



NBETPP

s t a t e m e n t s

Volume 2

Number 1

January 2001

Educational Testing as a Technology

George F. Madaus

National Board on Educational Testing and Public Policy

Carolyn A. and Peter S. Lynch School of Education

Boston College

Educational testing is a familiar, enduring part of our culture. But is it a technology? The term technology often conjures up visions of scientific experiments and industrial processes. But technology is also a tool, something put together to satisfy a need, solve a problem, or attain a goal in social, economic, and educational institutions.¹ More broadly still, technology is any body of special knowledge, skills, and procedures that people use, or even all “artificial aids to human activity.”² In these definitions, technology and technological devices are seen as tools, as means to some end. Thus testing is clearly technology, one embedded in such systems as education, government, and business.

Testing has its “hardware” – test booklets, answer sheets, and optical scoring machines that make the testing of large numbers efficient and economical – and like much of our technology, it has a community of practitioners, trained for membership in that community, and with a common language and set of practices.³ Further, testing is used as a means to important societal ends. For example, in education over the past two thousand years, tests have been used to help eliminate patronage, open or restrict access to various opportunities, establish and maintain standards of performance, hold teachers, students, or schools accountable for learning, and allocate scarce resources.⁴

The National Board on Educational Testing and Public Policy, located in the Lynch School of Education at Boston College, is an independent body created to monitor testing in American education. The National Board provides research-based information for policy decision making, with special attention to groups historically underserved by the educational system. In particular, the National Board

- Monitors testing programs, policies, and products
- Evaluates the benefits and costs of specific testing policies
- Evaluates the extent to which professional standards for test development and use are met in specific contexts

This statement develops the idea that testing is a technology and discusses how this affects the work of the National Board.

George Madaus is a Senior Fellow with the National Board on Educational Testing and Public Policy and the Boisi Professor of Education and Public Policy in the Lynch School of Education at Boston College.

Statements Series Editor:
Marguerite Clarke

Means Versus Ends

People generally do not concern themselves with how a technology like testing – or for that matter a light bulb, telephone, car, and the like – is devised, or how it works. They leave that to those who invent, develop, build, or repair the item in question. They are interested in the use of a technology. Once a test exists, users interact with it straightforwardly for specific purposes. Knowing what a test is, in a technical sense, is not viewed as terribly important. But this division between the device and its results can present a problem. Albert Borgmann uses a stereo system as an example:

Surely a stereo set... is a technological device. Its reason for being is well understood. It is to provide music. But this simple understanding conceals the characteristic way in which music is produced by a device.... To an apparent richness and variety of technologically produced music there corresponds an extreme concealment or abstractness in the mode of its production.... It is the division between the commodity, e.g., music, and the machinery, e.g., the mechanical and electronic apparatus of a stereo set, that is the distinctive feature of a technological device.⁵

So too the test. It produces results, and there is concealment or abstractness in its production. In addition, the device of testing allows the results to be used in completely unintended ways.

Borgmann also points out that while people claim technology is a value neutral tool, this position neglects the impact of the technology on the intended ends. "Putting technology in the context of political purposes is itself naive if one fails to consider ... the radical transformation of all policies that technology may bring about."⁶

If educational testing is used for the political purposes of accountability, allocating resources, or driving the system, we need to know how these policies impinge on the educational system as a whole. Are tests a means to an end, or have they become an end in themselves? Currently it appears that the latter is the case. There is a circularity in how tests are used and viewed. Often test results define a problem (e.g., low academic standards); new tests are then mandated to drive policy to solve the problem; coming full circle, the new test results then are used to show whether the problem is solved or has grown worse, or are discounted as corrupted or meaningless.

Testing Technology and Power

Technology affects our lives and our society in profound ways. Certainly we would argue that this is true of testing. If technology – and, we submit, testing – is a “mere instrument... the inquiry of what guides technology becomes a task in its own right.”⁷ Technologies are not merely an aid to human activity, but are “powerful forces acting to reshape that activity and its meaning” and doing so in uncontrolled ways, with most of us sitting on the sidelines allowing this reshaping to occur without our participation or even awareness of what is happening.⁸

This “technological somnambulism” has led to the prevalent, but mistaken, attitude that technology in general, and testing in particular, is “fundamentally neutral [in its] moral standing.”⁹ Testing may be used well or poorly, used for good or bad purposes; it is the particular use, in a particular context, that determines its moral standing. This emphasis on use rather than on the technology itself is captured perfectly in the dictum, “Guns don’t kill people, people kill people.”

Like all technologies, tests can be “judged not only for their contributions to efficiency and productivity and their ... side effects, but also for the ways in which they can embody specific forms of power and authority.”¹⁰ Langdon Winner describes the moral dimension of a technology apart from a particular use:

Indeed, many... technologies that have political consequences... transcend the simple categories of “intended” and “unintended” altogether.... The very process of technical development is so thoroughly biased in a particular direction that it regularly produces results heralded as wonderful breakthroughs by some social interests and crushing setbacks by others. In such cases it is neither correct nor insightful to say “Someone intended to do somebody else harm.” Rather one must say that the technological deck has been stacked in advance to favor certain social interests and that some people were bound to receive a better hand than others.¹¹

The late French philosopher-historian Michel Foucault points to the obvious connection between testing and the exercise of power. His 1979 book *Discipline and Punish: The Birth of the Prison* has ten intriguing pages on school examinations as a means of control, offering a quite different look at the role that testing plays in regulating schools and the people in them.



Testing may be used well or poorly, used for good or bad purposes; it is the particular use, in a particular context, that determines its moral standing.





Thus, the entire test development process – what we measure, the training of test developers, the material included, format, language, directions, the way cut scores are set, the validation process, and so on – might unintentionally stack the testing deck in favor of certain groups.

Foucault points out that the development of the written exam or test made it possible for the first time to accumulate student marks, organize them, rank them, classify them, form categories, determine averages, and fix norms. It was the beginning of a “comparative system that made possible the measurement of overall phenomena, the description of groups, the characterization of collective facts, the calculation of the gaps between individuals, their distribution in a given ‘population.’”¹²

The ability to form and describe groups in terms of their test performance can be used as a means of holding programs, schools, or school systems accountable. The information provided by a test gives those in charge of it the ability to not only objectify individuals, but also to form, describe, and objectify groups. It is not the tests per se that control the actions of teachers, students, and administrators. It is the coupling of the test results with important rewards or sanctions that gives those who control testing power over the action of others. This in turn makes a bureaucratic mechanism of program, or school-level, accountability possible.

Thus, the entire test development process – what we measure, the training of test developers, the material included, format, language, directions, the way cut scores are set, the validation process, and so on – might unintentionally stack the testing deck in favor of certain groups. Accordingly, the National Board on Educational Testing and Public Policy will foster consideration of testing as a technology, examine the assumptions and moral dimensions of test development, and study how a test is used and what the consequences are.

In addition to the proposition that testing is a technology, two additional characteristics of technologies must be considered by the National Board. The first is that technological endeavors tend to be directed by elites isolated from those who are not members of the elites, to the detriment of both groups. The second is that technological endeavor is firmly associated with a “religion of progress.”¹³ These two characteristics are discussed in the next two sections.

Elitism and Isolation

One result of the scientific and technological revolution of the past 150 years is that practitioners of technologies have formed their own communities with their own vocabularies and value systems. In the early days of our republic, scientists and technicians could speak almost as easily to lay people as to each other. With the rise of standardization in the 19th century, highly specialized vocabularies for the phenomena being studied began to evolve along with highly specialized groups associated with various technologies and social techniques.¹⁴

The formation of such elites and professional communities is a two-edged sword. It can enhance the members' accomplishments and facilitate communication among them, but it can also isolate them from and even alienate non-members. This may be due in part to the tendency of these groups to resist attempts by non-members to influence their activities. It is

a short step indeed from the claim that true science and technology must be allowed to operate free from irrational or self-serving outside influences to the claim that scientific and technological practitioners must be allowed to operate free from any challenge or critique from those outside their domains of expertise.... laissez faire's iron law – "never interrupt the working of the method by outside critique" – resides in Western consciousness at the primordial level of symbol and rhetoric.¹⁵

Moreover, "the very organization of a discipline ... often tends to cut its practitioners off from other disciplines."¹⁶

It is certainly true that a gulf has developed between the testing community and others. In testing, as in other technological areas, "there is almost no middle ground of rational discourse, no available common language with which persons of differing backgrounds can discuss matters of technology in thoughtful, critical terms... conversations gravitate toward warring polarities and choosing sides."¹⁷



It is certainly true that a gulf has developed between the testing community and others. In testing, as in other technological areas, "there is almost no middle ground of rational discourse, no available common language with which persons of differing backgrounds can discuss matters of technology in thoughtful, critical terms...





Too often, however, the technical elite thinks in terms of what is useful and efficient rather than what is good or just.

Criticism that requires the testing elite or the critics of testing to reorganize their view of the world provokes hostility and defensiveness.¹⁸ For example, an article entitled *The War on Testing: Detroit Edison in Perspective* argues that “The attack on tests is... an attack on truth itself by those who deal with unpleasant and unflattering truths by denying them and by attacking and trying to destroy the evidence for them.”¹⁹ Or consider a compilation of attacks on testing that include:

carelessness, hatred, favoritism, labor unrest, unprogressiveness, defective art, dishonesty, discontent, poverty, fraudulence, laziness, a generator of mental defectiveness and physical degeneration, serfdom, radicalism, suffering, death, strikes, and war.²⁰

Too often, however, the technical elite thinks in terms of what is useful and efficient rather than what is good or just.²¹ Thus, criticism of a sociotechnical practice often meets with defensiveness:

A typical response of engineers... is... “tell us the problem... We will find a solution. That’s our job. But you may not presume to question the nature of our solution. You are not a member of a technical profession and, therefore, know nothing of relevance. If you insist on raising questions about the appropriateness of means we devise, we can only conclude that you are antitechnology.”²²

When the testing community defends testing against strong attacks, they understandably fall back on their specialized vocabulary, their values, and their techniques – techniques that themselves embody values. They often overlook the fact that a highly technical psychometric defense of a test is itself very limited, and is based on a narrow set of technical values.

It is important to find methods of discourse between the testing community and its critics that will reveal the complexity of testing issues – technical and social – and the underlying values associated with various positions. As Robert Bellah and his associates aptly point out:

Complexity is real enough, but it should not be a cover beneath which undemocratic managers and experts can hide. Our culture or our institutions may lead us to believe that the big issues are beyond us; but then we need to change those assumptions, and a social science that takes its public responsibility seriously can help us do so.²³

To improve communication between the groups, words in the common lexicon that have been appropriated and narrowed by the testing community need to be abandoned so that we no longer talk on parallel, non-intersecting tracks. For example, test bias within the community's lexicon emphasizes the technical aspect – differential predictive validity – while the more common meaning of bias focuses on the concept of fairness.

Second, the social, technical, and value issues associated with testing should be examined by more people outside the testing community – not only scientists or technicians from other communities but non-technical communities as well. The National Board might study the social-technical issues that revolve around testing and how we might resolve conflicting values.

Winner points out that many fictional utopias propose government by a technical elite, and a number of works, such as *Principles of Scientific Management*, make similar arguments.²⁴ Winner notes that one group is excluded from membership in such an elite: the great mass of those who are considered to lack the knowledge or credentials to govern a technological society. The National Board must take pains to assure that the laity is not excluded from any of its undertakings.

The Religion of Progress

Coupled with the close relationship of technology to science and economic growth, our nearly boundless technological capabilities have contributed to an outlook that values efficiency and the rational and quantitative over the subjective and qualitative. But efficiency is a “systematically incomplete concept. For efficiency to come into play, we need antecedently fixed goals on behalf of which values are minimized or maximized. Those goals remain in the dark.”²⁵

Daniel Boorstin has pointed out that the problems stemming from technological development differ markedly from those traditionally posed in the political realm: “The problem of politics, then, is essentially the problem of man coming to terms with his problems. But our problem in the United States – and, generally speaking, the central problem of technology – is how to come to terms with solutions.”²⁶



It is important to find methods of discourse between the testing community and its critics that will reveal the complexity of testing issues – technical and social – and the underlying values associated with various positions.





What we value most about schools seems increasingly to be their pupils' performance on standardized tests. Valuing test scores so highly leads to policy decisions that are bereft of bases in theory or research, or worse, are contrary to research findings.

Not only are there technological solutions for problems that do not yet exist, but we accept the solutions as good because they represent progress. Further, we allow the solutions to dictate what the problems should be. Winner refers to our “fast belief in the religion of progress,” which led us to choose not to tailor technology to human needs, but instead to adopt the practice of “renovating human needs to match what modern science and engineering happened to make available.”²⁷

In the field of educational and psychological testing, we have developed the technology of standardized, machine scorable, objective tests to a high degree, and the availability of this “solution” has influenced both the nature of the problems we examine and the way we conceptualize general problems facing American society. For example, of all the problems that commentators tell us exist in public schools – drugs, violence, truancy, dropouts, teacher burn-out, racial and economic segregation, and so on – the one that gets the most attention seems to be test scores. Why? Simply because we have the technology to test students and thus reduce their performance and abilities to numbers that lend themselves to statistical analyses, graphs, and comparisons. This has had an undeniable impact on our value system. What we value most about schools seems increasingly to be their pupils' performance on standardized tests. Valuing test scores so highly leads to policy decisions that are bereft of bases in theory or research, or worse, are contrary to research findings. For example, a groundswell of support is building for national achievement tests – based on a touted “new,” but in reality quite old, assessment technology – that supposedly will cure the problems of public schools by inspiring pupils to perform better. A related movement is afoot to reverse a quarter century of employment law and allow employers to set test performance requirements that have no demonstrable relationship to job requirements – this, too, to inspire America's youth to stay in school and do well.

Another aspect of the faith in quantitative, technological methods is an unnecessarily narrow view of knowledge and the methods by which it is obtained. Standardized tests are part of the scientific method used in the social sciences: knowledge is gained through a linear progression of procedures aimed at discerning cause and effect. Ernest House sees a great danger in exclusive adherence to this method:

[It] implies that experts using the proper methods can ascertain the best programs and approaches for addressing social ills. This attitude delegitimizes knowledge derived from other sources and people, leading to a form of scientism: Only information derived from certain techniques is true knowledge.²⁸

The effect may well be to distort reality. The tools of behavioral science are not the neutral measurement devices people often assume them to be. Educational and psychological tests are not like yardsticks or scales for measuring pupils' heights and weights. They are more complex, less direct; and most important, they require that the objects of the measurement accommodate themselves to the tools. Again citing Winner:

Far from being neutral, uninvolved sensing devices, these technical ensembles have their own requirements.... Individuals and social institutions must adapt to these requirements or they cannot be adequately evaluated. The influences of standardized, centralized, computer-scored mass testing in education is a good example of how this works. The tests can measure only those qualities of a student's education that can be represented in pencil-marked squares on the test sheet during a four-hour examination. On top of this, students and teachers soon learn the game and its stakes. It is not unusual for high-school seniors in the United States to spend the better part of their time mastering the specific kinds of performance likely to appear on College Entrance Examination Board tests.... Thus, as a result of the structure of the instrument and human adaptation to it, techniques of measurement become purely self-fulfilling.²⁹

The National Board must grapple with the issue of values and with the unquestioning faith in the technology of testing's ability to solve our educational problems.



Educational and psychological tests are not like yardsticks or scales for measuring pupils' heights and weights. They are more complex, less direct; and most important, they require that the objects of the measurement accommodate themselves to the tools.





Testing is a complex technological system with its own infrastructure, akin to transportation, power, or manufacturing systems.

Conclusion

The National Board believes that tests, like computers or airplanes, are a technology with a well-developed technological community and technical underpinnings arcane to lay people. Testing is a complex technological system with its own infrastructure, akin to transportation, power, or manufacturing systems. To argue that whether to have state or nationally mandated testing programs is a political, and not a technical, question is to lose sight of the fact that policy decisions have technical implications, and technical decisions have policy ramifications. Moreover, while both policy decisions and the application of a technology can solve problems, they also create them. According to one observer of American society:

Americans need to fathom the depths of the technological society, to identify currents running more deeply than those conventionally associated with politics and economics. Indeed many of the forces that Americans need to understand and control in order to shape their destiny... are now not primarily natural or political but technological.... Americans [need to realize] that not only their remarkable achievements but many of their deep and persistent problems arise... from the mechanization and systematization of life and from the sacrifice of the organic and spontaneous.³⁰

notes

- 1 See: Basalla, G. (1988), *The evolution of technology*, New York: Cambridge University Press; Ellul, J. (1964), *The technological society*, New York: Vintage Books; Ellul, J. (1990), *The technological bluff*, Grand Rapids, MI: Williams B. Eerdmans Company; Staudenmaier, J. M., S.J. (1988), *Technology and faith* [Audio tapes Kansas City, MO: National Catholic Reporter].
- 2 See: Lowrance, W. W. (1986), *Modern science and human values*, New York: Oxford University Press; Winner, L. (1977), *Autonomous technology: technic-out-of-control as a theme in political thought*, Cambridge MA: MIT Press; Quotation from Winner, L. (1986), *The whale and the reactor: A search for limits in an age of high technology*, p. 4., Chicago: The University of Chicago Press.
- 3 Staudenmaier, J. M., S.J. (1985). *Technology's storytellers: Reweaving the human fabric*. Cambridge MA: MIT Press.
- 4 Madaus, G. F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 119-154). New York: Macmillan.
- 5 Borgmann, A. (1984). *Technology and the character of contemporary life: A philosophical inquiry*, pp. 3-4. Chicago: University of Chicago Press.
- 6 Ibid., p. 11.
- 7 Ibid., p. 11; emphasis in original.
- 8 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 6. Chicago: The University of Chicago Press.
- 9 Ibid., p. 6.
- 10 Ibid., p. 19.
- 11 Ibid., pp. 25–26.
- 12 Foucault, M. (1979). *Discipline and punish: The birth of the prison*, p.190. Harmondsworth, England: Penguin.
- 13 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 170. Chicago: The University of Chicago Press.
- 14 Gleick, J. (1987). *Chaos: Making a new science*. New York: Viking Penguin Inc.
- 15 Staudenmaier, J. M., S.J. (1990). *Science, technology, and empowerment: Who gets a say?*, p.3. Paper presented at the Technische Universiteit Eindhoven Technological Development and Science in the 19th and 20th Centuries meeting, the Netherlands.
- 16 Boulding, K., E. (1970). *Beyond economic: essays on society, religion, and ethics*, p.146. Ann Arbor, MI: The University of Michigan Press.
- 17 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 11. Chicago: The University of Chicago Press.
- 18 Gleick, J. (1987). *Chaos: Making a new science*. New York: Viking Penguin Inc.
- 19 Lerner, B. (1979). *The war on testing: Detroit Edison in perspective*, p.1. Princeton, NJ: Educational Testing Service.
- 20 O'Meara, J. F. (1944). *A critical study of external examinations and of their influence on secondary education*, p.10. Unpublished master's dissertation, University College Cork, Ireland.
- 21 Merton, R. K. (1964). Foreword. In E. Jacques, (Ed.), *The technological society* (pp. v-viii). New York: Vintage Books.
- 22 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 11. Chicago: The University of Chicago Press.
- 23 Bellah, R., N., Madsen, R., Sullivan, W. M., Swidler, A., & Tipton, S. M. (1991). *The good society*, p.20. New York: Alfred A. Knopf.
- 24 See: Winner, L. (1977), *Autonomous technology: Technic-out-of-control as a theme in political thought*, Cambridge MA: MIT Press; Taylor, F., M. (1914), *The principles of scientific management*, New York/London: Kegan Paul, Trench.
- 25 Borgmann, A. (1984). *Technology and the character of contemporary life: A philosophical inquiry*, p.9. Chicago: University of Chicago Press.
- 26 Boorstin, D. J. (1978). *The republic of technology*, p.9. New York: Harper & Row.
- 27 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 170. Chicago: The University of Chicago Press.
- 28 House, E. R. (1990). Methodology and justice, p.35. *Evaluation and Social Justice: Issues in Public Education. New Directions for Program Evaluation*, 45, 23-36.
- 29 Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*, p. 234-235. Chicago: The University of Chicago Press.
- 30 Hughes, T. P. (1989). *American genesis: A century of invention and technological enthusiasm*, p.4. New York: Penguin Books.



About the National Board on Educational Testing and Public Policy

Created as an independent monitoring system for assessment in America, the National Board on Educational Testing and Public Policy is located in the Carolyn A. and Peter S. Lynch School of Education at Boston College. The National Board provides research-based test information for policy decision making, with special attention to groups historically underserved by the educational systems of our country. Specifically, the National Board

- Monitors testing programs, policies, and products
- Evaluates the benefits and costs of testing programs in operation
- Assesses the extent to which professional standards for test development and use are met in practice

This National Board publication series is supported by a grant from the Ford Foundation.

The National Board on Educational Testing and Public Policy

Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Telephone: (617)552-4521 • Fax: (617)552-8419
Email: nbetpp@bc.edu

Visit our website at nbetpp.bc.edu for more articles, the latest educational news, and information on NBETPP.



The Board of Overseers

Antonia Hernandez

President and General Council
Mexican American Legal Defense
and Educational Fund

Harold Howe II

Former U.S. Commissioner of
Education

Paul LeMahieu

Superintendent of Education
State of Hawaii

Peter Lynch

Vice Chairman
Fidelity Management and
Research

Faith Smith

President
Native American Educational
Services

Gail Snowden

Managing Director, Community
Banking
FleetBoston Financial

Peter Stanley

President
Pomona College

Donald Stewart

President and CEO
The Chicago Community Trust



BOSTON COLLEGE