

# Academe Online

## Beyond Crude Measurement and Consumerism

*We ought to be up to the task of figuring out what it is that our students know by the end of four years at college that they did not know at the beginning.*

By Stanley N. Katz

Why should faculty members support efforts on their campuses to assess student learning outcomes? A great deal of ink has been spilled in recent years by a small number of professors and a much larger number of educational administrators arguing for assessment and pleading for greater faculty support of institutional assessment efforts. It is now a truism that faculty reluctance is the single biggest impediment to the adoption of systematic institutional efforts to measure learning outcomes. The next logical truism is that if we are to create an assessment culture in higher education, we must convince the faculty to support systemic assessment across the four years of college. Are these truisms true?

I first want to acknowledge the potential perils to faculty acceptance of outcome assessments. Paranoids do have real enemies, after all. Since the appointment of U.S. Secretary of Education Margaret Spellings in the most recent Bush administration, it has been clear that a number of those who support outcome assessment (like Spellings herself and Charles Miller, chair of Spellings's Commission on the Future of Higher Education) do so in the name of educational consumerism—they want assessment so that students and their parents can comparison shop. The metaphor Spellings repeatedly used was that Americans had much more information on buying used cars than they had in choosing educational institutions and that it was time they got a chance to “kick the tires.” This sort of crude consumerism (cost was the main criterion) is not in itself a threat to the autonomy of individual colleges or their faculty. Indeed, such an approach might have done a more adequate job than *U.S. News and World Report* and other media currently do. Carried out poorly, however, it could (and probably would) provide inaccurate and misleading information.

More significant for most of us in academia, however, has been the support for outcome assessment that comes from those committed to external imposition of accountability—those with the legal and financial power to reward or punish institutions that do not meet their expectations. This includes the federal secretary of education or state education officials. The concern here is that even beyond the consumerism mentality, crude measures of educational “success” would be developed through high-stakes testing or other blunt mechanisms and that these measures would constitute the primary methods for evaluating, rewarding, and punishing faculty “performance,” along the lines of the current movement to hold K–12 teachers “accountable” and to get rid of those who do not “measure up” to the prescribed standards

### Assessment Instruments

It is certainly possible to imagine that the imposition of such an assessment system could pressure faculty unions to abandon hard-won gains and reasonable teacher prerogatives in pursuit of incentives, such as those currently proffered in the Obama administration's K–12 “Race to the Top” program. The point is that accountability-based assessment, when imposed from outside the university, is understandably problematic and potentially liable to abuse.

But externally imposed assessment is the worst-case scenario. Faculty members should be capable of contemplating more benign, educationally helpful uses for sophisticated measurement of student learning outcomes. For the sake of argument, let us assume that we could agree on appropriate measures of student success and that we could measure them reasonably accurately. I do not see why either faculty members or their institutions should oppose such an approach to institutional evaluation. In fact, the widening acceptance of this idea has led recently to the development of outcome-assessment instruments such as the National Survey of Student Engagement (NSSE, or “Nessie”) and the Collegiate Learning Assessment (CLA). NSSE and the CLA are, however, very different sorts of assessment instruments. NSSE is designed to obtain, on an annual basis, information from large numbers of colleges and universities about student participation in programs and activities that the institutions provide for learning and personal development. It inquires about institutional actions and behavior, student behavior inside and outside the classroom, and student reactions to their own collegiate experiences. The CLA, in contrast, tries to assess an institution's contribution to student learning by measuring the outcomes of simulations of complex, ambiguous situations that students may face after graduation. The CLA attempts to measure critical thinking, analytical reasoning, written communication, and problem solving through the use of “performance tasks” and “analytical writing tasks.”

There are also other assessment instruments that use different techniques and measure different aspects of student learning. These technologies are still fairly new, but the most recent attempt to compare them seems to indicate that the most commonly used “produced comparable outcomes at the institutional level, based on having been administered at a diverse range of 13 institutions, big and small, public and private,” according to *Inside Higher Ed*, reporting on the results of a federally funded study in November 2009.

Anyone who believes that educators have a responsibility for critical and holistic self-evaluation of the process of education ought to support the most effective forms of outcome evaluation. That said, the jury is still out on what the most effective existing mechanisms are.

## The Faculty Role

On the whole, the United States has lagged behind Europe in the use of national assessment strategies, although the strategies currently used in the United Kingdom, Germany, and Australia purport to measure the quality of research, not the quality of undergraduate learning. Up to this point, public debate on assessment has been neither edifying nor helpful—the Spellings “debate” seemed to me like ships passing in the night, with no truly productive engagement on the most important issues, such as whether students are making satisfactory academic progress or how we would know whether they are making any cognitive progress at all. Most often, legislators and bureaucrats bluster and then do little to implement assessment strategies, while the universities dodge and weave in response to perceived threats, and then do little or nothing to carry out their boasts that they are fully capable of self-evaluation.

Should faculty members be expected to care about, much less take responsibility for, assessment of anything other than the work of students in their own courses (or, at most, their own departments)? Reasonable people disagree on the answer to this question. Of course, the underlying problem is that it probably makes little sense to generalize about “the faculty” in such a large and diverse system of higher education as we have in the United States. Still, the AAUP is an association of faculty members, and it is important to insist that we have common interests beyond our historically shared concern for academic freedom and faculty governance.

Are faculty members to blame for the slow progress of institutions in implementing assessment programs? Sure, but plenty of other factors also bear some responsibility. The first and most obvious is the technological limitations inherent in creating reliable measures (beyond the individual course grading process) of student learning outcomes. We have enough difficulty, after all, convincing ourselves (and sometimes our institutions) that we can grade with sufficient precision. The second problem is the expense of new instruments for institution-wide student outcome assessment, such as NSSE and the CLA. Beyond expense lie legitimate concerns about how effectively these instruments measure student learning. I will return to this important issue. A third and more basic problem is the faculty’s lack of agreement about just what it is we want to assess. Even if we can stipulate that the main objective of higher education is to produce student learning (not so clear, I am afraid), we agree neither on what it is we want students to learn (given the wide diversity of goals in U.S. higher academia, however, has been the support for outcome assessment that comes from those committed to external imposition of accountability—those with the legal and financial power to reward or punish institutions that do not meet their expectations. This includes the federal secretary of education or state education officials. The concern here is that even beyond the consumerism mentality, crude measures of educational “success” would be developed through high-stakes testing or other blunt mechanisms and that these measures would constitute the primary methods for evaluating, rewarding, and punishing faculty “performance,” along the lines of the current movement to hold K–12 teachers “accountable” education) nor on how best to articulate how the learning process works. A final problem is the broad range of faculty roles across the institutional map of colleges and universities. Should or can we expect the faculty in research universities and professors in community colleges to think identically about assessment? To whom does responsibility belong for either longitudinal or cumulative assessment: the faculty or the institutional research branches of university administration?

Student learning outcome assessment is a particularly delicate area, however, since it constitutes an intersection between individual faculty prerogative (the right to evaluate one’s own students) and the institutional interest in promoting learning across the curriculum and over the span of students’ college attendance. Historically, the dominant faculty attitude has been that so long as individual faculty members evaluate student course performance fairly and carefully, their responsibility to the institution should be considered fulfilled. I suppose one could also say faculty members had, in the recent past, an additional duty to assess outcomes in their field or department, sometimes through capstone exercises such as general examinations and senior theses. But the common thread of assessment has been one that runs from individual faculty members to individual students.

This faculty attitude probably has to change if we are to take seriously the emerging conception of institutional responsibility for overall student learning outcomes. Courses and capstone exercises are only a few of the many student learning experiences in an undergraduate career, and in many situations no “faculty member” is in an obvious position to evaluate the quality of the product or of the experience. Assessment instruments such as NSSE and the CLA have attracted wide interest because they attempt to evaluate the entirety of students’ collegiate learning. Is there any reason why the professoriate should be suspicious of (or opposed to) this new mode of evaluation?

## Formative Assessment

My brief is for the establishment of *formative* types of assessment—that is, for the use of the evaluation of learning outcomes to improve teaching and learning on an ongoing, continuous basis. This is not something that even the employment of NSSE or the CLA will necessarily provide. The consumerists and some other supporters of outcome evaluation are more interested in cross-institutional, comparative assessment data. The potential for political misuse of such information is substantial, and we certainly

need to keep that in mind. Many useful and legitimate uses for such comparative information exist, but they are not the uses that most interest me, nor will they be of interest to most faculty colleagues.

I can well imagine university teachers who oppose even formative assessment on the grounds that no one should be able to tell them how to teach, but we have all submitted to student course evaluations for many years, no matter how dismissive we may be of the current forms of that technology. It is hard to imagine a principled objection to careful evaluation of learning outcomes or to thoughtful suggestions for improvement in pedagogical strategies. The opposition will come when administrators either demand that individual faculty members employ specific pedagogical techniques or when they begin systematically to base decisions about tenure, promotion, and compensation on the evaluation of teaching solely as judged by learning outcomes, as is already happening in K–12 education.

My guess is that a more important and more sensitive area for faculty concern in an environment of genuinely formative assessment will be the articulation and specification of learning goals. To some extent, all meaningful assessment is based on the faculty's ability to identify specific benchmarks of success in learning. This implicitly engages the objective of grading in courses, after all, though many of us never think of the evaluation of papers and examinations in such a structured way. I think we mostly look for generalized demonstrations by students that they have read and thoughtfully considered the material we present to them. At least in the humanities fields I know best, we seldom have clear notions of more precise learning outcomes.

In my own field of history, for instance, no teacher I know values the capacity of her students to recite dates or to rehearse the arguments of the secondary materials assigned. Most of us would probably rather say that we are looking for broad demonstrations of historical understanding, without any expectation of a limited number of particular demonstrations. My sense is that, in history, we have not been very good at developing the sorts of new assessments that would enable us to estimate student achievement more precisely and to compare one student to another rigorously. In the early 1990s, when attempts were made in the name of high-stakes testing of high school students to draft national standards in the major secondary school fields of study, serious political conflict emerged over the content of the standards. The historians among us, for instance, will remember all too well the painful attempt, starting in 1992, by Gary Nash of the University of California, Los Angeles, to develop standards for assessing high school U.S. history. Political conservatives mounted a successful counterattack, led by Lynne Cheney, then chair of the National Endowment for the Humanities. In the end, the problem Nash confronted was not simply a battle in the culture wars but, more important, a lack of certainty within our own discipline about what constitutes particular learning outcomes for our students and academic historians' lack of both the political will and the political smarts to defend successfully the standards developed by disciplinary professionals.

I have for some time advocated institutional acceptance of formative outcome assessment, since I believe that undergraduate education ought to amount to more than an accumulation of work in separate courses—it consists of the whole mosaic of learning experiences over the years of college enrollment. If that is true, then we need to know what students have learned from the entire range of their learning experiences. Simply put, we should try to learn what the students know that they did not know when they entered college—and what they can do, intellectually, that they could not do when they completed high school.

Pat Hutchings of the Carnegie Foundation for the Advancement of Teaching [has recently written](#) thoughtfully about the need for greater faculty involvement in assessment, on the very plausible theory that faculty engagement with and support for the evaluation of long-term student learning is necessary to institutional acceptance of the assessment challenge. She acknowledges serious obstacles to such faculty engagement: “the language of assessment has been less than welcoming”; faculty members are not trained in assessment; “the work of assessment is an uneasy match with institutional reward systems”; and faculty members have not seen much evidence that “assessment makes a difference.” But she also sees signs of greater faculty interest and has recommendations for increasing faculty involvement, including building broader assessment using the practice of grading and the “regular, ongoing work of teaching and learning”; making “a place for assessment in faculty development”; building assessment into the training of graduate students; involving students in assessment; and so on.

I think Hutchings is right to pitch the notion of supporting assessment to the faculty. On the one hand, thoroughgoing assessment programs probably will not happen if faculty members do not believe in and support them. On the other hand, I think conscientious faculty members should feel an obligation to find out whether their students are learning what they are teaching—or learning anything at all. Do faculty members need more feedback from their students than they receive from the end-of-term student evaluation forms? Should we not worry whether students can relate what they learn in one course to other courses, and to the multiple other learning experiences of their undergraduate years? I would hope that most teachers would answer “yes” to these questions. Of course, most of us take some responsibility for the integration of student learning within our departmental and disciplinary specialties, although my own view is that we could frequently be more systematic in our evaluation of this aspect of student learning. Some faculty members will say, further, that existing capstone exercises give us sufficient information on four-year student learning accomplishment, but in my experience this is not a persuasive response. My institution, for example, requires all seniors to write a demanding thesis, and this experience certainly does test their research capacity. But the thesis is still a single exercise on an ordinarily narrow topic, and it frequently does little to test the student's capacity to integrate the knowledge learned over four years. Neither does it assess the extent to which students can bring the full range of their learning experiences (including experiential learning, for example) to bear on their “academic” knowledge.

Additional demonstrations are necessary for us to assess the full range and accomplishment of learning, and it is my impression that too few colleges and universities provide opportunities for those demonstrations. As I have already indicated, instruments such as the CLA are intended precisely to determine what mental operations students are capable of at the end of their fourth year of college that they could not manage in their first year. I am no expert on the construction of assessment mechanisms, but I feel confident that there are various techniques that would enable us to get a more secure sense of the specific ways our students have grown intellectually during the course of college. Some will doubtless be newer and conceivably better versions of instruments such as the CLA. Others will be more intensive and sophisticated capstone exercises devised by each institution.

If we could put a man on the moon, we ought to be able to figure out reasonable solutions to the technology of assessing student learning. And if we can, we ought to be able to improve quite dramatically the quality of the education we offer our students.

Thus, my conclusion is that those of us who want to take ownership of the evaluation of undergraduate education must devote considerably more time, effort, and ingenuity to the assessment of student learning over the life course of undergraduate education than we have been doing. We need to enter into the debate over which of the newer assessment technologies will tell us what we need to know (at reasonable cost). And we must also join the effort to ascertain how this feedback can be used on a continuing basis to adapt our educational strategies to what we learn about how students learn. We know more than we used to about learning outcomes, but not enough. We know far too little, however, about how to put the knowledge we do have to practical use in transforming both our pedagogical technique and curriculum design to enhance student learning. I can think of no more important challenge for college and university faculty members as teachers.

Stanley N. Katz is president emeritus of the American Council of Learned Societies and director of Princeton University's Center for Arts and Cultural Policy Studies. His recent research focuses on the relationship of civil society and constitutionalism to democracy and on the relationship of the United States to the international human rights regime. His e-mail address is [snkatz@princeton.edu](mailto:snkatz@princeton.edu).

Comment on this article by writing to [academe@aaup.org](mailto:academe@aaup.org).

Note: Comments are reviewed prior to posting and will not appear on this page immediately.

## Comments

Stanley makes some very interesting, but wholly un-compelling arguments for university/college faculty to engage in formative assessment beyond faculty or departmental scholarly interest. His points regarding the real outcomes behind the exhortation of comparative assessments of politicians, federal officials in the Obama and its predecessor administration of George Bush, and higher education administrators are an important backdrop to the "rock and hard place" faculty members are placed when even considering examining the effects of their teaching and what they believe their students should learn (and do, I might add). Faculty resistance to pressure to "be reasonable" is in reality a sword of Damocles. As long as higher education officials and federal governments are essentially pro-business and pro-profit before they are pro-learning and pro-human need, engaging in development of assessments with anything other than a faculty or departmental interest and at a discipline level tied to faculty scholarly pursuits will only result in opening the door to the "consumerism" Dr. Katz correctly decries.

The primary issue, implied by Dr. Katz, but not wholly stated, is that determining the "worth" of our teaching in an environment where that worth is challenged by antagonists to the pursuit of academic excellence is an unworkable goal. Moreover, it is indeed premature because a cardinal rule of assessment is that one must always begin by asking "why"; that is what is the purpose of the assessment (e.g., Salvia & Ysseldyke, 2010)? To determine one's purpose for assessment presumes there is a common base of knowledge (never mind philosophical agreement) that any assessment worthy and viable in which to engage exists. "What do faculty members at the University of Minnesota want their students to know and do?", How about those at Harvard, St. Paul Technical College, Pomona Community College . . . or, for that matter, the Center for Arts and Cultural Studies at Princeton?

To be sure, I actually do believe that it is possible for university and college faculty to forge a path to excellence in academic pursuits for all their students. It just must actually follow a path that accomplishes our common "historically shared concern for academic freedom and faculty governance". Without a free university system firmly in the control of people who produce and facilitate the learning—students, faculty, and staff—we can never hope to achieve the common purpose that obviates "consumerism" and promotes the democratic ideals and humanism inherent in the pursuit of learning.

M. B.

Hello,

Near the end of Stanley Katz's recent column, "Beyond Crude Measurement and Consumerism," he writes, "my conclusion is that [we] must devote considerably more time, effort, and ingenuity to the assessment of student learning." What do we stop doing in order to shift time, effort, and ingenuity into this dubious and notoriously vague endeavor known as "assessment"? Do our

teaching loads lighten? Do we cut back on research? Do we walk away from other service commitments? This is the obvious question that administrators, legislators, and bureaucrats have no answer for. No one "makes time." We can only use the finite amount of time that the planetary rotation determines for us.

Sincerely,  
R. R. L.

---

The three feature articles in the September-October article of *ACADEME*, by Katz, Eaton, and Gilbert, dealing with assessment and accreditation, deal with only half the inherent problem. Thus, much of the reason for faculty opposition to the expansion of these activities is missed. Present and proposed activities in these fields focus primarily on assessing the extent to which student have learned facts, theories and skills and the extent to which institutions produce students that have learned facts, theories, and skills. But there is an entirely different dimension of a liberal arts education that these articles do not mention and for which there are neither creditable means nor proposals for assessment and accreditation. It is the neglect of this second dimension that causes and should cause much of the opposition.

A liberal arts education is meant not only to produce "educated" persons but also to produce persons who live a "good" life. To assess whether an institution is successful in this respect would require assessing the lives of alumni after a major part of these lives have been lived, say 25 or 30 years after graduations. Further, there seem to be no agreed criteria for such an assessment. Perhaps we should ask questions such as:

How many ballets, operas, or symphonic concerts did you attend last year? How many serious books have you read in the last six month? How often have you been divorced? Are you currently acting as a volunteer worker or a civic or charitable institution? Have you held any leading positions in government? How many persons work, directly or indirectly, under your supervision? Etc.

The danger is that, when all the attention is paid to one dimension, the other, equally valuable, dimension disappears.

W.H.

**American Association of University Professors**  
1133 Nineteenth Street, NW, Suite 200  
Washington, DC 20036  
Phone: 202-737-5900 | Fax: 202-737-5526