

Quantification and Intelligence Testing: A Reassessment,

Richard Bishirjian, Yorktown University
Peter Wood, National Association of Scholars
Presented to the Virginia Political Science Association
Old Dominion University
Saturday, November 20, 2009

I'm not a student of the history of testing for intelligence, but I have always thought that intellectual ability, the ability to think conceptually, is primarily shaped by culture and that testing for intelligence reflects the culture of the West, not human intelligence qua human.

"I got rhythm" is a cultural statement, not a description of a genetic trait.

This thought returned to me when I saw that Charles Murray has been given prominence at several conferences I attended recently and at the American Enterprise Institute. Since Charles Murray's and Richard Herrnstein's *The Bell Curve* (1994), Murray has become the most listened to, if not the reigning authority, on intelligence and education.

This paper will give a critique of the strong hereditarian bias of Charles Murray whose arguments defy just about every philosophical and theological truth of the Western philosophical tradition and a warning about the dangers to academic freedom by the push for measurement of learning outcomes that now dominates the accreditation of higher education degree programs and institutions.

With respect to the accreditation of higher education the U.S Department of Education under former Secretary Margaret Spellings engaged in a systematic effort to dumb down college education by the forced imposition of a progressive education ideology.

The debate about higher education is framed by Charles Murray who advocates a college education for only those in the top 20th percentile of IQ and the U.S. Department of Education that wants to measure the learning outcomes of every course taught in every college in America.

Murray's focus on genetics even embarrassed his neoconservative colleague, Nathan Glazer, who writes: "I wish Herrnstein and Murray had pressed further other explanations for these differences among groups before taking up differences in biological inheritance. Indeed, I wish they had dropped resort to such explanations totally: little would have changed in their argument if they had. For the nonbiological explanations will carry us far, to the point perhaps we need make no reference to genes at all." (Glazer, *Scientific Truth and the American Dilemma*, p. 143 in Steven Fraser, ed., *The Bell Curve Wars*, Basic Books, 1995)

There are other reasons to demote Murray's celebrity.

Let me begin by reading some statements of Richard Nisbett, a cultural psychologist at the University of Michigan, that conflict with the strong hereditarian approach of Charles Murray. Two of his works are relevant:

The Geography of Thought: How Asians and Westerners Think Differently (2003)

Intelligence and How to Get It: Why Schools and Culture Count (2009)

I take the following statements from *The Geography of Thought*:

Economic and social factors can affect perceptual habits which accounts for the differences between agricultural peoples and those living in industrial societies. (42)

Asians' feeling good about themselves is tied "to the sense that they are in harmony with the wishes of the groups to which they belong." (49)

Contrary to Westerners, for Asians choice is not a high priority. (49)

Asians prefer jobs in which everyone works together, and no one person is singled out for personal honor. (63)

In the West, the purpose of negotiation is to achieve a desired result. In Asia negotiation avoids either/or choices. (75)

Differences between the sexes are greater in the West. (99)

Chinese regard change more likely than Americans. (104)

Westerners think in linear terms. (108)

Asians organize their perceptions in terms of perceived relationships. (140)

Children learn nouns more rapidly than they learn verbs, but East Asian children learn verbs at about the same rate they learn nouns. (149)

East Asians live in an interdependent world in which the self is part of a larger whole. (76)

And finally, and I quote: "For Westerners it is the self who does the acting; for Easterners, action is something that undertaken in concert with others or that is the consequence of the self operating in a field of forces." (158)

First of all, let us become aware that the strong hereditarian approach of Charles Murray gives little consideration to the historical and cultural origins of cognitive ability.

For many of us trained in the Western philosophical tradition intelligence is best understood as an historical discovery, a discovery discussed by Bruno Snell in his seminal essay, "Homer's View of Man" a chapter in *The Discovery of the Mind: The Greek Origins of European Thought* (1960) and particularly Werner Jaeger's *The Theology of the Early Greek Philosophers* (1964) where Jaeger captures the mystic aspects of early natural science.

Those studies along with the interpretation of cosmological consciousness in Eric Voegelin's *Israel and Revelation*, *World of the Polis* and *New Science of Politics* capture the essence of Voegelin called the "break with myth" and the discovery of man's unity of soul and mind (intelligence).

Eric Voegelin fashioned a philosophical anthropology that defined Western society as "macroanthropos" and counterposed them to societies and cultures preceding the advent of philosophy in Hellas and revelation in Israel which he described as "cosmological."

Ancient man, living within a world of experience that was "cosmological" did not see the world around him as a universe of differentiated objects, but as a cosmos composed of other beings whose living presence was manifest in the progress of daily life. That consciousness still resides in Asia, Africa, amongst aborigines and American Indians and explains, I believe, why science progressed in the West to the degree it has.

Nor did ancient, pre-philosophic, man understand himself as a 'person'—a later concept reflecting differentiated consciousness—consisting of unified body and mind.

Everywhere ancient man turned, he encountered the gods and interpreted his own actions by reference to their decisions.

Political order was also understood as standing in direct relationship with the gods. It was not an order which was autonomous or independent of an order higher than itself. Rather it was perceived as an extension of cosmic order. Political community was experienced as a smaller portion of a larger sacred order or cosmos.

This worldview can be traced back to the early evidence of uniquely human existence in the Neolithic age, a cosmological consciousness that was discarded by the Greek mystic philosophers who searched for a first principle that expressed the sacred *arche* in terms that were, by varying degrees, non-mythic.

[Still very close to mythic descriptions, Thales suggested that the origin of the process of coming into being, growth, and death was water, a symbol of generation in all mythic cultures. Anaximenes, perhaps more revolutionary, said that it was air; and Anaximander made the complete break with the formulation that the *arche* was infinite (*to apeiron*) and that the infinite *arche* of being was divine (*to theion*), which symbol, itself, was a philosophic revolution. No longer from that point could the question of the beginnings be answered in terms of a mythic god. Anaximander had abstracted the essence of the genderless divine (*theion*) reality from the mythic gods and chose the neuter article (*to*) to express that absence of myth.

The *arche* of nature is not a god (*theos*), he said, it is the divine (*to theion*) reality.

Socrates stood in this constructive tradition of criticism of myth, founded upon a new concept of divine reality differentiated intellectually from the previous mythical forms.]

The intellect manifest in our understanding of human intelligence was shaped by an intellectual shift (Voegelin calls it a "leap") from mythic consciousness to philosophic in Hellas and through revelation in ancient Israel and the Gospels.

Culture shaped by these Western influences of what Voegelin calls *pneumatic* and *noetic* experiences sustain the differentiated consciousness of post-mythic order and contributed to the development of the skills that are measured today by IQ testing.

Let's look more closely at the assumption that intelligence be measured.

If intelligence can be measured, we must first agree on what it is we are measuring.

In the language of Eric Voegelin, intellect is differentiated consciousness of the nature and limits of mind or *nous* experienced as in-between the divine and the world.

So an examination of the scientific methods of testing of intelligence necessarily takes us back to the Presocratics, Socrates, Plato and Aristotle.

[Bruno Snell (*The Discovery of the Mind*) traces the discovery of the human person and parallel changes that occurred in Greek representations of the human form.

That change in ancient Greece from 10th century depictions of men composed of stick-like members called "geometric art" to 2nd century depictions of the human form as a unified body is the result, Snell shows, of Heraclitus' discovery of the soul (*psyche*) that transformed the previously dominant Homeric vocabulary by

which ancient Greeks understood themselves as men and in turn altered the Greek depiction of man in the plastic arts. (Bruno Snell, *The Discovery of the Mind: The Greek Origins of European Thought*, T.G. Rosenmeyer, trans. New York: Harper and Row, Publishers, Torchbooks, The Academy Library, 1960), 1-22)]

And it raises this question: If intelligence is an artifact that was discovered by the Greek philosophers, it seems a bit strange to say that it is inherited by genes?

If we think of intellect as a learned capacity that was discovered in a process in which the *physiologoi* (natural philosophers) challenged the myths of ancient Greece, we see that nurture, not nature, plays the dominant historical role in shaping the “intelligence” of cultures.

I like to refer to the West Indies and India which were subjected to British colonial rule for hundreds of years as evidence that civilizations shaped by mythic consciousness can traverse the divide between myth and differentiated consciousness through cultural imperialism of the West.

West Indians and Indians from India are comfortable in the West and have progressed in material and scientific ways that nations in grip to tribal myths have not.

Richard Nisbett demonstrates that citizens of Hong Kong can change how they think by shifting from Western to Asian modes.

In Western philosophic terms, the transition from tribe to community is one of movement to consciousness of the common rational part of the soul (*logos*) that holds community together, not physical or material growth in size.

For that reason education (*paideia*) is of supreme importance for the survival of community.

In *Real Education*, however, Charles Murray asserts his belief that a real college education is attainable only by persons in the upper 20th percentile. Low ability children (10th to 37th percentile) cannot be educated.

That too conflicts with Richard Nisbett’s findings that IQ is not immutable, that IQ can be improved, that IQ is higher amongst persons raised in upper middle class families and lower among lower socioeconomic classes, that lack of schooling reduces IQ, that schooling improves memory and that studies of IQ of persons raised in adoptive families reflects also an environment characteristic of adoptive parents, not merely inherited IQ.

The Greek discovery of our common reason that all Greeks shared (but not foreigners) was differentiated in the West by consciousness of the universality of

mankind. “All men are created equal” was a phrase that immediately came to mind in 1776, not “all Americans are created equal.”

The history of intelligence testing, therefore, counters centuries of thinking about education as a cultural artifact (*paideia*) and achieves full impact after World War II when Europe lay prostrate and American behaviorists played a role in the reconstruction of European higher education.

The danger for society at large, and Western culture at large, is that if we participate in a science that denies non-objective truth then everything non-objective is mere opinion.

That, of course, is an upside down world in which the lovers of truth are supplanted by the lovers of opinion.

I conclude, therefore, that the science of testing for intelligence misinterprets the meaning of science *and* intelligence and the assertion that intelligence testing is “scientific” is simply ignorant of philosophy, anthropology, culture and, of course, philosophical anthropology.

Not everything can be measured quantitatively.

Character education, instilling virtue, love for one’s fellow man, and the nuanced truths of classical philosophy are not measurable in terms of behaviors but they are important aspects of what the Greeks called *paideia*.

Aristotle asks in the Nicomachean Ethics, “What is the measure of what is right?” He replies, “the good man” (*spoudaios*).

He touched upon an important truth.

In life we are guided by good men—and women—whose judgment we trust and whose qualities of character compel us to use them as a standard and measure of what is right, wrong, true and untrue, honorable and dishonorable, just and unjust. We ourselves may not know what to do, but we can know what the good man would do.

In American higher education today, entire universities are being compelled to measure the learning outcomes acquired by their students.

In engaging in this research and reporting, the compilers of these measures are instructed to state learning as “behavior.”

Students don’t learn, understand or appreciate, they define, compare, contrast and analyze.

These emasculated learning outcomes devalue what education is all about and revive what Eric Voegelin described as the “derailment” of philosophy by propositional metaphysics.

In my own courses in the history of political theory and modern ideologies that I teach online at Yorktown University I have described the learning outcomes of every session of my course.

What you’ll find listed there are conclusions described in behavioral terms of some of the greatest intellectual events that occurred in the West—the break from mythic consciousness, the turning around of the soul depicted in Plato’s Myth of the Cave, the discovery of the mind, the discovery of the human person in relation to transcendent divine reality, the measure of what is right and just, and the growth of the administrative state.

In stating these discoveries as “learning outcomes” I have ‘reified’—made into things—experiences of reality that have no “thingness,” cannot be defined in terms of learned *behaviors* and which nevertheless are the source of all that we admire in Western civilization.

Yorktown University has been directed by the U.S. Department of Education to engage in a process of hypostatization by which truth, justice, God, heaven and hell become things. Every institution up for re-accreditation will encounter similar pressures.

That begs the question, “How many learning outcomes can stand on the head of a pin?”

[Enter Peter Wood]

I also am not a student of the history of testing for intelligence, but unlike Dick, I come from a discipline—anthropology—in which attempts to measure intelligence have a long pedigree. My disciplinary ancestors filled up skulls with little beads to measure cranial capacity and sailed off to the far ends of the world to test the natives’ powers of analysis and deduction. Anthropology also has a much more variegated view of what life in the grip of tribal myths is like. An important anthropologist named Paul Radin wrote a book in 1927 titled *Primitive Man as Philosopher* that presented the cases of men who gained at least as much critical distance from their tribal myths as the pre-Socratics but who, unlike the pre-Socratics, had no literate civilization in which their insights could flourish. Context matters, as does literacy.

I am also the head of an organization that has complicated and sometimes conflicting views on testing. We are generally on the side of such tests as the AP Exams and the SAT as tools for helping to decide who gets into college. The current movements to eliminate from the SAT sections on which minority students tend to do poorly and the decisions of many colleges to drop the SAT

altogether are not based on considerations of what is best for students. But while higher education is demoting some tests, it is avidly promoting others.

In September 2005, then Secretary of Education Margaret Spellings announced that she was forming a Commission on the Future of Higher Education. Its focus was to be how well colleges and universities prepare students for the workplace. Spellings herself described the Commission as an extension of the No Child Left Behind policy, which of course had put standardized testing at the center of federal policy on the nation's schools.

The nineteen-member Commission issued its 76-page report a year later, on September 26, 2006. It addressed matters of "access" and affordability, as well as standards of instruction and accountability. The Commission had been chaired by Charles Miller, a former chairman of the University of Texas Board of Regents and fervent enthusiast for standardized testing. Miller founded the National Center for Educational Accountability, a private foundation that promotes such testing and he was a principal architect of No Child Left Behind. After Spellings received the Commission's report, she created a steering committee to implement it and appointed Miller to it, too.

That is to say that under Secretary Spellings, the view that some form of national standardized testing should be at the heart of American higher education was hardwired into the policy deliberations. In March 2007, Spellings hosted a summit in Washington titled "A Test of Leadership," in which she assembled 250 leaders from business and academe, in an effort to win them over to her plan. Higher education responded with some skepticism. David Ward, president of the American Council on Education, for example, resigned from the Commission and refused to sign its report. And the president of the Association of American Universities rejected the report for its lack of "nuanced understanding." The regional accrediting associations also expressed some misgivings. But Spellings' effort to promote more widespread testing began to bear fruit.

By September 2007, the *Chronicle of Higher Education* reported that "hundreds of U.S. colleges" were using standardized achievement tests in response to the federal initiative. And the public university system of Mr. Miller's home state of Texas embraced the new testing regime.

The center of that regime is a test developed by the RAND Corporation called the Collegiate Learning Assessment or CLA. Another test that is widely used is the National Survey of Student Engagement, but CLA is the principal outgrowth of the Spellings Commission's call for quality control and assessment in higher education. In 2007-2008 there were 210 colleges and universities using the CLA, including CUNY, the California State system, Macalester, George Washington, Ohio State, Spellman, Bob Jones, Florida State, Pepperdine, and Syracuse. CLA's sponsors describe it as "uniquely designed to test for reasoning

and communications skills that most agree should be one outcome of a college education.”

CLA is not intended as a measure of the success of individual students. Rather, it tests a sample of freshmen in the fall and another sample of seniors in the spring, and then attempts to give a measure of how much, on average, a college has succeeded in adding to its students’ reasoning and communication skills. More specifically, CLA describes itself as combining two instruments:

Performance Tasks such as writing “a memo or policy recommendation by using a series of documents that must be reviewed and evaluated. Completion of these instruments does not require the recall of particular facts or formulas; instead, the measures assess the demonstrated ability to interpret, analyze and synthesize information.”

Analytic Writing Tasks which are used to “Evaluate students’ ability to articulate complex ideas, examine claims and evidence, support ideas with relevant reasons and examples, sustain a coherent discussion, and use standard written English.”

On its face, CLA seems wholesome. Surely we would like college graduates to be able to write memos or policy recommendations, and to articulate complex ideas; and surely we would want a form of higher education in which graduating seniors are better at these sorts of things than entering freshmen.

Moreover, experts have been poring over the data from the first rounds of the CLA tests and have discovered interesting results. The Social Science Research Council (SSRC) published in November 2008 *Learning to Reason and Communicate in College: Initial Report of Findings from the CLA Longitudinal Study*, a study of 2,300 students at 28 colleges. The key findings were that on average:

1. Students who start out academically ahead increase their lead. Students with stronger high school grades and Advanced Placement outperform other students on the CLA as freshmen, and outperform other students by an even bigger margin as seniors.
2. Students who study in groups lose ground relative to students who study alone. Students who study alone improve their performance on the CLA substantially, and the more hours they study alone, the better their improvement. Students whose habit is to study in groups improve their CLA scores too, but by much less, and the more hours they study in groups, the more they erode their improvement on the test.
3. Working at an on-campus job for no more than 15 hours per week improves CLA performance; working more than that hurts CLA performance, as does any off-campus job.

4. If students perceive faculty to have high expectations of them, they do better on the CLA.
5. Students studying science, math, social sciences, and the humanities improve more on the CLA than students studying education, human services, or business.
6. It makes a big difference what college you attend. “29 percent of variation in longitudinal growth in CLA performance occurs across schools.”

Every one of those six findings comports with common sense, and one might ask whether large-scale and expensive testing is needed to reach such conclusions. I suppose the answer is while we didn't exactly need the additional evidence, it may be useful to have it.

The CLA substantiates that the intellectual capital students gain by working hard in a good K-12 program pays off in college, at least in terms of the sorts of things the CLA measures. It tells us the work-in-groups pedagogical fad of the last twenty years is educationally worthless. It says college education accomplishes more the more students focus on it, strive to live up to high expectations, and put themselves in the company of fellow strivers.

But there is actually more here than first meets the eye. I am in what follows indebted to my NAS colleague Tom Wood (no relation) for drawing out the implications of CLA for the argument made by Charles Murray. Murray, in his recent book, *Real Education*, argues that, since intelligence is mostly determined by genes and therefore being fixed, there is little point to sending large numbers of average and sub-average students to college. Tom Wood responds by citing the CLA data which show that large numbers of average and sub-average students actually learn a lot in college. The SSRC study I've been quoting found that students who took the CLA test in fall 2007 had improved their performance by .18 standard deviations by spring 2007. If the trend continues, their gain by graduation earlier this month would have been 0.36 standard deviations—which is substantial.

But that's not all. The CLA test is not measuring achievement in a given subject, but abstract skills such as analytic reasoning and problem solving. That means that students have intellectual gains in their capacity to think in rigorous ways—a crucial dimension of what we in the West call intelligence. This rather flatly contradicts the idea that intelligence is genetically fixed. Rather it demonstrates what educational theorists call “learning transfer.” Students who dive into studying a particular subject learn not only the specifics of that subject, but they “transfer” the patterns of thought they acquire to other tasks. In doing so, they get smarter.

That's what liberal arts instruction in college always purported to do. Now we have some substantial evidence that it actually happens—indeed that it generally

and routinely happens. It happens to some students a lot more than others, but it is the pervasive trend across all classes and all students and all colleges.

Tom Wood, who has written quite a bit for NAS in response to Charles Murray, has also drawn attention to two other phenomena that Murray downplays or ignores: neuroplasticity and the Flynn Effect. Neuroplasticity is the discovery over the last few decades that our earlier model of the brain as fixed at a very early age was wrong, and in fact, the brain grows and changes through a person's life. The Flynn Effect, named after the New Zealand researcher James Flynn, is the baffling observation that generation by generation IQ scores have risen. In fact, not just IQ scores, but scores on a wide variety of intelligence tests. It is as though everyone has gotten smarter—and not just a little but a lot smarter. Researchers have picked over this for an explanation but so far without success. The results, however, make no sense. If IQ were a measure of true intelligence, we would have to conclude that half of all children in 1932 were mentally retarded and that a person who was rated as very bright in 1900 would today be classified as severely disabled.

It would seem that people like Edith Wharton and Henry James, the Wright brothers, and Albert Einstein coped pretty well with their limitations. Either that, or there is some deep problem with the conception of the test. We can pretty safely rule out Darwinian selection. The Flynn Effect applies across the whole population, and evolution doesn't produce such sweeping changes in a few generations. Flynn's own answer is that intelligence tests are not measuring intelligence per se but rather a certain collection of cognitive skills that we are getting better at because of constant practice.

Or to put this another way, modern American society has emphasized and rewarded certain capacities that generally involve manipulating symbols on paper or in your head, and engaging in abstract reasoning. While people vary in their mastery of these skills, the actions in which they are embedded are the stuff of everyday life for us. Our technology gets inside our heads.

This conclusion seems to dovetail with Dick Bishirjian's views about the cultural predicates of intelligence. "I got rhythm" is a cultural statement, not a description of a genetic trait, says Dick. I'll add that, "I've got a high IQ" is no less a cultural statement, not primarily a description of a genetic trait.

I don't want to banish genes from the picture altogether. Anthropology has assembled a compelling picture of the development of the general cognitive capacity of our species from hominid ancestors. Roughly the hominid brain has more than tripled in size over the last 4 million years, from Australopithecus with a brain of about 400 cubic centimeters, to the modern human brain of 1,350 cubic centimeters. Brain size is not a simple index of intelligence. Neanderthal brains were considerably larger than ours, but Neanderthals remained stuck making exactly the same stone tools for 400,000 years without innovations.

Smaller brained Homo sapien was, by the lithic technology test of IQ, a whole lot smarter. Virtually every archaeological site of early Homo sapiens gives evidence of new techniques, new inventions, new combinations, new applications.

So if we take the broadest picture of humanity, we have good reason to think that we have an evolved capacity for intelligence and that part of the legacy evolution has given us is a kind of intelligence that is extraordinarily flexible. We can not only learn the same old stuff, like Neanderthals, but we can learn whole new ways of thinking that are suitable to the conditions we find ourselves in.

This doesn't necessarily mean that theorists who posit what they call *g*, the basic underlying intelligence regardless of the forms in which it is expressed are wrong. I am perfectly ready to say, for example, that Homo sapiens is, over all, smarter than Homo erectus was. But measuring *g* appears to defy current scientific ability. Remember the Flynn Effect.

A rule of thumb for modern Americans is that we tend to get better at what we can measure. We have much less patience for and orientation to achievement in the realm of the immeasurable. Confronted with that which the calipers cannot encompass or the galvanometers register, we turn our ingenuity to the search for proxies and approximations. Psychologists test pupil dilations of eyes that gaze on a succession of faces to discover in what exact proportion of lip and nostril lies beauty. Statisticians slice the play of professional athletes into ever more precise components in search for the formula that wins games. The commendable desire *to know* is channeled into the stultifying impulse to systematize and count.

I say stultifying, but is that a fair characterization? Several months ago the *New York Times Magazine* ran a cover story, "The No-Stats All-Star," about Shane Battier, a basketball player with the Houston Rockets who seems somehow to help the teams he plays on to win despite his mediocre profile in the categories that usually get counted: "points, rebounds, assists, steals, and blocked shots." But the story was really as much about Daryl Morey, as Shane Battier. Morey is an analyst who works for the Rockets and who specializes in developing new sports statistics. It was Morey who discerned Battier's mysterious knack for making teams he plays for play better. So, of course, Morey has undertaken the task of breaking Battier's performance into countable parts. If Morey can figure out how Battier does it, presumably the Rockets can teach the same skills to other players.

Basketball statistics are a long way off from the Collegiate Learning Assessment, but there is a common thread. Daryl Morey has dedicated himself to discovering measures of success hidden below the surface of a game that involves, as the article puts it, a peculiar tension "between the interests of the team and the interests of the individual." In Morey's view, individual players know what they

are rated on and work hard to improve their statistics in those categories, regardless if doing so degrades the team's overall performance. The rule of thumb applies: we get more of what we measure, whether it is IQ points or sensational dunks.

Of course, Morey would like to count the less obvious things that capture a player's contribution to the function of the whole team. If he succeeds, he will create a new set of incentives and we may end up with leagues full of Shane Battier-style basketball players—a kind of Flynn Effect for the parquet.

It is not that the systematizing and counting don't work. They *do* work, but they work primarily to foster whatever it is that falls within the system and gets counted. They exert an invisible pressure towards standardization. They dislike whatever is anomalous, original, new, or exceptional, and have no capacity at all to register the magnificent. Artistic splendor, wisdom, and justice are foreign to assessment in all its forms, although it is not beyond the assessors to give it a try.

This seems to me the central problem with the initiative that Secretary Spellings and Mr. Miller set in motion and that now lives on in higher education. It is a system that ignores the most important things, and by ignoring them, marginalizes them. We get more of whatever it is we measure, and we sure aren't measuring the attainment of capacious character and the internalization of a civilization. Those are intrinsically un-measurable, though not un-discernable.

I don't mean to say the tests are useless. CLA has its virtues. It says to those for whom nothing is real unless it is counted that liberal education really does shape students' minds for the better, and that the traditional techniques of higher learning—high standards, good teachers, long-term preparation, dedicated study, and individual concentration—are solid and better than the alternatives, such as remediation, self-esteem-oriented teachers, and group study.

Spellings introduced this current wave of testing out of sincere concern that college graduates frequently are short-changed. They receive diplomas, but not much in the way of education. Since the federal government has no Constitutional claim of authority over higher education and even her statutory authority was sparse, Spellings looked around for whatever tool she could find that might help her in a campaign for more educational rigor at the college level. What she found was accreditation. The Department of Education authorizes the accreditors, and colleges must have accreditation in order to receive federally guaranteed student loans and grants. So accreditation was a point of leverage. Fortuitously, most of the college accreditors and all of the large regional accreditors were already committed to "assessment," an educational spin-off of the old Total Quality Management approach. Adding outcomes assessment to this regimen was not a great stretch.

That is how we came to this point of attempting to quantify the results of a liberal education. I fear it will have a Neanderthalizing outcome: it will tend towards doing the same thing over and over to stone-tool perfection. However much it seems to leave open to colleges and universities the choice of how to proceed, it is in fact setting the metric of what will be measured. For better or worse, we are going to get a whole lot more of whatever is on that scale. It just won't be, to borrow a phrase, *real education*.