

Assessment as a Subversive Activity

Dave Porter

The 2011 volume of the *AAUP Journal of Academic Freedom* contained two articles critical of “the relentlessly expanding assessment movement.” Accompanying these provocative essays was a challenge for someone “on the other side of [this] question” to answer these criticisms. The essays by John Champagne and John Powell are packed with “philosophical, political, and pedagogical” concerns about assessment. Powell’s final sentence aptly summarizes his perspective:

Outcomes assessment’s ... origins are suspect, its justifications abjure the science we would ordinarily require, it demands enormous efforts for very little payoff, it renounces wisdom, it requires yielding to misunderstandings, and it displaces and distracts us from more urgent tasks, like the teaching and learning it would allegedly help.¹

Champagne’s conclusion focuses even more ominously on the potentially dire consequences of assessment activities:

Given the current political climate ... I fear that assessment will ultimately provide the corporate university with another alibi for silencing dissent, subjecting faculty members to increasing surveillance, and eroding faculty rights.²

Both articles contain a plethora of arguments, anecdotes, and assertions supporting these disturbing conclusions. After reading these essays, many faculty members might conclude that the appropriate response for any sane and conscientious educator would be to withdraw and await the passing of what both authors castigate as being the latest management fad. While I am sympathetic to some of the observations and opinions offered by Powell and Champagne, an explication of “the other side” will require a reframing of many of the issues they raise. Problems they identify should not be attributed to outcomes assessment itself but to more general problems such as the lack of institutional integrity and the manipulative managerial style of administrators who do not understand the learning process and educational systems sufficiently to implement assessment programs effectively. To use John Powell’s term, “ignoramusness” is not a malady limited to our students; we need to recognize that this is a condition that afflicts faculty and administrators as well. We cannot improve unless we are willing to accept the fact that we are imperfect—and our blots and blemishes are what assessment, when done well, can show us. Used in this way, assessment has power; one might even consider it the ultimate subversive activity. It provides a mechanism through which the authority of the institution might even contribute to the kind of transformation and liberation most valued in the liberal arts tradition.

Assessment has been an integral part of my involvement in higher education over the last three decades. As a classroom instructor and fledgling educational researcher at the Air Force Academy, I observed how the systematic collection of data relevant to student learning could be used to increase awareness and eventually improve institutional policies and programs. As an accreditation committee member and liaison for three very different institutions (the U.S. Air Force Academy, Western Governors University, and Berea College), I’ve seen how greater emphasis on the use of evidence to assess student learning by accreditation bodies created conditions conducive to institutional enlightenment and growth. As a senior academic administrator, I’ve witnessed the power of evidence to alleviate prejudice, overturn antiquated institutional policies, and raise student retention and graduation rates to record levels at Berea

College. As a member of accreditation teams in three regions and as chair of two of these teams visiting institutions in the California State University system, I've observed instances where effective assessment has helped align resources with institutional priorities and increase conversations across campus about how to enhance as well as measure student learning.

There are many individuals better qualified to address the serious concerns raised by the articles in the last volume of the *Journal of Academic Freedom*. My positive experiences with assessment are just as anecdotal as the litany of egregious assaults attributed to assessment contained in the essays by Champagne and Powell. We all fall short of Powell's quite reasonable request for rigorous and comprehensive scientific analysis. However, I'd like to revisit some of the problems they present in broader social and organizational contexts (some of which were introduced in their essays). Hopefully, the injection of a few explanations, alternative interpretations, and definitions might lead a few educators to conclude that assessment in general, and outcomes assessment in particular, represents a greater opportunity than threat to academic freedom and shared institutional responsibility.

There are several parts to my argument: 1) assessment is an integral part of learning (and hence education), 2) assessment is a necessary function of effective and adaptive organizations, and 3) involvement in assessment activities is particularly important for the AAUP and its members. Assessment is about creating a culture of evidence that is much more than merely collecting piles of data and accumulating a multitude of meaningless measures. (It seems to me that Powell's claim that the mere use of this phrase shuts down thinking reflects his own misunderstanding of its meaning.) A relatively high level of both mutual trust and systemic understanding are prerequisites to developing an effective assessment program. When assessment fails, more often than not, it is a reflection of deeper individual and organizational deficits. If one plans to speak truth to power, there is no better ally than the evidence assessment can provide.

Assessment Is Integral to Learning and Education

Education is both a craft and a profession. In addition to professing the wisdom of their particular disciplines, university professors should understand and be able to apply knowledge about the process of human learning. Unfortunately, many of the most critical features of human learning are often glossed over or forgotten altogether by the intense focus on mastering disciplinary lexicons and epistemologies. Our collective failure to recognize that human learning is at the heart of most things we educators do weakens our academic communities as well as our profession. The next few paragraphs present a brief review of several important ideas about human learning that I believe are relevant to all educational systems, our role as professors, and our appreciation of the potential benefits of assessment.

Human curiosity is innate; learning is natural. We humans are born with relatively few instincts, so we must acquire most of the knowledge and skills necessary to survive and reproduce through learning. Learning activities typically consume childhood—some learning occurs through structured experiences, but many things are discovered through unplanned and informal interactions with the environment and other people. Although B.F. Skinner and other behaviorists once argued that human behavior could be explained solely by environmental consequences and schedules of reinforcement,³ most contemporary psychologists appreciate the critical importance of social and cognitive factors. Humans have the extraordinary capacity to construct complex internal representations of the external world, that is, “mental models.”⁴ We test these models repeatedly against the real world by choosing actions and experiencing or observing consequences. Through reflection, we use feedback to refine our internal representations and gradually increase the effectiveness of our actions to achieve the results we intend.⁵

Developmental psychologist Jean Piaget’s distinction between the complementary processes of *assimilation* and *accommodation* is important. Most of the time, our mental models are adequate to guide us through the external worlds we inhabit (physical, social, and conceptual). As long as this is so, we maintain the models we have with only occasional small adjustments. This is the process of *assimilation*. In contrast, there are times when our internal representations

of reality just don't fit the circumstances in which we find ourselves; they do not allow us to predict the outcomes of our actions. This is when *accommodation* occurs and we must make significant adjustments to our internal representations of the external world. Piaget suggested such accommodation marks children's transition from one stage of cognitive development to the next.⁶

Many others have used the notion of stages of development to describe differences in mental models associated with different types of activities at progressive levels of moral development, psycho-social development, and even spiritual development.⁷ Psychologist William Perry developed a stage model of human learning and development that relates specifically to higher education. After extensive interviews with male undergraduates at an Ivy League college, Perry suggested the ways in which undergraduate students perceive the world are likely to change in a particular sequence. Many students start with a firm but naive attachment to dualism (a belief in the black-and-whiteness of all issues and subjects). However, some students, realizing that such a view is incompatible with the real world, gradually come to accept increased complexity and multiplicity and eventually arrive at a perspective Perry labeled "contextual relativism."⁸ A subsequent study of undergraduate women disputed some of the finer points about Perry's later stages of cognitive development and suggested distinctive differences in the paths of most women undergraduates.⁹ However, the basic scheme remained similar to Perry's model.

The process of observing the consequences of one's actions or choices involves *assessment*, and, without it, learning simply does not occur. Students need feedback on their performance—especially when they are entering a new field or discipline. Most of this initial feedback should be in the form of reassurance and encouragement that the mental models they already possess can help them frame and resolve problems in a new realm. However, students also need to recognize the ways in which the models they already possess may be inadequate or insufficient for resolving more complex problems in new knowledge domains. It is at this critical juncture that the most intense and transformative learning is likely to occur. It is only as students attain

the higher levels of intellectual development that the complicated and contentious aspects of knowledge and its relationship to power can be meaningfully examined. Higher learning certainly appears to be more complex, more subtle, and more nuanced than what occurs during Piaget's initial stages. However, there seems to be no reason to assume that it is a fundamentally different process or any less measurable.

Champagne asserts that the essence of liberal arts education is "a fostering of the life-long attempt to interrogate, understand, and be unsettled by the limitations of one's own thinking." His statement captures much of what I also have come to value about higher education. It seems to be a general description of the process of *accommodation* that occurs as students transition from one way of thinking or knowing to the next higher, and more sophisticated, level or stage. It is Champagne's next assertion that appears to me to be the greatest flaw in his argument. He asserts that this lofty objective "is not a set of contents or even a skill but rather a praxis that cannot be measured by any test."¹⁰

Powell characterizes some of those who oppose assessment as fearing that "outcomes assessment would turn all art into calculi and all wisdom into idiot savant checklists."¹¹ In fairness, Champagne makes it clear that he does not believe that evaluating student learning is completely impossible (otherwise, as he admits, he could not assign grades). However, he argues that the two particular "labor-intensive techniques" (daily response papers and daily classroom conversations) he currently uses are valid, appropriate, and sufficient, then grumbles that those in "the corporate university" impede his ability to use these tools effectively by requiring him to use other, less effective ones.¹² Unfortunately, this position precludes the opportunity to assess critically the degree of convergence between the measures Champagne accepts and those he rejects. Why not use evidence rather than a rhetorical argument and fervent opinion to settle the matter? (My guess is that there would be much greater convergence between the two measures than Champagne would predict.)

I suspect if we questioned students in any class, including one of Champagne's, we would find differences in what they know, what they can do, and their awareness of, and relative

comfort with, their own intellectual limitations. Champagne would, in all likelihood, be able to distinguish students who “got it” from those who did not. (One of the great challenges of classroom teaching is this: How do we create coherent conversations that diminish the gaps between our students by increasing the knowledge and skills of those who lag behind while simultaneously challenging those who are more advanced to accept greater responsibility for the learning of others in addition to new substantive intellectual challenges?) The methods and the tools Champagne describes appear to be an effective use of “formative assessment,” which provides him with all the information he believes he needs to guide the learning process.¹³ I suspect he does this very well. However, this is not a reason to ignore or disparage the potential added value of additional summative assessment.

Formative and summative assessment can serve complementary roles; both can inform the process of education. Assessment, like the scientific method, can be seen as a means to overcome our considerable capacity for self-delusion. Convergence is the key. To the extent that we can distinguish students who “get it” from students who don’t, examining patterns across classes, programs, and departments can benefit students, faculty members, and the institution. Barbara Walvoord and Virginia Johnson Anderson, for example, offer a number of techniques and tools through which the reliability and validity of grading can be increased to provide viable evidence for subsequent higher level assessment.¹⁴ Similarly, Thomas Angelo and K. Patricia Cross provide many examples of effective formative classroom assessments tools that can contribute important evidence to overall assessment plans.¹⁵ The notion of “embedding” assessment within course work is itself a powerful pedagogy as well as a relatively low-cost way to systematically gather evidence for broader assessment plans.

One of the concerns Powell raises about “outcomes assessment” is the consequence of assessing outcomes independently of inputs and processes.¹⁶ Both science and assessment require an understanding of causal contexts (systems). For educational systems, information about inputs and processes is vital to be able to interpret and apply the results of outcomes assessment. John Darley’s Law of Criteria Control Systems suggests that when authorities

impose statistical controls based on selected measures of outcomes (such as in “dashboard” reporting systems), the consequences can be dire.¹⁷ Such systems inevitably distort the data and corrupt the subordinates expected to provide it. The first chapter of Steven Levitt and Stephen Dubner’s popular work, *Freakonomics*, provides several vivid examples of how Darley’s Law undermined the high stakes testing associated with the No Child Left Behind Act of 2001. Teachers (especially those in low-performing classrooms) began cheating. Levitt and Dubner also explain how appropriate statistical analyses were used to discover what was really going on.¹⁸ One of the major difficulties with educational outcomes assessment is that it is often used by those who have little understanding or appreciation of the educational systems they are trying to assess. They search for solutions without ever really understanding the problems.

However, there are measurable differences in knowledge and skills related to important liberal arts outcomes. Student attitudes (which can be measured as well as other outcomes) also relate to the attainment of these outcomes. I suspect that the mysterious and illusive “praxis” Champagne claims cannot be measured may be the product (that is, represented by a statistical interaction) of these three dimensions: students’ knowledge, skills, and attitudes.¹⁹ Claiming that these dimensions of growth are measurable is not, however, a claim that any of them can be measured perfectly or even precisely. Even measures of content knowledge, which many educators assume to be relatively easy and objective, are likely to be fraught with error and bias. However, both critical thinking and social scientific research require that we learn to contend with imperfect information and make decisions in the face of uncertainty. As Donald Broadbent concluded his defense of empirical psychology:

It seems to me inevitable that an approach ... through the armchair, by the exercise of fallible human reason, intuition and imagination, will lead one to hostility and disagreement If we refuse to use experiment and observation on other human beings, we start to regard them as wicked or foolish.²⁰

A brief account of a large assessment project at the United States Air Force Academy conducted during the mid-1990s may help illustrate this point.²¹ This project was facilitated by

an institutional interest in total quality management and continuous quality improvement (both of which Powell would undoubtedly identify as “management fads”). However, the general principles upon which this project was developed are also consistent with basic principles and processes of more general approaches such as the scientific method or critical inquiry. Over a nine-month period, fifty faculty members, assigned to seven independent assessment teams, reviewed course portfolios from each of the academy’s thirty-five “core” (required general studies) courses and evaluated the contributions each of these courses made to cadets’ fundamental knowledge, critical thinking ability, and intellectual curiosity. Each course portfolio contained information from the College BASE (a standardized test of subject knowledge and critical thinking ability given to all freshmen and seniors); student course critique ratings for two consecutive semesters; course characteristic descriptions from course administrators along with syllabi, sample assignments, and tests; instructor emphasis surveys from all faculty members teaching the courses; and critical thinking climate surveys completed by students at the end of each course. Although most assessment team members spent less than twenty-five hours during the year-long project, collectively, this represented a huge institutional investment in assessment. Some argued that the actual results did not justify the expense, but reports from most of those who participated in this process consistently touted its benefits as a faculty development program and an opportunity to enhance communication among faculty members across academic departments and disciplines. Subsequent events, including two successful accreditation visits, also supported the institutional value of this ambitious assessment project.

Two of the most important questions concerning any type of measurement are reliability and validity. The question of reliability is one of control, precision, and consistency. Since the design of this project required two different faculty assessment teams to assess each of the thirty-five core courses, simple measures of inter-rater (group) reliability were easy to compute. For each of the three criteria (the distinctive contributions of the course to cadets’ knowledge, skills, and intellectual curiosity), the common variance between the two rating groups was

between 45 and 48 percent. While this level of convergence is less than one might hope in developing a nationally standardized test or experimental protocol, it is far better than would be expected to have occurred by chance. A subsequent study asked all graduating seniors to identify the core courses that had contributed the most to their attainment of each of the same three educational outcomes. The results of this survey showed that about the same proportion of variance in ratings (about 50 percent) was common to the ratings that had been provided by the faculty assessment teams three years earlier.

These results do not prove that these assessments are valid; however, they support the notion that there is a great deal of commonality in the way teams of faculty and graduating students evaluate the relative contributions of academic courses to these three outcomes (knowledge, skills, and intellectual curiosity). I agree with Champagne that there is no comprehensive test of “praxis.” However, from my experience, I would expect that conscientious efforts by educators over an extended period of time will identify many measurable attributes associated with this illusive construct. Furthermore, I would argue that efforts to find ways to measure “praxis” will enrich our understanding of educational processes and allow us to develop more effective policies and programs to enhance its attainment.

A current study using a very different methodology provides further evidence that outcomes important to liberal arts educators can be measured. Charles Blaich and Kathy Wise of Wabash College are coordinating the efforts of nineteen liberal arts colleges from across the country to study the effects of liberal arts education.²² All these schools are using the Defining Issues Test “to gauge student’s moral reasoning” along with the CAAP Critical Thinking Test to measure the effects of liberal arts education. On the one hand, the increases in scores on these two instruments during students’ first year of college are disappointingly small. However, before assessment’s detractors rush to label this as yet another indication of the inadequacy of all such standardized measures, it is important to look at the results more closely. The correlations between students’ particular classroom experiences and increases in their scores on these measures were found to be quite strong. It appears as though increases in student scores

(that is, student learning) and certain pedagogies are closely related. Two of the six sets of classroom attributes examined in the study showed consistently strong positive correlations with both moral development and critical thinking. The first of these, “Good Teaching and High-Quality Interactions with Faculty,” included student responses to twenty-three statements about their perception of teacher engagement, the timeliness of feedback, and teacher clarity and organization. The second factor, “Academic Challenge,” included student perceptions of course difficulty, complexity, and academic standards. Together, these components seem to converge with psychologist Wilbert McKeachie’s classic emphasis on enhancing student learning by balancing the levels of classroom challenge and support.²³

The literature provides extensive documentation of efforts by many different educational researchers to measure higher order thinking and appropriately enlightened discomfort characteristic of what William Perry labeled “contextual relativism.”²⁴ None of these studies claim to have measured these outcomes with great precision. However, the fact that there is substantial convergence within these measures and also the repeated identification of observable associations with widely accepted “best” educational practices, suggests that contrary to Champagne’s claim, such outcomes can be measured by a variety of tools, tests, and instruments. There is reason to believe that outcomes assessment can be done—learning is a natural process, not a supernatural or magical one. The question remains, however: Under what conditions and circumstances is assessment likely to contribute to rather than detract from the goal of enlightening and inspiring students?

Assessment Is an Essential Organizational Function

Just as learning is likely to be greatest in healthy individuals, effective educational assessment occurs most easily and naturally within healthy organizations. From my experiences in consulting with a variety of organizations in several industries, I’ve found the general conceptual structure of the learning organization archetype provides a useful framework for organizational assessment and development.²⁵ Learning organizations are ones where the

relationships among members are simultaneously differentiated and integrated. Although each individual finds ways to make unique contributions, all share a commitment to realizing a common vision of the future. The five disciplines which characterize learning organizations are: personal mastery, mental models, shared vision, team learning, and systems thinking. Although the names of these five disciplines may sound like mere management jargon to some, I have found that once faculty and staff are given the opportunity to engage these concepts, they discover a great deal of relevance to their work as educators. This is not the only model that promotes organizational health and allows for increased effectiveness, but it is a set of ideas and principles I have seen work at many different levels (for example, in the classroom, department, division, college, and even for aggregate councils and committees).

In such an organization, leadership is broadly shared; administrators listen and are influenced by the aspirations as well as the concerns of students, faculty, and staff. Trust and knowledge increase steadily as everyone comes to accept that leadership's fundamental obligation is to create opportunities for everyone to take pride in the work they do.²⁶ In a chapter in Gary Kramer's *Fostering Student Success in the Campus Community*, colleagues and I describe the ways in which progressive leaders at the Air Force Academy and Western Governors University were able to create this kind of organization.²⁷ I would also nominate Bob Maxson, a former president of California State University, Long Beach, as an executive administrator who seemed to epitomize the kind of leadership that helped to transform his institution into a learning organization. As one senior faculty member confided in me, "If a motion of no confidence in this president was ever brought before our faculty, I don't think it would receive a single vote."

I recognize that the learning organizations described in the preceding paragraph and the "corporate university" Champagne rails against in his essay appear to be mutually exclusive entities. Similarly, Powell's criticisms of outcomes assessment assume that all assessment is externally generated and hierarchically imposed. Both of these actions are incompatible with the philosophy and principles of learning organizations. Despite the inherent incompatibility of

these two models, aspects of each archetype are likely to be found in every institution. Perhaps the challenge is not to decide which of these two categories best fits one's own institution, but to interrogate the ways in which our collectives (at every organizational level from the classroom to the boardroom) manifest the characteristics of these two extremes and how these might be reconciled in favor of increased understanding and trust.

In my opinion, it all begins with values. Twenty years ago, as a relatively junior faculty member, I had the unique opportunity to facilitate the Air Force Academy's senior leaders' work together to formally articulate the values that would serve as the foundation for this institution. There were biweekly meetings over several months and lots of input from focus groups and constituent surveys. There was no doubt that each of these general officers was sincerely committed to the project. In the end, this is what they agreed upon: "Integrity first; Service before self; Excellence in all we do." These values were accepted and acknowledged by all the Air Force Academy's "stakeholders" and later adopted by Secretary of the Air Force, Dr. Sheila Widnall, as the core values of the United States Air Force. Of all the good things I received from my thirty-four years in the military, these values are among the most precious. I have also found them to be among the most useful in my work with assessment in a variety of roles over the past two decades.

While such values are necessary for a learning organization to develop, values alone are not sufficient. Two other aspects of a learning organization are also essential: understanding and trust.²⁸ Organizations need to know what they are doing; they need to understand the functional interdependencies involved. As theorist Michel Foucault so famously quipped, "People know what they do; frequently they know why they do what they do; but what they don't know is what what they do does." Work done by organizations to clarify the organization's mission and vision are critically important. In order to be effective, mission statements need to be succinct and convey a clear sense of institutional priorities. Mission statements composed of a multitude of commitments which sprawl over several pages, promising that the institution will be all things to all constituencies, are of little practical value.

Even if those within an organization share a vision of the future, they also will need to share an understanding of the path by which this vision is to be realized. (Outcomes alone are insufficient to establish organizational identity; it is often the paths that organizations choose that most distinguish them from one another.) All organizations have action (or process) models, even if they are not explicitly articulated or publicly embraced.²⁹ Organizations willing to do the hard work necessary to develop a clear and unequivocal statement of their mission, and also the manner and methods by which it will be accomplished, have a distinct advantage over organizations that do not.

As the mission is clarified, the methods and mechanisms for achieving the mission must be systematically tested and evaluated. The question, “How well is the institution doing?” is one of evaluation and becomes important to everyone in a learning organization since the organizational vision and mission reflect their own values and aspirations. However, many of the examples and issues cited by Powell and Champagne suggest that the essential foundation provided by creating a shared, inclusive, and common vision was never established. Under such conditions, the fact that senior faculty members balk at administrative initiatives and bridle at administrative requirements is not surprising.

Assessment also allows organizations to address another and perhaps even more critical question, “How is the organization doing well?” By paying attention to processes (and not just outcomes), assessment provides vital diagnostic information that creates a shared awareness of opportunities for improvement. Together the evaluative and diagnostic information provided by comprehensive assessment allows for continuous refinement of the mental models that underlie organizational activities. Learning organizations learn by making their implicit theories of process and operation or action explicit and then testing them—over and over and over again. In learning organizations, many people accept the responsibility for ongoing efforts to make all processes more transparent. In such organizations, accountability is not a bludgeon used to terrorize faculty, staff, or students, but a value modeled by those in leadership positions and sustained throughout the organization.

Understanding the essential organizational processes is necessary but not sufficient to sustain a learning organization. Learning organizations also continuously increase the level of trust—trust both within and between organizational constituencies.³⁰ Such trust can only develop if all parties are committed to being trustworthy. “Integrity first” requires that what is proclaimed publicly be consistent with what is whispered privately in the multitude of micro-meetings that characterize academic administration. Even impeccable integrity and complete trustworthiness, however, are not enough to ensure trust grows within an organization. Those in positions of leadership (in the classroom or in the executive suites) must be willing to convey trust in others. Micromanagement and manipulation are simply not effective long-term strategies for influencing faculty members. Faculty and staff, as well as students, know when they are trusted and respected and when they are not. The failure to convey trust in others will impede the development of trust in the organization regardless of the integrity or good intentions of those in charge. Without trust, the effectiveness of all internal functions, including assessment, are likely to be limited.

Institutional effectiveness depends on values; without such enlightened and inclusive educational values, and leaders willing to take them seriously, organizations are likely to founder. Many of the incidents recounted in the essays by Champagne and Powell reflect the consequences of institutions eager to use assessment to create appearances that are far better than organizational and operational reality would justify.

My transition from the military to become the provost of a small private liberal arts college was difficult. I had seen the positive impact of academic program assessment at the Air Force Academy. (I could not help but wonder if the positive impression of assessment described by Gerald Graff, then president of the Modern Language Association, in 2008 had been influenced by his visits to the Air Force Academy.³¹) My new college had not done well during its previous accreditation visit, especially in the area of assessment, and seemed eager to improve. I was eager to share broadly accepted assessment principles and programs with my new colleagues. However, a lack of understanding of general educational processes among administrators and

faculty due to relative isolation was exacerbated by a lack of trust throughout the organization. I was amazed that an organization could simultaneously be so hierarchically structured with regard to organizational and personnel decisions and yet so laissez-faire and lackadaisical when it came to issues of student learning. In my thirty-four years in the Air Force, I had not encountered anyone as autocratic as the college president for whom I worked. The effect of his intrusive managerial style on assessment was ubiquitous and chilling.

Berea College is a unique institution in several ways—it was the first college south of the Mason-Dixon Line to educate black and white, men and women together. It continues to offer full-tuition scholarships to all its students; most come from the southern Appalachian region. All students have significant economic disadvantages (basically, Pell Grant eligibility is a requirement for admission), but only those whose records show evidence of strong academic potential are selected. Every Berea College student is required to work at least ten hours per week for the college in exchange for a scholarship that pays the entire \$22,000 annual cost of education and also provides support for room and board expenses when needed. African American student enrollment is nearly 20 percent and international student enrollment is just below 10 percent. Despite the difficulties of ambient low trust across campus, program assessment data were used to increase the focus on support for faculty development and student success.³² Within two years, graduation rates increased to 60 percent, a level that had been attained only twice in the school's 150-year history. The current 60 percent graduation rate is just a little above average nationally. However, when one considers that base rate graduation for students similar in age and income to those who attend Berea College is below 10 percent, this graduation rate represents a very significant achievement. The institution's next accreditation visit was also very successful due, at least in part, to the apparent strong commitment the college had made to assessment. By mutual agreement, I stepped down as provost and returned to full-time teaching after four long years.

Unfortunately, several autocratic personnel decisions coupled with the financial exigencies associated with the institutional response to the recent "great recession" undermined nascent

growth in organizational trust. The comprehensive assessment program that had been developed to impress the accreditation team was abandoned so abruptly that even our students noticed.³³ Current administrative priorities appear to be aimed at increasing apparent institutional effectiveness and efficiency by buying out senior tenured faculty with lucrative early retirement offers, dissolving existing academic departments, and rearranging the remaining faculty into new academic divisions. Despite the turmoil and commotion, educators in many programs are still using classroom and program assessment to make incremental improvement in teaching and learning, and graduation rates above 60 percent have been sustained now for nearly a decade.

I know that the somewhat authoritarian executive leadership style at Berea College is not unique. Most universities and colleges appear to have become more corporate in recent years—the public has come to see higher education as a private gain rather than a public good. Traditional emphases on mission, integrity, and leadership have been supplanted by acquiescence to the influence of money, marketing, and management. Government funding continues to decline steadily. Corporations are where the money is; thus, higher education institutions' increasing need for funds has created pressure for them to take on increasingly corporate characteristics and seek to fill executive positions with those familiar with corporate perspectives and values. Dealing with the corporate world is likely to become increasingly necessary. For most, this is not a prerogative; it is imperative. Is it possible that assessment itself might provide a way in which faculty, staff, and students could exert greater influence on the internal processes of emerging corporate universities?

The Importance of AAUP Involvement in Assessment Activities at All Levels

Over the last five years, I've taught a course that introduces students to the behavioral sciences. We cover the scientific method and evolution by natural selection in some detail before delving into distinctive aspects of anthropology, sociology, and psychology. While each of these social sciences contains many examples of theories and theorists who adhere closely to the scientific

method, students learn that there are also those within each discipline whose ideas and activities are at considerable conceptual distance from the rigors imposed by the scientific method. Among anthropologists, the postmodernists constitute an obvious example of those who deny even the possibility of objective observation which lies at the heart of all empirical science.³⁴ Similarly, psychoanalyst Sigmund Freud's provocative speculations about the psychosexual development of distinctive metaphorical components of personality are viewed with considerable skepticism by most of us currently engaged in psychological science. Karl Marx and others who adopted partisan perspectives before searching for evidence to support their particular theories also belong to the not-so-scientific brand of social scientists. Thus, I would question Champagne's understanding of science when he makes the casual assertion that "Marxism [is] a science and not an ideology."³⁵ This assertion also reveals other, more general, difficulties with the critiques of assessment provided by Champagne and Powell.

The conflict perspective represents an important aspect of contemporary sociology, but it is recognized by many sociologists as not being particularly scientific.³⁶ Absolute adherence to any theory or particular perspective (such as extreme partisanship) can lead to a variety of insights and revelations. Such an approach, however, lacking the potential of falsifiability, cannot claim to be scientific.³⁷ To some extent, the criticisms of assessment that Champagne and Powell express reflect an apparent animosity toward science itself or at least the application of the scientific method to the social milieu. Many of the examples they cite resonate with negative stereotypes that some faculty members in the humanities hold about disciplines such as science, social science, engineering, and management (particularly the emphasis on the mindless manipulation of numbers devoid of inherent value or significance). Helping faculty members to get over the negative stereotypes they hold of one another (as well as of the administration) is one of the greatest challenges to developing the trust necessary to create learning organizations (and thus appropriate contexts for the development of effective assessment programs).

Examining assessment as a labor issue (that is, from a conflict perspective), Champagne struggles with the appropriate way to classify faculty members. Are we bourgeoisie or

proletariat? As scholar Alex Callinicos points out, faculty work contains many characteristics of both categories.³⁸ Despite our relatively low position in the institutional hierarchy, faculty members have a great deal of autonomy and often considerable time to pursue their own research interests and activities without much direct supervision. However, as Champagne hastens to point out, outcomes assessment and the hierarchical imposition of its findings will inevitably reduce faculty autonomy. Thus, he concludes that college professors really are members of the proletariat. While this conclusion might garner general support within academia, a view from the outside is unequivocally to the contrary. A chapter entitled “Lunchpails and Laptops” in Joe Bageant’s *Rainbow Pie: A Redneck Memoir* provides a scathing critique of the role academic institutions and the professoriate, as members of an intellectually elite “overclass,” have played in oppressing America’s true proletariat over much of the last century. As I am about to argue, the claim that we, as professors, are members of the proletariat is disingenuous as well as inaccurate. As Bageant points out, only members of the bourgeoisie would ever consider taking a shower before going to work rather than having to take one when they returned home from work.³⁹

There is also an implicit assertion lurking within the criticisms of assessment offered by Champagne and Powell. They both imply that since higher education has been doing just fine without assessment, it is unnecessary. Whether the declines in higher education over the past few decades have been caused by the assessment movement’s negative effects or have occurred in spite of its many positive contributions, it is difficult to argue that there have not been dramatic declines. A recent Pew Research Center survey of the general public found that “a majority of Americans (57%) say the higher education system fails to provide students with good value for money An even larger majority—75%—say college is too expensive... .”⁴⁰

Richard Arum and Josipa Roksa’s provocative book *Academically Adrift* provides a broad view of just how dire the higher education situation has become. Nearly all faculty members claim that developing critical thinking skills is an “essential” or “very important” goal for undergraduate education. However, Arum and Roksa’s studies find that nearly half of all

college students show no significant gains in this ability during their first two years of college, and over a third show no measurable gains even after four years of college.⁴¹ The Wabash National Study found that students' academic motivation and interest in academic subject matter declined significantly during their first year in college. In 1961, full-time undergraduate students reported studying twenty-five hours per week; in 2003, full-time undergraduates reported studying only thirteen hours per week. Over a third of today's full-time undergraduate students reported spending less than five hours per week preparing for classes. When asked about the level of faculty expectations, one-third of students enrolled in four-year baccalaureate programs reported that none of the courses they took in the previous semester required them to read more than forty pages per week and half reported that no course taken during the previous semester had required them to write twenty or more pages during the entire semester.⁴² Arum and Roksa's study also found evidence that students who did take courses that required at least forty pages of reading per week or more than twenty pages of writing per semester showed significantly greater gains in critical thinking ability.⁴³ This seems to be an interesting example of how assessment itself can have both summative and formative value. While these data suggest that we university professors collectively are not doing well, the results also indicate that providing greater challenge and more effective support can lead to improved learning outcomes for our students.

While Powell and Champagne imply that the corporate university is the prevailing model in higher education today, it is worth examining this assertion. Education, as an industry, exhibits several characteristics that limit the effectiveness of traditional hierarchical corporate models. The education process is complicated; developing and administering educational programs effectively (including teaching courses that transform student thought and perception) is a very complex process. It is also a process that requires regular adaptive expert involvement. The naïve notion of education through osmosis that seems to underlie some conservative critiques of higher education is simply untenable. Similarly, if the faculty is committed to doing only the minimum in each course (that is, just more than that for which they might be punished), the

likelihood of enhancing student learning or achieving educational success is very small. If the university seeks regional accreditation, it needs to have the active support of many of its faculty members. If it seeks to actually educate students, it requires even greater faculty support. Unlike many industries, much of higher education has been tied to the place where it occurs; this has limited the corporate university's opportunity for outsourcing. Similarly, the general practice of academic tenure and notion of shared governance restrict administrative prerogatives to a much greater extent than is true of other industries dominated by corporations. One might conclude that institutions of higher education fit even less comfortably under the label "corporation" than college professors fit the label "proletariat." Nonetheless, there are obviously those who are working assiduously to overcome these "obstacles" and increase the commodification of higher learning and the corporatization of our colleges.

Let's now return to the alternative organizational model provided by learning organizations. Unlike corporations, learning organizations are not driven exclusively by the profit motive; they tend to take a much broader view of the ways in which they can create opportunities for individuals to take pride in the work they do. Perhaps this model provides a chance to renegotiate, and even reconcile, some of the inherent contradictions between the corporate model and educational processes. Author of *The Fifth Discipline*, Peter Senge suggests that a discipline is a practice that builds capacity over time.⁴⁴ To the extent that assessment is done well, it is an organizational practice that can garner the evidence institutions need for adaptation and survival. Arguably, assessment is the best way to test implicit theories of how educational systems work and identify opportunities for improvements. In contrast to assessment's potential benefits, Powell and Champagne describe situations where assessment programs became impediments to effective organizational functions and processes. Rather than building community, the hierarchical imposition of externally generated, but ill-conceived, assessment programs and inadequate protocols fractured academic communities. These initiatives also distracted and frustrated educators throughout the organization. Most of these

problems, however, appear to be a consequence of the misuse of assessment principles and processes rather than deficiencies in assessment itself.

Doing assessment right requires that the appropriate goals must be identified and then measures must be developed and refined over time to capture relevant data. Identifying the institutional priorities is a leadership function: it necessarily involves all constituencies in an interactive conversation about those outcomes the institution values most highly. Appropriate goals cannot be externally imposed nor should they be post hoc creations that are generated from the adoption of particular assessment instruments or tools. Instruments must be locally developed, tested, and refined to serve institutional priorities. As the late total quality guru, W. Edwards Deming was so fond of saying about standardized approaches to quality, "This is not instant pudding!" All too often, administrators, eager to impress accreditation agencies or governing boards, identify popular instruments or methodologies, then try to bend the institutional priorities to fit the particular assessment technology. Similarly, the lack of adequate controls on the administration and collection of data together with superficial statistical and qualitative analyses are creating piles of data across the country which have little real relevance or value to educational organizations.

I was curious as to what the American Association of University Professors' position on assessment might be. On the one hand, publication of two articles critical of assessment might indicate the perception by the Association of a fundamental incompatibility. On the other hand, the invitation (or challenge) for an essay from "the other side" suggested that there might be some openness to debate. Just recently, Gary Rhoades, general secretary of the AAUP from 2009 to 2011, coauthored an occasional paper for the National Institute for Learning Outcomes Assessment entitled "What Faculty Unions Say about Student Learning Outcomes Assessment." I found his articulation of AAUP policy to be surprisingly close to the perspective that I've tried to develop in this essay:

To some observers as well as some faculty, the AAUP's principles and policies might suggest that the association encourages its members to resist the assessment of student

learning outcomes, including acting on that data to reform curriculum and instruction.

That is a fundamental misreading and a misapplication of the association's basic principles and policies Of principal interest to the AAUP is the *process* by which assessment metrics are developed and applied and the process by which the findings of those assessments are translated into instructional and curricular reform.

Assessment of student learning and reform of teaching and academic programs are core academic activities. As such, the AAUP sees them as being the primary responsibility of faculty—individually and collectively.⁴⁵

Academic freedom appears to be a particularly difficult issue embedded within the debate about assessment. I suspect this issue remains an unnamed motivation by some outside the educational establishment who are pushing for greater outcomes assessment as well as some within academia who are resisting most frantically. The argument that learning cannot be assessed should be abandoned; there is too much evidence to the contrary. However, the particular assessment results one observes are likely to depend on the protocols and instruments employed. Careful, deliberate, and gradual synthesis is the only appropriate way to “close the loop” and apply the results of assessment to improving programs and policies. It is of vital importance that faculty members be included and engaged in the process all along the way. Powell makes an excellent point in one of his essay's later sections, “Neglected (Because Unknown) Outcomes: What Gets Left Out.”⁴⁶ Left on their own, those outside the educational process, and this could include many administrators as well as politicians and trustees, lack the knowledge and understanding necessary to develop meaningful educational assessment programs. The success of the comprehensive assessment program at the Air Force Academy was contingent on the amount of latitude given to the faculty committees to collect and interpret evidence they agreed was relevant to each of the institution's learning outcomes.

As important as academic freedom is, it is not absolute. The individual faculty member's freedom must be considered in the context of the faculty's collective goals and objectives. Once again, this is a position the AAUP seems to endorse:

In the classroom, a core element of academic freedom is the autonomy of the individual faculty member to determine what and how to teach. At the same time, AAUP emphasizes the collective responsibility of faculty as a whole for academic programs, suggesting that an academic department, for instance, can adopt pedagogical or curricular standards that colleagues teaching the courses need to adopt.⁴⁷

Academic freedom should not include freedom from the consequences of one's choices. This is true of executive administrators as well as faculty, staff, and students. Robust and effective assessment (which includes inputs and processes as well as outcomes) is the best way to evaluate the overall quality of our programs and improve them. In fact, the AAUP's 1968 statement *The Role of Faculty in the Accrediting of Colleges and Universities* not only acquiesces to the need for accreditation but asserts that periodic accreditation visits should include assessment of issues and policies relating to academic freedom, tenure, shared governance, faculty working conditions and morale.⁴⁸ I suspect that although the influence of these issues on student learning may be indirect, comprehensive assessment may identify numerous causal connections. The evidence of these connections is what is needed to persuade administrators and external constituencies to develop policies and programs that will actually contribute to increases in student learning by including and engaging faculty members more fully.

The AAUP has also wrestled with the tensions between assessment for improvement and assessment for accountability. Rhoades points out that if outcomes are defined too narrowly, assessment results can lead to policies that undermine student learning. For example the push for increases in faculty "productivity" as measured by insufficient assessment indicators might lead to reductions in standards or increases in recruiting selectivity. Obviously, neither of these actions is likely to have a positive effect on student learning. Rhoades also suggests that the current emphasis on increasing faculty accountability is being translated into "attempts to educate more students with fewer full-time faculty, increasing class size, and decreasing the amount of time faculty have to be available to students."⁴⁹

For a disturbing example of the dangers of the kind of minimalist analytics that might accompany ideologically driven measures of “productivity” data, please consider the Center for College Affordability and Productivity’s *Faculty Productivity and Costs at the University of Texas at Austin*. In this report, Richard Vedder and his coauthors speculate hungrily on the enormous economic savings to the system if every faculty member was as “productive” as those who were in the top quintile of productivity at the University of Texas. This analysis completely neglects any consideration of the impact on the quality of student learning if all faculty members were to be assigned responsibility for teaching 896 student credit hours or 318 students per year. While the word “teaching” occurs eighty-seven times in this seventeen-page report, unfortunately, the word “learning” cannot be found.⁵⁰ As Ted Marchese, former editor of the American Association for Higher Education’s journal *Change*, often pointed out, “Teaching without learning is just talk.” Clearly, assessment activities require the inputs of reasonable and responsible faculty members.

One final area of concern is that an administration’s emphasis on immediate improvements may seriously distort the assessment process. Once again, the gradual synthesis model is to be preferred to shortsighted preliminary analyses or knee-jerk reactions to patterns of evidence that appear to be contrary to closely held aspirations. In the short term, investments in things like faculty development and service learning may appear to cost more than their immediate benefits would justify. However, with time and continuous, collaborative engagement, such investments have the potential to yield substantial gains. As Rhoades concludes:

It takes investment to substantially enhance the yield in student learning outcomes. That means concentrating attention on and tracking patterns in personnel and other expenditures that are designed to stimulate greater learning. It means investing in enhancing student learning outcomes, not just assessing them. To do otherwise, to have assessment without investment, is to have academic equivalent of an unfunded mandate.⁵¹

On the other hand, it seems equally inappropriate in the present perilous circumstances to invest resources anywhere in higher education without the development of a robust and inclusive program of educational assessment to guide and inform these choices. Ultimately, this is a collective professional responsibility of all faculty members. It is time to get over our fears, and get on with our work.

Dave Porter is professor of psychology at Berea College.

Notes

¹ John Powell, "Outcomes Assessment: Conceptual and Other Problems," *AAUP Journal of Academic Freedom* 2 (2011): 21, <http://www.academicfreedomjournal.org/VolumeTwo/Powell.pdf>.

² John Champagne, "Teaching in the Corporate University: Assessment as a Labor Issue," *AAUP Journal of Academic Freedom* 2 (2011): 24, <http://www.academicfreedomjournal.org/VolumeTwo/Champagne.pdf>.

³ B. F. Skinner, *About Behaviorism* (New York: Vintage Books, 1976).

⁴ Philip Johnson-Laird, *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness* (Cambridge: Cambridge University Press, 1983).

⁵ Richard Hughes, Robert Ginnett, and Gordon Curphy, *Leadership: Enhancing the Lessons of Experience*, 2d ed. (Chicago: Irwin, 1996).

⁶ Jean Piaget, *The Language and Thought of the Child* (New York: Meridian, 1927).

⁷ For moral development, see Lawrence Kohlberg, "The Development of Children's Orientations toward Moral Order: 1. Sequence in the Development of Moral Thought," *Human Development* 6, no. 11 (1963): 11–33. For psychosocial development, see Erik H. Erikson, *Identity: Youth and Crisis* (New York: W.W. Norton, 1968). For spiritual development, see James W. Fowler, *Stages of Faith Development: The Psychology of Human Development and the Quest for Meaning* (San Francisco: HarperCollins, 1981).

⁸ William G. Perry, *Forms of Intellectual and Ethical Development in the College Years: A Scheme* (Troy, MO: Holt, Rinehart, & Winston, 1970).

⁹ Mary Belenky, Blythe Clinchy, Nancy Goldberger, and Jill Tarule, *Women's Ways of Knowing: The Development of Self, Voice, and Mind*, 10th anniversary ed. (New York: Basic Books, 1997).

¹⁰ Champagne, "Teaching in the Corporate University," 2.

¹¹ Powell, "Outcomes Assessment," 3.

¹² Champagne, "Teaching in the Corporate University," 10–11.

¹³ Champagne, "Teaching in the Corporate University."

¹⁴ Barbara E. Walvoord and Virginia Johnson Anderson, *Effective Grading: A Tool for Learning and Assessment* (San Francisco: Jossey-Bass, 1998).

- ¹⁵ Thomas A. Angelo and K. Patricia Cross, *Classroom Assessment Techniques: A Handbook for College Teachers*, 2d ed. (San Francisco: Jossey-Bass, 1993).
- ¹⁶ Powell, "Outcomes Assessment."
- ¹⁷ John M. Darley, "The Dynamics of Authority in Organizations and the Unintended Action Consequences," in *Social Influences on Ethical Behavior in Organization*, ed. Darley, D. M. Messick, and T. R. Tyler (Mahwah, NJ: L.A. Erlbaum, 2001), 37–52.
- ¹⁸ Steven D. Levitt and Stephen J. Dubner, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (New York: HarperCollins, 2005).
- ¹⁹ Champagne, "Teaching in the Corporate University," 2.
- ²⁰ Donald E. Broadbent, *In Defense of Empirical Psychology* (London: Methuen, 1973), 207.
- ²¹ David B. Porter and S. M. Eisenhut, "An Integrated Approach to Educational Outcomes Assessment," *Journal of Adult Assessment* 6, no. 2 (1996): 1–12; David B. Porter, "Educational Outcomes Assessment: The Good, the Bad, and the Ugly," *Adult Assessment Forum: Journal of Quality Management in Adult-Centered Education* 8, no. 3 (1998): 1–14.
- ²² Charles Blaich and Kathy Wise, "Overview of Findings from the First Year of the Wabash National Study of Liberal Arts Education" (2011), http://www.liberalarts.wabash.edu/storage/Overview_of_Findings_from_the_First_Year_web_07.17.09.pdf.
- ²³ Wilbert J. McKeachie, *Teaching Tips: Strategies, Research, and Theory for College and University Teachers*, 9th ed. (Lexington, MA: DC Heath, 1994).
- ²⁴ Perry, *Forms of Intellectual and Ethical Development*.
- ²⁵ Peter M. Senge, *The Fifth Discipline* (New York: Doubleday, 1990).
- ²⁶ David B. Porter and Deanna G. Sergel, "Institutional Assessment: The Indispensability of Diversity," in *Proceedings: Sixteenth Applied Behavioral Sciences Symposium*, USAF-TR Publication No. 98-1 (Colorado Springs, CO: United States Air Force Academy, 1998), 86–90.
- ²⁷ David B. Porter, Joe Bagnoli, Janice Burdette Blythe, Donald Hudson, and Deanna Sergel, "Organizing Student Services for Learning," in *Fostering Student Success in the Campus Community*, ed. Gary L. Kramer (San Francisco: Jossey-Bass, 2006).
- ²⁸ Porter and Sergel, "Institutional Assessment."
- ²⁹ Chris Argyris and Donald Schön, *Organizational Learning: A Theory of Action Perspective* (Reading, MA: Addison-Wesley, 1978).
- ³⁰ Porter and Sergel, "Institutional Assessment."
- ³¹ Gerald Graff, "Assessment Changes Everything," *MLA Newsletter* (Spring 2008): 3–4.
- ³² Porter et al., "Organizing Student Services for Learning."
- ³³ Megan Rodgers, "A Call for Student Involvement in the Push for Assessment," *Assessment Update* 23, no. 1 (2011): 4–5.
- ³⁴ John Monaghan and Peter Just, *Social & Cultural Anthropology: A Very Short Introduction* (New York: Oxford University Press, 2000); Keith E. Stanovich, *How to Think Straight about Psychology*, 7th ed. (Boston: Allyn & Bacon, Pearson Education, 2004).
- ³⁵ Champagne, "Teaching in the Corporate University," 22.
- ³⁶ Richard T. Schaefer, *Sociology: A Brief Introduction* (New York: McGraw-Hill, 2008); Steve Bruce, *Sociology: A Very Short Introduction* (New York: Oxford University Press, 1999).
- ³⁷ Karl Popper, *Conjectures and Refutations* (London: Routledge and Keegan Paul, 1963), 33–39; Stanovich, *How to Think Straight about Psychology*.

- ³⁸ Alex Callinicos, "The New Middle Class and Socialist Politics," in *The Changing Working Class*, ed. Callinicos and Chris Harman (London: Bookmarks, 1987).
- ³⁹ Joe Bageant, *Rainbow Pie: A Redneck Memoir* (Melbourne, Australia: Scribe Publications, 2010).
- ⁴⁰ Paul Taylor, Kim Parker, Richard Fry, D'Vera Cohn, Wendy Wang, Gabriel Velasco, and Daniel Dockterman, "Is College Worth It? College Presidents, Public Assess, Value, Quality and Mission of Higher Education," *Pew Social and Demographic Trends* (May 15, 2011), 1, <http://pewsocialtrends.org/2011/05/15/is-college-worth-it/>.
- ⁴¹ Richard Arum and Josipa Roksa, *Academically Adrift* (Chicago: University of Chicago Press, 2011).
- ⁴² Blaich and Wise, "Overview of Findings."
- ⁴³ Arum and Roksa, *Academically Adrift*.
- ⁴⁴ Senge, *The Fifth Discipline*.
- ⁴⁵ Larry Gold, Gary Rhoades, Mark Smith, and George Kuh, "What Faculty Unions Say about Student Learning Outcomes Assessment," occasional paper of the National Institute for Learning Outcomes Assessment (May 2011), 7, <http://learningoutcomesassessment.org/documents/Union.pdf>.
- ⁴⁶ Powell, "Outcomes Assessment," 10–13.
- ⁴⁷ Gold et al., "What Faculty Unions Say," 7.
- ⁴⁸ Gold et al., "What Faculty Unions Say."
- ⁴⁹ Gold et al., "What Faculty Unions Say," 11.
- ⁵⁰ Richard Vedder, Christopher Matgouranis, and Jonathan Robe, *Faculty Productivity and Costs at the University of Texas at Austin* (Washington, DC: Center for College Affordability & Productivity, 2011), http://www.centerforcollegeaffordability.org/uploads/Faculty_Productivity_UT-Austin_report.pdf.
- ⁵¹ Gold et al., "What Faculty Unions Say," 12.