

# Is Outcomes Assessment Hurting Higher Education?

by James F. Pontuso and Saranna R. Thornton

Whenever we meet with colleagues from other schools, the conversation soon focuses on one pressing issue: not our research projects, not a promising new pedagogy, not which books inspire students, but *what to do about assessment*. Assessment has become such a critical problem for higher education, especially for teachers of the liberal arts, that we spend our time talking about little else. Could it really have been the intention of the architects of “outcomes assessment” to *distract* educators from their primary responsibilities?

Until recently, undergraduate college assessment primarily evaluated the environment in which students study—asking the question: Do the various elements that constitute the academic “playing field” combine in a way that relatively motivated students can succeed in their programs of study? Accrediting institutions examined colleges’ financial records to ensure that institutions had the resources to provide promised areas of study. Faculty credentials were appraised to determine whether they reflected required subject-area expertise. Libraries, laboratories and other physical facilities were evaluated to ensure adequate support existed for academic programs. Unlike outcomes assessment, the logic behind this sort of assessment of the environment in which students learn is realistic because it recognizes important differences between education and so many other products available for sale in a free-market economy.

For example, the fat, sodium, and calorie content of a twelve-ounce bag of Ruffles potato chips doesn’t change based on who is eating them. The lead con-

---

*James F. Pontuso is Patterson Professor of Government & Foreign Affairs at Hampden-Sydney College in Virginia. He has authored or edited six books and published more than sixty articles, reviews, and essays. In 1993-94 and again in 2003, he was a Fulbright lecturer at Charles University in Prague, the Czech Republic. Saranna R. Thornton is Elliott Professor of Economics at Hampden-Sydney College in Virginia. A nationally recognized expert on how the Family and Medical Leave Act (FMLA) and Pregnancy Discrimination Act impact higher education, she also serves as a consultant to the International Association for Feminist Economics.*

tent of a gallon of gasoline, the wattage of a light bulb, or the voltage of a battery doesn't vary across individual users. But in other cases, product performance varies dramatically based on the combination of provider outputs and user inputs. The high cholesterol of patients who won't follow their doctors' prescribed diet isn't an outcome that reflects the quality of services provided. Getting only 20,000 miles out of brand-new tires isn't an outcome that measures tire quality, if the owner drives with underinflated tires.

The measured "outcomes" of a student who completes minimal amounts of school work and who graduates from college learning little—or a student who

---

*The academic abilities that a student possesses at matriculation will dramatically affect how much that student is able to learn in a four-year course of study.*

doesn't do any work and flunks out before graduation—are not reflective of the quality of education that their colleges offered. Outcomes assessment also fails when students are motivated to study, but are unable to take maximum advantage of their academic environment for personal reasons, such as being single parents, or having to work long hours while attending college. The academic abilities that a student possesses at matriculation will dramatically affect how much that student is able to learn in a four-year course of study. Students who enter college from the bottom 10 percent of their class are capable of learning, but it is unlikely that the body of knowledge they will demonstrate upon graduation will be as great as that of students who enter college from the top ten percent.

To adopt an analogy from college football, consider two quarterbacks who play for teams in the same athletic league. One was born talented. He skips practices, doesn't do pre-season workouts, and drinks too much beer the night before games. Yet his outcomes—pass completion rate, total yards gained passing, touchdowns, etc.—are impressive. The other quarterback doesn't have much natural talent, but is motivated and self-disciplined. He never skips practice, works hard on the field and during pre-season, never breaks training rule, and seeks extra help from the coach; he produces the exact same outcomes as the first quarterback. The phenomenon of observational equivalence makes it impossible for us to judge which coach is better. We cannot infer from the outcomes assessment data alone which football program has coached the better quarterback.

Because of observational equivalence, the metrics of education outcomes assessment—the percentage of students who graduate within four years, the employment rates of alumni post-graduation, standardized tests of knowledge acquired in specific academic disciplines—are inaccurate and, thus, not credible

regarding the performance of individual colleges. Metrics do not exist that would allow individual colleges to assess student learning—while simultaneously controlling for student aptitude, willingness and ability to do college level academic work, and other such variables.

While the common-sense goals of assessment are laudable, the actual consequences of the process are far from beneficial. Ongoing assessment diverts teachers from teaching. Instead of preparing their courses, meeting with students, or grading papers—in short, executing their teaching duties—instructors must

---

*Academic deans, instead of overseeing assessment activities, might be better engaged in useful activities such as developing young faculty or securing grants.*

spend a substantial amount of time worrying about how to assess what they teach. Moreover, academic deans, instead of overseeing assessment activities, might be better engaged in useful activities such as developing young faculty or securing grants. No one, to my knowledge, has done a serious cost-benefit analysis of whether the innumerable hours faculty and administrators expend on outcomes assessment could be better used on activities that directly benefit students. No one knows what opportunities have been lost to the demands of devising and implementing assessment instruments.

The real scandal of outcomes assessment, the one nobody talks about, is that the methods used to assess usually produce very little worthwhile data. Departments and programs create assessment tools through a process that: 1) sets goals for student learning, 2) gathers evidence of whether students have learned what is expected, 3) interprets the information gathered, and 4) adapts teaching methods in light of the evidence. But every social scientist knows that the only valid way to measure human phenomena is with double-blind experiments in which neither those who actually administer the tests nor those who take them know what's being tested.

Of course, this is impossible when assessing college programs. Students know exactly why they are being assessed. Even worse, the faculty members whose programs are being assessed are the very people charged with devising, administering, and appraising assessment instruments. Such a system is easily abused since no one wants to look bad. In almost every case, measurement tools are constructed that simply validate what teachers and administrators are already doing.

But that is not the worst of it. The dirty little secret is that teachers pay almost no attention to assessment outcomes. They learn little from the exercise—considering it only another (usually uncompensated) onerous administrative duty—and

they often dismiss the findings because of the way accrediting agencies structure the activity. Since assessors cannot be experts in every academic field, they require that every department and program aggregate information. The people doing the assessing are not capable of judging the merits of syllabi, tests, and papers from outside their field of study. Thus, they make all departments homogenize the “outcomes” into a form comprehensible to a generic reader. The problem is that students learn chemistry differently than they do a Dostoyevsky novel. Assessment measurements that attempt to aggregate information across disciplines may miss this important difference. In fact, the only people who truly pay attention to assessment results are accrediting agencies. But because all institutions are fearful of losing accreditation, everyone at the institutions under the microscope acts as if assessment matters.

Since the current methods of assessment are little more than a pretense, why not replace them with a useful instrument and process? One reason is the inordinate amount of time required to create valid methods of assessing student learning. Teachers would hardly have time to teach if they were required to undertake true double-blind experiments testing the effectiveness of their teaching.

Inevitably, since everyone is aware that faculty-administered assessment is faulty, the push will come to ascertain student outcomes using nationally standardized tests. This, indeed, has already begun, as reported by Daniel Golden in the *Wall Street Journal*. National tests are constructed by experts in the various subject areas. Why would a group of experts know what and how to teach better than the professor in the classroom? Many people’s lives have been affected by good teachers, but no one’s soul has ever been touched by a committee of test writers. As high



school Advanced Placement exams demonstrate, national tests not only standardize how student achievement is measured, they standardize what students are taught. Professors would inevitably teach to the test. American higher education—arguably the best in the world—would lose the diversity that is a major source of its strength.

The onerous effects of outcomes assessment are felt most heavily in the traditional liberal arts, where the educational goal is not merely to impart a body of knowledge, but to convey a way of thinking. Liberally educated students must be taught to analyze problems, evaluate data, critically appraise arguments and beliefs

---

*In a sense, the object of the liberal arts is to prepare young people to cope with problems and challenges that do not yet exist.*

and, most importantly, weigh alternatives. In a sense, the object of the liberal arts is to prepare young people to cope with problems and challenges that do not yet exist.

The ills created by outcomes assessment are even more striking in the humanities where the aim is not much different than the objective sought by Socrates. As the name implies, professors teaching the liberal arts are not merely attempting to transfer knowledge to students. Instead, they are trying to liberate them from the tyranny of the merely conventional. The problem with teaching young people is not that they do not know enough; the real problem is that they think they know the answers to all of life's most important questions.

The goal is to show people who do not know—but think they do—that they do not know. The best teachers understand that their mission is to prepare students to lead interesting and satisfying lives; equip them to appreciate the comedies and tragedies that will surely occur in their later years; and instill a sense of wonder and longing in their souls for what is true, just, and beautiful. The task of the liberal arts is to make students aware that a good career may not be the one that earns the most money, that beauty is a charming but fragile quality, and that happiness is gained only when the soul has mastery over both joy and sorrow. How does one explain to assessors that the best possible “outcome” is for students to leave college less certain about what they “know” but more curious about what they have yet to learn?

The model of national testing that will be thrust on colleges is similar to the No Child Left Behind Act. To be fair, No Child Left Behind has forced some states and localities to make much-needed changes in a few substandard schools, mostly by holding teachers and principals accountable for student achievement.

No Child Left Behind often uses standard of learning tests (SOLs)—a form of outcomes assessment. But even before the SOLs were administered, it was clear that the schools being tested were not adequately preparing students: graduation rates were low, the number of students entering college was minimal, and SAT performance was below par. No Child Left Behind created a large federal bureaucracy to fix a few failing schools. Educators did not need outcomes assessment to see that these schools were failing; both the malady and the cure were obvious to common sense. For schools that were already fulfilling their educational mission, the SOLs have created nothing but red tape for administrators and useless mem-

---

*Usually, good teachers are too busy teaching to waste time—unless coerced—on what they consider at best a marginally useful endeavor.*

orization for students. Teachers who must prep students for the SOLs have renamed the policy No Child Left Awake.

Another argument from proponents of outcomes assessment is “the accreditors are us.” That is, the accrediting agencies are composed of faculty and staff from the schools they scrutinize. But, in fact, are accrediting committees really us? It might be useful to explore how many teaching awards members of accrediting committees have won. Usually, good teachers are too busy teaching to waste time—unless coerced—on what they consider, at best, a marginally useful endeavor. Cynics have charged that outcomes assessment has spawned a cadre of accrediting “professionals” with a financial interest in making the assessment process difficult, burdensome, and costly. The charge has some merit: some academics who have served on accrediting committees have later parlayed their expertise into consulting fees, selling their services to institutions uncertain about how to pass muster with accrediting agencies.

I doubt that outcomes assessment is a plot by self-interested academics intent on making a fast buck on the latest academic fad; most of the people involved sincerely believe they are helping to ensure college students “get what they pay for.” But God save us from the earnest. As Michael Crozier points out in *The Bureaucratic Phenomenon*, bureaucracies are burdensome not because their members are lazy and inefficient, but because their talented members want to reform the world. In the case of assessment, such people assume more and more responsibility, convinced that they’re improving something—the students’ learning, the professors’ commitment to quality instruction, the institutions’ reputation. In doing so, they give bureaucracy more and more work, its staff greater authority, and teachers more inane paperwork.



Education Secretary Margaret Spellings embodies the thinking of those who, with the best of intentions, oversee the outcomes assessment movement. In 2006 she led a charge “to streamline and improve the accreditation process, to support innovation, to promote consistency in accreditation standards, and to increase accountability.” But how is it possible to standardize educational results and at the same time promote innovation? Under pressure from Congressional leaders, Secretary Spellings later reversed the goal of standardization, noting that: “No one-size-fits-all measures, no standardized tests;” yet, Secretary Spellings has not given up the goal of assessing whether colleges and universities have achieved

---

*The implication is that change is better than the status quo, even if that means changing what has been demonstrated to work well.*

“acceptable levels of institutional success with respect to student achievement.” In other words, she wants to continue and to strengthen outcomes assessment.

The expansion of bureaucratic tasks that Crozier predicted is just what has happened in accreditation. Accrediting bodies not only examine what colleges are doing, some—like that Southern Association of Colleges and Schools (SACS)—now require Quality Enhancement Programs (QEPs) that force colleges to “enhance” their programs in various ways. Most colleges have responded to the QEP requirement by attempting to revise their curriculum. QEPs have become an unfunded mandate to engage in curriculum reform and—of course—to assess the results of those reforms. In fact, the SACS requirement comes close to stating that improving the quality of learning is the same as assessment.

SACS’ “Principles of Accreditation” states: “Accreditation requires institutional commitment to the concept of quality enhancement through continuous assessment and improvement.” To put it bluntly, the argument goes, even if students are really learning, colleges can do better, and they can do better if they assess to find out how to do better. The implication is that change is better than the status quo, even if that means changing what has been demonstrated to work well. Even when colleges decide not to revise their curriculum, faculty must waste valuable time debating the issue.

No serious study has ever been undertaken to determine whether outcomes assessment actually improves education. While some evidence seems to imply that assessment improves what assessors are looking for, such studies use the information constructed on the same pseudo, made-up data that accreditors demand in the first place. A more interesting and useful study would measure whether colleges that are not required to undertake outcomes assessment are as successful teaching

students as colleges that are.

There is a kind of assessment that teachers have been doing for years; it is called grading. If students learn, they get good grades; if they don't, they get poor ones. As Harvard's Harvey Mansfield has pointed out, rampant grade-inflation has undermined confidence that grades actually mean something. The outcomes assessment craze is a direct consequence of grade inflation.

The traditional method of grading may not be a perfect way to discover whether teachers have touched students' souls, but it has the virtue of forcing students to undertake the kind of discipline and confrontation with great ideas that

---

*Student evaluations are helpful to teachers in finding out if the way they are presenting the materials to students is working.*


may liberate their minds. At least grading takes into account the specific subject matter of the course. Most liberal arts teachers grade students not merely on their ability to regurgitate information, but also on their capacity to master the materials through term papers, exams, and class discussion.

Teachers also have been taking student evaluations of their courses for generations. Student evaluations are helpful to teachers in finding out if the ways they are presenting the materials to students is working. Moreover, independent of accrediting agencies and outcomes assessment, most colleges currently ask professionals from other schools to evaluate specific disciplines and programs—economists review economists, biologists review biology—and because the reviewers are knowledgeable in the field, a look from an eye outside can often be revealing and useful.

There is a very simple way for accrediting agencies to assess whether students are learning. Rather than reinventing the wheel by using slick language (such as “student outcomes”) that doesn't mean anything, they should look at randomly selected course syllabi, reading assignments, term papers, tests, and student evaluations. It would be more work for assessors, since they would have to learn something about fields of study not their own, but it is a more effective way to judge whether students are learning. There may be an added benefit to this method: perhaps evaluators will discourage the grade inflation that has become the other scandalous bane of higher education.

Teachers assess all the time. They read student papers and exams to discover if students have learned. They ask questions in class and engage students in discussion. They look over student evaluations to see if the way they are interacting with students is being well-received. They are always trying to find better ways to



help students grasp the material. Why do they need to spend time in another elaborate and meaningless type of assessment? For good teachers, outcomes assessment is mostly a distraction; for bad ones, it provides a bureaucratic cover to validate what they are doing. 

## WORKS CITED

- Basken, Paul. (2008, January 11). "Colleges and Their Accreditors Escape Tougher Scrutiny, for Now," *Chronicle of Higher Education*, January 11, 2008. <http://chronicle.com/free/v54/i18/18a02002.htm>.
- Crozier, Michael. *The Bureaucratic Phenomenon*. Chicago: University of Chicago Press, 1967.
- Golden, Daniel. "Colleges, accreditors seek better ways to measure learning." *Wall Street Journal*, November 13, 2006. [www.post-gazette.com/pg/06317/738027-298.stm](http://www.post-gazette.com/pg/06317/738027-298.stm).
- Southern Association of Colleges and Schools. "Revisions to the principles of accreditation: foundations for quality enhancement." [www.sacscoc.org/pdf/Proposed%20Principles.pdf](http://www.sacscoc.org/pdf/Proposed%20Principles.pdf).
- U.S. Department of Education. Press Release. November 29, 2006. [www.ed.gov/news/pressreleases/2006/11/11292006.html](http://www.ed.gov/news/pressreleases/2006/11/11292006.html).

A different version of this article, entitled *The Assessment Waltz: How Outcomes Assessment is Hurting Liberal Arts Colleges*, by James Pontuso appeared in the online journal *First Principles* ([www.firstprinciplesjournal.com/articles.aspx?article=578&doc=qs](http://www.firstprinciplesjournal.com/articles.aspx?article=578&doc=qs)) April 9, 2008.

