08/2001080801u.htm

McKinsey & Company. (2001). *U.S. productivity growth, 1995–2000*. Retrieved April 19, 2002, from http://www.mckinsey.com/knowledge/mgi/feature/index.asp

Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114, 449–458.

Mislevy, R. J., Steinberg, L. S., Almond, R. G., Breyer, F. J., & Johnson, L. (2001). *Making sense of data from complex assessments* (CSE Technical Report 538). Retrieved April 19, 2002, from the UCLA CRESST Web site: http://www.cse.ucla.edu/CRESST/Reports/RML%20TR%20538.pdf

Moe, M. T., & Blodget, H. (2000). *The knowledge web: People power—Fuel for the new economy.* San Francisco: Merrill Lynch.

National Association of State Boards of Education (NASBE). (2001). *Any time, any place, any path, any pace: Taking the lead on e-learning policy.* Retrieved April 19, 2002, from http://www.nasbe.org/Organization_Information/e_learning.pdf

National Center for Education Statistics. (2000). *Teacher use of computers and the Internet in public schools* (NCES 2000-090). Retrieved May 29, 2002, from the U.S. Department of Education Web site: http://nces.ed.gov/pubs2000/2000090.pdf

National Center for Education Statistics. (2001). *Internet access in U.S. public schools and classrooms: 1994–2000* (NCES 2001-071). Retrieved April 19, 2002, from the U.S. Department of Education Web site: http://nces.ed.gov/pubs2001/2001071.pdf

National Commission on Excellence in Education. (1983). *A Nation at risk: The imperative for educational reform.* Retrieved April 19, 2002, from http://www.ed.gov/pubs/NatAtRisk/

Negroponte, N. (1995). *Being digital.* New York: Vintage.

Neuburger, W. (2001, April). On-line assessments: A state's perspective. In R. E. Bennett (Chair), *The "Three W's" meet the "Three R's": Computer-based testing in the schools—Why, what, and how.* Symposium conducted at the annual meeting of the National Council on Measurement in Education, Seattle.

Oliner, S. D., & Sichel, D. E. (2000, March). The resurgence of growth in the late 1990s: Is information technology the story? In *Structural Change and Monetary Policy.* Conference sponsored by the Federal Reserve Bank of San Francisco and Stanford Institute for Economic Policy Research, San Francisco. Retrieved April 19, 2002, from http://www.frbsf.org/economics/conferences/000303/papers/resurgence.pdf

Olson, L. (2002, January 23). Idaho to adopt 'adaptive' online state testing. *Education Week.* Retrieved April 19, 2002, from http://www.edweek.org/ew/newstory.cfm?slug=19online.h21

Oregon Department of Education (ODE). (2002). *Technology Enhanced Student Assessment (TESA) System.* Retrieved April 19, 2002, from http://www.ode.state.or.us/asmt/tesa/

Pellegrino, J. W., Chudowski, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academy Press.

Powers, D. E., & Potenza, M. T. (1996). *Comparability of testing using laptop and desktop computers* (RR-96-15). Princeton, NJ: Educational Testing Service.

President's Information Technology Advisory Committee Panel on Transforming Learning. (2001). *Report to the President: Using information technology to transform the way we learn*. Retrieved April 19, 2002, from the National Coordination Office for Information Technology Research and Development Web site: http://www.itrd.gov/pubs/pitac/pitac-tl-9feb01.pdf

Quality Education Data. (2001). *QED's school market trends: Internet usage in teaching—2001–2002*. Denver, CO: Author.

Russell, M., & Haney, W. (2000). Bridging the gap between testing and technology in schools. *Education Policy Analysis Archives,* 8(19). Retrieved April 19, 2002, from http://epaa.asu.edu/epaa/v8n19.html

Russell, M., & Plati, T. (2001). Effects of computer versus paper administration of a state-mandated writing assessment. *TC Record.Org*. Retrieved April 19, 2002, from http://www.tcrecord.org/Content.asp?ContentID=10709

Schaeffer, G. A., Bridgeman, B., Golub-Smith, M. L., Lewis, C., Potenza, M. T., & Steffen, M. (1998). *Comparability of paper-and-pencil and computer adaptive test scores on the GRE General Test* (RR-98-38). Princeton, NJ: Educational Testing Service.

Schaeffer, G. A., Steffen, M., Golub-Smith, M. L., Mills, C. N., & Durso, R. (1995). *The introduction and comparability of the computer-adaptive GRE General Test* (RR-95-20). Princeton, NJ: Educational Testing Service.

Schwartz, K. D. (2001). Learning is mandatory; presence is optional. *Mobile Computing Online*. Retrieved April 19, 2002, from http://www.mobilecomputing.com/showarchives.cgi?145

Shea, C. (2001, September 16). Taking classes to the masses. *The Washington Post Online*. Retrieved April 19, 2002, from http://www.washingtonpost.com/wp-dyn/articles/A19591-2001Sep12.html

State Board of Education. (2001). *Assessment brief: Preliminary information about computerized adaptive testing system*. Retrieved April 19, 2002, from http://www.ncpublicschools.org/accountability/testing/briefs/computerizedadaptive/

State of Utah. (2001). *Request for proposal (LW1039): Development of the Utah Core Curriculum Elementary and Secondary End-of-Level and End-of-Course Core Assessment Series*. Salt Lake City, UT: Author.

State University's premiere education program reports enrollment at all time high. (2001, May 18). [Press release]. Retrieved May 1, 2002, from http://test.sln.suny.edu/sln/public/original.nsf/58f6ddfae8daaa10852567b00054e290/22cb3669f3d4a9d485256b76005ae997?OpenDocument

Steinberg, D. (2001, September 1). The smart business 50: Wal-Mart. *eWEEK*. Retrieved April 19, 2002, from http://www.eweek.com/article/0,3658,s=1884&a=10427,00.asp

Taylor, C., Jamieson, J., Eignor, D., & Kirsch, I. (1998). *The relationship between computer familiarity and performance on computer-based TOEFL test tasks* (TOEFL-RR-61). Princeton, NJ: Educational Testing Service.

Technology counts 2001: The new divides. (2001). *Education Week on the Web*. Retrieved April 19, 2002, from http://www.edweek.org/sreports/tc01/charts/tc01chart.cfm?slug=35challenges-c10.h20

U.S. Department of Commerce. (2002). *A nation online: How Americans are expanding their use of the Internet*. Washington, DC: Author. Retrieved April 19, 2002, from http://www.ntia.doc.gov/ntiahome/dn/nationonline_020502.htm

U.S. Department of Education. (2002). *FY 2003 budget summary—February 4, 2002*. Retrieved April 19, 2002, from http://www.ed.gov/offices/OUS/Budget03/Summary/SectionII/A.html#top

Virginia Department of Education. (n.d.). *eSOL Web-based assessment*. Retrieved April 19, 2002, from http://www.pen.k12.va.us/VDOE/Technology/soltech/soltech.html

Virginia Department of Education. (2001). V*irginia Algebra Readiness Diagnostic Test (ARDT) participation resources*. Retrieved April 19, 2002, from http://www.accessardt.com/

Wainer, H., & Eignor, D. (2000). Caveats, pitfalls, and unexpected consequences of implementing large-scale computerized testing. In H. Wainer (Ed.), *Computerized adaptive testing: A primer* (2nd ed.). Mahwah, NJ: Erlbaum.

Walfish, D. (2001, May 21). Chinese government predicts strong growth in online education. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2001/05/2001052101u.htm

Wise, S. (1997). *An evaluation of the item pools used for computerized adaptive test versions of the Maryland Functional Tests*. Retrieved April 19, 2002, from the Maryland Assessment Research Center for Education Success Web site: http://www.marces.org/mdarch/pdf/M032045.pdf

World Bank Group. (2001, August 3). Virtual Colombo Plan. *DevNews*. Retrieved April 19, 2002, from http://www.worldbank.org/developmentnews/stories/html/080301a.htm

Young, J. (2001, June 13). Web site tracks statewide virtual university projects. *The Chronicle of Higher Education*. Retrieved April 19, 2002, from http://chronicle.com/free/2001/06/2001061301u.htm

About the Author

Randy Elliot Bennett
Educational Testing Service
Princeton, NJ 08541
Email: rbennett@ets.org

Randy Elliot Bennett is Distinguished Presidential Appointee at Educational Testing Service in Princeton, NJ, a nonprofit organization dedicated to research and service in educational measurement. Dr. Bennett began his employment at ETS in 1979. Since the 1980's, he has conducted research on the applications of technology to testing, on new forms of assessment, and on the assessment of students with disabilities. Dr. Bennett's work on the use of new technology to improve assessment has included research on presenting and scoring open-ended test items on the computer, on multimedia and simulation in testing, and on generating test items automatically. Dr. Bennett is the editor or author of seven books and many other publications including, "Reinventing Assessment: Speculations on the Future of Large-Scale Educational Testing" (ftp://ftp.ets.org/pub/res/reinvent.pdf) and "How the Internet will Help Large-Scale Assessment Reinvent Itself" (http://epaa.asu.edu/epaa/v9n5.html). He has made presentations on this and related topics throughout the world. Dr. Bennett is currently co-directing the Technology Based Assessment Project, a series of studies designed to lay the groundwork for introducing computerized testing to the U.S. National Assessment of Educational Progress. This project is believed to be the first in the world to have tested nationally representative samples of school children on computer.

J·T·L·A

The Journal of Technology, Learning, and Assessment