

Outcomes Assessment: Conceptual and Other Problems

By John W. Powell

Abstract: Despite their plausibility and their having become common practice, programs of outcomes assessment are fraught with problems. Disclosing the problems requires critical thought and a skeptical attitude at odds with most of the relevant literature. The practice lacks a basis in research, though circular, question-begging faux research attempts to caulk the gaps. Its real basis, discernible on investigation, raises more problems. There is a further problem of map versus territory, in that accountability encroaches on, and becomes confused with, that for which we are accounting—practitioners confuse outcomes with education, with the result that education becomes routinized and mechanical to-do lists. In addition to displacing teaching and learning, outcomes assessment divides the faculty, contributes to a misrepresentation of higher education as more narrow and simple than it in fact is, encroaches on faculty decision making and the ability of professors to claim curriculum and the means of its delivery as the purview of faculty members, and distracts us from more important and more urgent needs for action to improve education.

Outcomes Assessment: Conceptual Problems

The basic argument for doing outcomes assessment (OA) in higher education is almost Amish in its plainness. That plainness helps garner assent even from those who do not understand outcomes assessment or like it. The argument also influences what counts as outcomes assessment. The argument's plausibility displaces other arguments as superfluous—for example, those grounded in controlled-studies-based research—or, indeed, any kind of research.

Here is the argument: *Higher education uses the money of other people, people who are neither students nor staff nor administrators nor teachers. Those who pony up the money deserve evidence they can understand that their money is being used wisely. Therefore higher education should include structures that track measurable differences produced in students who get degrees, and it should then report the results.*

This argument has been bought by legislators, boards of trustees, US secretaries of education, and accrediting agencies. It has been accepted by university presidents. Programs of outcomes assessments are a result. In those programs, each university, each college within those universities, each department, each course, and each professor teaching that course are required to articulate measurable outcomes showing the success of these courses, departments, colleges, and so on, as well as the changes produced in students. This evidence is then gathered and reported. Enormous amounts of work are going into efforts to implement and follow up on these programs at universities across the country.

This argument provides a good test case for students to review critical thinking skills of clarifying issues and describing arguments. In fact, I can assess my teaching of these skills by whether students can see whether and how this issue is problematic and whether they can articulate the difficulties. I win some and lose some. If students whether parents, legislators, taxpayers understand grades, transcripts, and letters of reference as appraisals of what students have learned, that's a plus. If they ask whether there's been research on how outcomes

assessment improves education over education without assessment, that's another plus. If they ask how teachers can otherwise improve their teaching, yet another. (One student, an anthropology major, citing William James, argued that paying attention to one's teaching might improve it but that we still need to distinguish different kinds of paying attention before claiming that one [OA] is particularly effective.) If they cite studies claiming that graduates earn an average forty thousand dollars more a year than those who didn't attend college, I wince at the danger of a narrow view of education but still nod. If they ask whether the push toward outcomes assessment is really the result of a grassroots movement, I remind them of dangers of genetic fallacies but encourage them to look at Robert Birnbaum's *Management Fads in Higher Education* (2000). And I tell my students that if they get it completely wrong that's okay because everybody's doing it.

But the argument's problematic assumptions, once exposed, eat it alive. Various critical tools are relevant to assessing outcomes assessment, including an archaeology of its sources, an exploration of its effects, a conceptual analysis of its assumptions, and research into its supporting arguments. Sometimes the first alarm bell, a *duh!* moment, comes with realization that calls for increased accountability in higher education require wisdom and education on the part of those who call and those who would do the appraising. Technicians, managers, administrators, and accreditors who would assess the outcomes of education need themselves to be broadly educated and thoughtful regarding education. That outcomes assessment has become so widely implemented might be thought disheartening, but its prevalence may also make it easier to reevaluate now.

I pass by measurability problems, mostly because I see many faculty members shut down in such discussions. The wars over operationalizing goals into measurable outcomes are strongly reminiscent of fifty-year-old controversies surrounding behaviorism, in which many on each side demonized the other (and possibly still do), then shut down with fingers in their ears. One side apparently thinks those insisting on outcomes assessment would turn all art into calculi and all wisdom into idiot savant checklists, while the other thinks its opponents would insist on reverence toward antiscientific will-o'-the-wisps and mass delusions while insisting on

contempt toward evidence and arguments. These problems could use fresh appraisal, but it will have to be smooth, mediating, and persuasive in ways for which I have no patience, no patience for either side.

In addition to the question of whether OA fetishizes measurement are other issues that are logically prior, more urgent, less sclerotic, and (because they have received less press) more promising of progress. They offer relevant considerations to policymakers deciding whether to implement or to reassess outcomes assessments. Advocates and opponents can work together sorting through pros and cons.

Where Is the Research?

What are the measurable effects of outcomes assessment programs?

Outcomes assessment is undeniably succeeding on its own terms, in that it is becoming widely implemented and is generating data. Also undeniably, outcomes-assessment-based education lacks evidence that it is an improvement over traditional education based on any other terms of evaluation. That is, when OA-based programs are compared with other programs lacking OA, one would expect to see some kind of results. Instead, the main results one can find are question-begging measures of whether OA has indeed been implemented. One wonders whether we are seeing another educational fad go nova, and what structures will remain among the clouds of gases. Advocates have succeeded in their efforts without the benefits of controlled studies. Those advocates include U.S. secretaries of education, accreditation agencies, and agencies that could have funded scientific research on these issues, and they make their case without appeal to hypothesis testing, without research projects, without publications of successes and failures, without comparison studies of any kind except those comparing before and after implementations, where indicators of success *at implementation* are assumed to be evidence for OA. No one has assessed the outcomes of implementing programs of outcomes assessment by comparing the results of such programs with the results of other forms of education.

Some to whom I have made this claim find it implausible. Perhaps we can count on rebuttals after this is published. Part of the puzzlement about how advocates of outcomes assessment could have succeeded without research can be resolved by realizing that many advocates think doing outcomes assessment *is* doing research because it results in data. But searches of the Internet, the archives of the *Chronicle of Higher Education*, recent books on outcomes assessment, and their bibliographies yield a clear picture of the state of research, which is that it is made up of substitutes for research. One is directed to manifestos claiming that outcomes assessment makes sense, is needed, is succeeding, and is the wave of the future. Judging from the number of accounts of universities, colleges, school systems, and accreditation agencies in the throes of implementing outcomes assessment, one might be excused for thinking all these claims are true. Among some of these, the job of implementing a program of stipulating outcomes and assessing them is taken up by entities with names like Office of Research and Analytic Studies, again suggesting that implementing outcomes assessment is somehow doing research. A review of literature also turns up journalistic commentary on the manifestos, usually written as though simply chronicling waves of the future. There are a few people dragging their feet and worrying about implications of implementing outcomes assessment, among whom I would place myself with this essay. Somewhere lower, that is, less common, on the list of what one finds in the OA literature are inquiries wondering, "Where the hell is the research?" Occasionally there are reports of educational organizations, such as the school system in New South Wales, Australia, withdrawing from or questioning their commitment to outcomes assessment initiatives.

Readers may still find it implausible that no controlled studies have addressed whether outcomes assessment improves education. There are journals and metastasizing bookshelves of literature, much of it originating in that part of education in which one can expect insecurities to result in scientific rigor. I finally wrote to the chair of the only PhD program in assessment and measurement in the country, at James Madison University, asking if she could direct me to any controlled studies. The first sentence of her reply was, "Your search for the holy grail and disappointment in finding it is fairly widespread." She told me of the careful work there to do outcomes assessment testing on students at orientation and as sophomores and juniors at an

annual Spring Assessment Day, and said, “We believe that the best comparison is with yourself over time, and we can document progress in a variety of ways.” She pointed out that finding institutions which do not do outcomes assessment will now be very difficult, since all regional accreditors are now mandating OA. Although she did not put it this way, in effect, any possible control group is now vanished. “The pressure from the feds on all accreditors is now unrelenting and growing in force” (Sundre 2011).

It is striking how the question about lack of research bewilders advocates of OA — the question and the advocates seem ships passing in the night. By now we would expect reports comparing results of doing outcomes assessment with programs in which there is no outcomes assessment. But outcomes assessment is seldom mentioned now except in contexts in which the speakers push to implement it. Advocates, convinced that the results of implementation will produce clear data, present before and after snapshots as evidence that implementation in fact increases an institution’s ability to meet its declared outcomes. The easy availability of that data is evidence that the advocates do not yet understand this question. Declaring outcomes and success at implementing outcomes is, after all, already a part of programs of outcomes assessment. What is needed is testing whether doing all that is educating better. The concern is partly whether the faith in OA is built on wishful thinking. And articulating outcomes is, after all, requires that those who do the declarations be wise about what they do. Where is the research showing that this step leads to better education of students? If you are convinced going in that better education is education which is meeting measurable objectives, then we’ve a no-brainer: of course OA-shaped education will be better (provided it produces evidence that it really is OA-shaped). On any robust conception of evidence, however, in which we insist on evidence that is not part of circular reasoning, the evidence is alarmingly absent. Indeed, when I have asked administrators what got them to endorse OA, all of them have cited their own individual anecdotes, mostly recognitions of shortcomings which came from reexamining their purposes in teaching, with no acknowledgement that they are overgeneralizing or reaching hasty conclusions based not on outcomes assessment or methodical checking but on the confounding variable of having reexamined how they teach — a much more broad and flexible

process than articulating measurable student outcomes, testing for them, and accumulating reams of documentation.

A review of those alarmingly large sections of library bookshelves dealing with outcomes assessment also reveals large overlaps with corporate calls for increased productivity in education. For that large camp of education reformers who think that higher education should be more businesslike, these overlaps seem to be part of the appeal. But calls for increased productivity should be informed by knowledge of what is being produced — what, in other words, education is for. I return to this below in the sections “Neglected (Because Unknown) Outcomes” and “Appropriate Actions for Us to Take Now.”

If Not Research, What Basis Is There?

The surprising lack of research raises the question of what explains advocates’ faith in the efficacy of outcomes assessment. We could here recall from Freud a relevant part of the argument for doing critical thinking: Often we believe not on the basis of arguments but on the basis of desire. Why would advocates want to believe in outcomes assessment? This takes digging. There are suggestions in the history of the idea and in relationships of the arguments for outcomes assessment to the research-based arguments we would expect but do not find.

Attention to the history of the idea is not a tangent. When movements hold their block parties, intellectual historians don’t get invited. And if intellectual historians do show up at the party, everyone starts observing what an ex-nun friend of mine calls chastity of the eyes, looking at the ground (e.g., Bracey 1995: 239). There’s a reason for hostility to historians. Historians are powerful corrosives of the work of True Believers (as Eric Hoffer’s philosophical and historical work of that title is corrosive of the general concept). Further, reading histories of higher education tends to cause cynicism and despair regarding reform. The history of outcomes assessment is rootbound in the same pot with American positivism and with behaviorism, and it is closely related to histories of management theories, which apparently mostly make progress by renaming themselves. Thorstein Veblen’s *The Higher Learning in America: A Memorandum on the Conduct of Universities by Business Men* (1918) analyzes how the secularization of universities

led to imposition, at many levels, of values and administrative structures that are extraneous to the academy's work. From Richard Hofstadter's *Anti-intellectualism in American Life* (1962) we learn, among other things, how insistence on accountability can be based on hostility toward and lack of understanding about education. Gerald W. Bracey's *Final Exam* (1995) presents five powerful, historically based essays anatomizing contemporary movements in education. Three of those are directly relevant to issues regarding outcomes: (1) the move to broaden access to include all, but insisting that this all achieve at a minimum level based on historical levels to which only more selective processes provided access (an insistence that has since led to "No Child Left Behind"); (2) the move to reform based on the claim, recurrent since Adam and Eve's kid flunked the unit on brotherhood, that schools are going to hell in a handbasket, always have been, and that the fix will start with better testing and assessment; and (3) the move to fix education, not by new commitments to teacher education or better schools but by articulating and imposing new testing. **It is a theme in Bracey's work that claims driven by ideology become resistant to data.** Jacques Barzun's *House of Intellect* (1959) and *The American University* (1968) attempt to distinguish core purposes of higher education from administrative tasks, money management, and oversimplified models of education which distract us from that core. Learning, guarding learning, and teaching are not the same as managing budgets, getting students, publishing, offering to name classrooms for people with money, or testing whether learning has happened. Barzun tells us that learning must even be distinguished from the excitement which is an effect of learning, and from other effects. Barzun's recommendations for bringing us back to what we are about are bracingly radical and specifically counter to historical trends calling for external accountability (see Barzun 1968: chap. 8, "The Choice Ahead," esp. 247–59). And of course we can remind ourselves of how terms such as *strategic*, *evidence*, *culture of evidence*, and *accountability* carry an authority that puts to sleep the critical faculties of both those who hear them and those who speak them.

One of the main lessons of history is that we have to keep our critical faculties. The history of educational reform, as Bracey points out, proceeds like the Bill Murray film *Groundhog Day*—we do the same day over and over and over as long as the lessons do not get learned. But if we step

back and regard not the dreary cyclical history of educational reforms but the analytical historical literature, we see a striking division between the great number of those who have remedies to sell and those who think. For those who are selling something, the world is a simple place. Critical faculties, as Veblen teaches us, are out of order when the agenda is advertising. But history offers powerful therapy for oversimplification. (J. L. Austin says of philosophers, and we could include most educational theorists, “Oversimplification is the occupational disease of philosophers—unless it is the occupation.”) Historians tend not to make good salespersons—though it’s not necessarily being a historian that makes a person thoughtful and critical, or we would not have to choose our histories carefully. Perhaps some of the historians’ critical attitude comes from finding that their best work comes with a certain distance. Outsiders who are not historians, like bell hooks, James Herndon, Socrates, and Paula Gunn Allen, but who offer powerful critiques of their and our cultures, provide confirmation of this.

The arguments for assessing outcomes are expressions of models of corporate business management. These models are startlingly temporary, in the sense that there is always a new one that gets attention and aspires to capture the hearts and minds of managers and administrators. (Though many of them fade from prominence only asymptotically: like behaviorists and Jehovah’s Witnesses, a blessed remnant remain.) Joel Best’s *Flavor of the Month: Why Smart People Fall for Fads* (2007) gives examples for higher education, and Robert Birnbaum’s *Management Fads in Higher Education* (2000) anatomizes several of these. One example is total quality management (TQM), whose blood quantum shows in outcomes assessment. TQM presents itself as a new, humane, compassionate, but still hardheaded approach to doing business by bringing to each activity not a stopwatch but the Sherlock Holmes lens of quality. Some impetus for that model may have come when Robert Pirsig made the word *quality* into a mantra, beginning with *Zen and the Art of Motorcycle Maintenance* (1974).

We, too, might be well advised to meditate on the word. One of the most difficult parts of implementing outcomes assessment is appraising quality. The oversimplifying dichotomy, quantitative (good, measurable, objective) versus qualitative (marginally scientific, difficult, subjective) still shapes these debates. Further, this issue of whether we should try to appraise

quality is going systemic in higher education. The current rash of program prioritization efforts on campuses is least regimented and least tied to criteria when it comes to providing evidence for a program's quality. Robert C. Dickeson's *Prioritizing Academic Programs and Services* (1999), which my administration has handed out like necklaces at Mardi Gras, includes a guarded note about the move to assess outcomes rather than inputs or processes in its section on appraising program quality. A few suggestions are made for possible outcomes to observe, but this section is much less prescriptive (which gets translated on campus as "less helpful") than most others. The section's last paragraph owns up to this lesser aspect by acknowledging that "assessing quality outcomes is generally regarded as more difficult and less precise than assessing quality inputs" (66–67). Assessing quality is difficult partly because, in the history of the term, it takes in the dirty laundry from quantitative, where objective evidence and measurability helped guarantee that the appraisers were not subject to delusion. The subjective-objective distinction, then, with its polluted sources in Cartesian dualism, still wells up to influence our work.

These sources are not new. Indeed, one characterization of them is present in Veblen's *The Higher Learning in America*. He anatomizes, provides examples, and then predicts pathologies that result from putting captains of industry and cocaptains of the pecuniary system in charge of higher education. One of his prophecies we can see fulfilled in reports surveying college graduates' ignorance. Over the last hundred years the alleged steady rise in the ignoramushood index among school and college graduates (see, e.g., Gatto 2008) could easily be an effect of increases in power and control by nonacademics over universities. External calls for increased accountability erode the quality of education. This helps explain the dreariness of the history of educational reform. Calls for increased accountability may be especially corrosive (and particularly consistent with Veblen's account) when they come from those whose interests are not deeply informed by education: narrowly educated theoreticians, managers, government bureaucrats, legislators, regents, and trustees. We need to produce inhabitants of those categories who are more broadly educated.

Neglected (Because Unknown) Outcomes: What Gets Left Out

Applying outcomes assessment to academic programs requires that those who assess know and understand what the programs do. That is, undergraduate degree-granting institutions need to know what it is they produce. This turns out to be problematic. I have become convinced that very few academics, much less trustees and administrators, have a clear picture of the contemporary bachelor's degree and what it means. If they did, general education would not be such an ineffective mess. What are the causes of this tunnel vision? We can round up all the usual suspects: the withering of interdisciplinarity and the metastasizing of specialization; the hardening of turf boundaries in the face of declining budgets; economic hardships in the larger society leading to emphasis on preparation to serve the labor market despite that market's regularly being beaten bare by winds of change; general education programs being shoved off the end of the bench by majors; an alleged decline in the number and quality of scholars who focus on the largest issues facing the academy.

In Humboldt State University's preparation for our accreditation renewal, a group of smart people with surprising goodwill and an astounding lack of agendas worked hard to come up with a brief list of the university's goals for graduates with a baccalaureate. We have been engaged for several years now in implementing outcomes assessment plans for these goals. I think ours are as good as it gets, though the Web is rife with these lists. Boiling all our work down to a list of seven goals may be—no, is—problematic. If it is not already obvious, we can see this by looking again at the process we followed and asking what gets left out.

We began the work of articulating these goals by asking groups of people to reflect on their own undergraduate degrees, on what had been the most important results of those years, and on particular outstanding students we have now, as well as what makes them outstanding. What struck me then was how varied the responses were, how insightful but also how idiosyncratic and difficult (at least for me) to regiment the answers into categories. Much of what we said back then does not show in final lists of goals and objectives.

Here are some of the missing factors, including some mentioned in those beginning conversations. College provides a safe vestibule for fledglings being kicked out of the nest, and its libraries and texts provide an entrance to the world. College is still a large part of many

adults' coming of age. It provides a substitute for apprenticeships and community rituals. It is instrumental in helping people find a calling, a vocation, a set of foundational values. Jobs may come and go, but college often still helps shape interests and directions that influence the rest of a person's life, and provides abilities and encouragement to pursue them. Small classes and conversations with faculty members expose our mistakes and give us practice in arguing (and joking) with professional people in which we are treated seriously, perhaps even treated as peers. This kind of education prepares students not so much for taking a place in existing power structures or class systems as for analyzing, influencing, and changing power structures. Reading Homer, Aeschylus, Sappho, Chuang Xu, Plato, Lesley Marmon Silko, and Maimonides forcibly decenters provincial points of view. Faculty members who share their academic lives with us, who welcome us into doing research, provide us models of scholarly involvement and activism. Faculty members who share their personal lives with us provide models of interests and enthusiasms that help us take our own seriously. Faculty members who are broadly educated, broadly activist, and broadly interested help us keep in perspective the regimented disciplinary boundaries that still exert powerful influence over the degree. Faculty members who manage to teach and write and serve on committees and still raise children and drink and coach Little League, provided we get to know them, help us integrate academics and community, gown and town. Faculty members who pay attention to analysis of social structures and who raise consciousness of the roles of gender, race, class, culture, money, and provincialism prepare students to cross boundaries imposed by these factors. Faculty members who take our young professors under their wings—mentors—were mentioned repeatedly in those first conversations.

Other important omissions in the thinking behind outcomes assessment are demonstrated in research on attempts to close gaps in educational achievement between successful and less successful groups. Closing these gaps requires attention to lives outside the classroom, to informal social processes, to relationships with parents and peers—none of these easily measurable in the classroom and almost none of them amenable to outcomes assessment in most courses (Ferguson 2006; Kim Abood and Kim-Hall 2006). (We would then imagine the

following: “Outcome two for History of Ancient Philosophy: Students will have hung out with the Plato wonks. Assessment measure: # hours students spend in informal settings with the Plato wonks.”)

None of this is new—we could regard the foregoing list as an application of John Dewey’s recommendations that a school be part of a community and part of the world. Of course that means a lot of different things to different people, but one main meaning is a repudiation of the factory model emphasizing measurements focused on inputs, productivity, and outcomes. Assessing the items on this list via outcomes is a parody of education. R. G. Collingwood (1939), the philosopher of history (among other things), teaches us that a primary determiner of the answers we get is the questions we ask, and that another determiner is the questions we don’t ask. Not asking what education is for may doom us to a stunted list of alternatives. Not asking what assumptions shape programs of outcomes assessment results in a parody of what we want, and distracts us from our own goals.

OA Pulls Resources from Teaching and Learning

As implemented, programs of outcomes assessment shift resources and emphases within programs away from those results that have been largely unarticulated and toward those that are being measured. Departments and faculty members teach to the tests. In order that course content not be eroded, the tests must be fair and complete. But in fact, completeness of testing interferes with completeness of courses, at least to the extent that assessment becomes a larger part of the courses. As a result, the completeness of both is routinely sacrificed for sampling efforts, and teaching to the test becomes a more dangerous possibility because the evidence of good teaching comes from a smaller portion of the course than before. After circulating drafts of these remarks to faculty members and administrators, vehement agreement regarding this point was the most common response I got.

OA Divides the Faculty

One problem, with effects on morale, is that many who are required to implement OA

understand the motives behind it and feel rebuked. Boards of trustees (which, as Veblen predicted a hundred years ago, are composed almost entirely of captains of industry), the chancellors and presidents they hire (and deans, if trustees can pull it off down that far on the organizational chart), and the legislators and governors and governors' higher education advisors still often look at four-year colleges and universities with envy and uncomprehending distrust. Workweeks measured by contact hours are short; faculty salaries are higher than factory workers who do real work; faculty members get a quarter of the year off; faculty members sit staring off into space and claim it's work; and new graduates keep demonstrating breathtakingly new levels of ignoramusness. Those who foot the bills and look over the budgets don't trust the faculty. Outcomes assessment gets taken up by such players as a Missouri "Show Me" project or a prenuptial agreement. It is a result of distrust of the university on the part of those who pony up the money, or rather those who administer for those who pony up the money. It also reinforces those faculty members who enjoy rebuking each other in underhanded ways (at times I am in this number as well). The fact that we are in the habit of preaching skepticism as part of critical thinking helps lower our guard to this kind of attitude. Saying, "Oh yeah? Show me," is an easy way to shove the burden of proof onto others and to slough off the burden on myself.

Standard practices from the civil service and corporations also illuminate this source for OA. If an employee is not performing, is dragging his or her feet, or is otherwise unproductive, then performance reviewers may put the employee on a plan to become better. These plans are called by names such as Personal Development Plan or Performance Improvement Strategic Plan. The plan has to move down the ladder of abstraction from the job description under which the employee was hired. It will contain more specific goals, often called "objectives" for this purpose, and clear criteria for whether the employee at the next performance review can be judged as improving. Putting employees on such development plans serves notice that their work is not good enough and is a forcible reminder that employees are underlings and that their continued employment is at stake and a way for the manager to reduce respect for them in the workplace. Such plans provide outcomes and outcomes assessment measures as ways to make

the subsequent decision, up or out, into a public, verifiable process based on measurable evidence. These measurements are meant to forestall grievances, answer charges of unfairness, and provide a basis for a case if there is a hearing. In higher education, then, these measures nudge up against retention, promotion, and tenure processes.

Some faculty members, with such practices as background, feel misunderstood, threatened, unrecognized, and defensive in ways that do not promote cooperation. They often point out that they give grades, read papers, and give tests, and that students take away transcripts and recommendations. These are assessments. Further, they do assess, rather than inputs, outcomes. Ignoramuses can be predicted and in fact are predicted from their transcripts, their grade point averages, and their faculty letters of recommendation. These processes are well understood by all involved—professors, students, those who hire graduates, parents. Faculty members feel universities are being pressured to become more like on-the-job training programs. This diminishes higher education to the extent that what higher education really is about is preparing students to handle a world in which their jobs will keep changing and their roles as citizens and thinkers will require more information, more investigative, creative, and critical skills than just one generation ago. Further, the supervisory functions are already performed within the academic world. Faculty members are not very good at trusting each other, and so they can point to policies, procedures, committees, and the work they already have to do to show each other they are putting out a “quality product.” Those policies and committees can be offered for consideration as effective tools of supervision and quality assurance. For example, policies are in place that stipulate procedures, forms, and required approvals anytime a faculty member wants to claim that a course should be a part of a university’s general education program. The main drawback of these processes, as Veblen points out (he questions the need for deans and deanlets) is that they require few professional managers and do not add to levels of administration on the organizational charts of universities.

Against oversimplifications, faculty members insist on neglected aspirations in higher education. Higher education is about not just memorization or skills but also evaluating what has been memorized. Higher education is about liberation, about questioning and arguing

values and authorities, about criticizing and refining citizenship roles, about acquaintance with the best ideas humans have had and adding to these, and about developing better peripheral vision to get past the blinders we don with our cultures and dogmas. Some may think we exaggerate the importance of these goals. Certainly for many students the goal of economic independence, preparation for a job, is the crucial first goal and should not be discounted. These other goals, though, are important to a larger picture, and faculty members are right to remind us of them. Looming global issues face our graduates in ways that make the most profound, pervasive, and deadly threats of video games look like good analogies for the world in which their main weapons will be a cap and gown.

Faculty members do disagree. Some professors regard it as reasonable that those who pony up the money be given evidence that their money is put to good use, and that the evidence be easily understandable by those who do the ponying. Some not only accept this view but endorse it with enthusiasm. Other faculty members roll their eyes but regard the push for outcomes assessment as inevitable and not outrageously unreasonable and so work with goodwill to put together assessment programs that will pass muster among the campus assessment coordinators and committees. But others, a great many among those with whom I talk, regard outcomes assessment as evidence that the town does not understand the gown (it should be clear by now that OA does not really originate in the town but in the administration and in accreditors and legislators). They regard the demand for evidence that the academy is doing its job well as unfair and as based on willful refusal to investigate how in fact professors do track and provide evidence of success appropriate to what they do. The strength of these negative reactions varies, including those who are merely depressed and those who are ready to hoist flags over the ramparts. These divisions among the faculty are of course not all bad; like committee work and strategic planning, they keep faculty members off the streets and out of the halls of the legislature. Most professors would agree, though, that OA and other productivity, efficiency, and accountability initiatives have real impact on teachers and classrooms. They tend to dilute education and do divide the faculty. The stronger versions of these views turn into indictments: outcomes assessment initiatives add to workload, increase class sizes, reduce conversations

between teachers and students and between students and other students, reduce the actual accountability of the students, reduce the amount of writing involved in a degree, push the students out the door more quickly, and hide from students the degree to which they are being cheated.

Nobody Reads Them

First, some comments which might seem to contradict the claim that no one reads outcomes assessments. I now serve on several campuswide committees by virtue of having been elected general faculty president, and have served on more. The university's Curriculum Committee, the Provost's Council, and an accreditation task force routinely look at program goals and outcomes assessment plans, and have been involved in implementing requirements to include statements of goals and objectives in syllabi and in program review documentation. All curriculum changes now are accompanied by statements of how the changes will accomplish goals, how the goals generate objectives, and how these objectives will be assessed. I read a lot of syllabi, and the syllabi are getting longer. I help write objectives so the Western Association of Schools and Colleges can see how we at my university accumulate evidence that we are accomplishing our goals. I read program reviews in general education as well as for departments and majors. The California State University is once again engaged in a systemwide strategic planning effort, "Access to Excellence," in which the language of goals, objectives, and assessment is prominent. All over the four-hundred-thousand-student California State University, the talk is of what our objectives are and how we will assess them, and faculty and staff members are writing statements about how we will do these things. I keep hearing proud talk, coming from the accrediting agencies and from the secretary of education, about "nurturing a culture of evidence." In many departments evidence of how those objectives have been assessed is accumulating in drawers and in large computer files. Departments do exit interviews and tests, students accumulate portfolios, and teachers collect data from each course and forward that data to department chairs and to committees. Each committee rows through tides of paperwork, higher tides in winter and bad weather, writing reports. The university's

Integrated Curriculum Committee has a two-and-a-half-inch binder each year for each member, and for each meeting we receive a packet including more outcomes. As I write, this two-thirds through the year, my packets no longer fit into four folders. Outcomes and assessments comprise about a fourth of those folders. But reports of assessment of outcomes are a much smaller part of the whole than are the stacks of promissory notes. In each department, there are one or two faculty members who have to review assessments and forward them. On each campus, a few administrators have to review what has been forwarded. These reviews are done for someone else who is far away. These efforts began with people who are off campus—systemwide administrators, the board of trustees, influential businessmen, legislators, staff in the governor’s office, accrediting agencies, the U.S. secretary of education, and national education organizations operating at a level of abstraction to which one must bring one’s own oxygen. Not one of these is going to look at a single document my department generates, nor will my department’s results even find their way into a summary of all the documents written by my committees or my college. The culture of evidence has turned into a culture of reams of paper, a culture of file cabinets and large pdf files. Our students’ transcripts get looked at seldom enough—at job interviews, by graduate admissions committees—but they do get looked at. They are blunter and quicker than the results of my department’s exit interviews or exit exams or the course outcomes assessment data or portfolio reviews. While transcripts do get read because they are useful, no one will ever read OA reports unless someone pays for it and requires doing so as part of some poor soul’s position description. The simple argument at the beginning of this essay, which keyed this kind of accountability to an alleged audience of taxpayers, legislators, parents, students, is producing warehouses of materials that are not clearly any more important to that audience than a campus’s signage and landscaping.

The Argument Disguises Its Speakers

Given the talk about educators using other people’s money (governments’, taxpayers’), it might seem as though this argument is being put forward by these other people. Strikingly, it is not. One has to watch out for the genetic fallacy here—the argument’s source need not be relevant to

its force. But when the argument is partly about the existence of a need – the people of the state need the academy to render an accounting – then the question of who is speaking for the people becomes germane. Do average taxpayers have problems with traditional means of assessing students, namely grades, majors and transcripts, and letters of reference? They clearly do not. Does this issue arise because taxpayers or employers find fault with college records? Clearly not. Perhaps these concerns arise with claims of rampant ignoramusness among graduates? Those who do pursue this argument are certainly happy to cite surveys emphasizing ignorance, which remain consistent since the 1920s (Bracey 1995). But these calls for reform and for increased accountability do not have their origins with the public. Instead they show up among administrators under the influence of schools of thought articulated by management theorists in business. These schools of thought are discussed in books by the historian Gerald Bracey, by Robert Birnbaum, and by Joel Best. All three provide bracing criticisms. Birnbaum and Best explicitly call them fads. All three are dubious about their final value and point to serious problems in the processes by which they are justified and by which they get implemented. History provides evidence that these remedies will likely yield to other inflated promises, other oversimplified cures, other fads. Fads give their devotees the appearance of being involved with a Cure, a movement, a Help, and administrators are under pressure, as Veblen points out, to produce something visible that will mark their campus or agency as belonging to such a movement and which will mark it as different from others in ways that increase prestige. One cannot help but wonder how many millions of dollars and hours of faculty and staff time could be saved if, instead of letting new presidents and provosts initiate strategic plans and new initiatives, we could just march them around the boundaries of the campus and have them pee on the corners.

Appropriate Actions for Us to Take Now

If the foregoing is at all correct, then there are clearly actions we can take, both to improve our situation and regarding outcomes assessment. Those actions will not appease those calling for more or faster implementation. They will involve thoughtful analysis rather than accelerating

the training of faculty members. They involve, wait for it, philosophical reexamination. Fourteen years ago, I wrote a piece for the National Education Association's *Thought and Action* on what education is for, claiming that when you are wrong about what education is for it is mad to insist on increased productivity, and that even though it looks a lot like doing nothing, doing some philosophy might save us from that madness. That point applies here (Powell 1998).

Scientific analysis would also help. A controlled study of whether degree programs that implement outcomes assessment in fact produce better educated graduates than programs that do not has not been done. One reason such a study cannot quite be done yet is that it will take some prior reflection. Such a study will need a noncircular, non-question-begging description of what it is to be "better educated." The study would be a major undertaking and would require careful design work and money. The design would include the possibility of disproving the claim that outcomes assessment programs help educational effectiveness. This means such a study would be unlike studies in which measures of implementation are taken as measures of success—which concept of research makes research like a dog incapable of barking in the night or the day, much less biting. One would need to control carefully for other variables. One variable William James points out that can confound results at a subliminal level is paying attention—if an institution can be improved simply by paying attention to what it is doing, then this is not evidence that a particular form of paying attention is the main factor in any resulting improvement.

One might save work, even, by doing a thought experiment: take an institution of unquestionably high quality, such as, perhaps, Harvard or Stanford or Columbia or St. John's; select some programs at that institution for implementation of outcomes assessment and select others as controls for a several-year study. We might have to begin by overcoming doubts as to whether such an effort would be profitable, but perhaps we could threaten their continued accreditation.

We also need an analysis that compares calls for accountability urged from outside higher education and those that emerge from the faculty. We need such an analysis because of the possibility that such an analysis might reveal that calls for accountability from outside the

academic world are based on misunderstandings and oversimplifications of education (for instance, thinking that education is about learning facts and job skills). Historians of education, especially of educational reform, could help. A history of higher education that focuses on separating changes which have been most beneficial from enthusiastic but ultimately futile reform efforts might illuminate our current situation. The history of reform in higher education can be matched for dreariness only by history of housing reform, and articulating the relations of meaningful and positive reforms to theories provides mostly cautionary lessons about theory. We might find that calls for accountability are a factor in recent reductions in educational quality.

Finally, there is an opportunity now for real reform of a quite different kind, not driven by theory and much more meaningful. Right now, there is a relevant, interesting, and perhaps pregnant pause in discussions of general education, that core part of the degree not presently included in majors. One gets the impression (something similar may be happening in linguistics) that discussants are waiting for a crucial innovation. Part of this is because doing outcomes assessment in general education is particularly difficult, raising more recalcitrant issues than the issues within most majors (Katz 2008). Part of it is captured in the report from the University of California's Center for Studies in Higher Education, *General Education for the Twenty-First Century*, which includes useful surveys of the state of general education in American universities and accounts of many attempts at reform. The report offers recommendations, but they do not address a recurrent problem: the reforms they document that have actually been pursued to completion mostly turn out to be piddling. Careful work by a strong task force—or perhaps just implementation by a strong institution—to articulate what is needed and what a coherent general education program would look like might find a receptive audience and might even influence discussions of outcomes assessments.

For instance, it may be that general education, after decades of being nearly pushed off the bench by expanding majors and by trustees pushing for faster throughputs, and in the face of increased and pressing world complexity, should be expanded by about a year. Perhaps each student should be required to take a minor, completing one organized cluster of courses in

humanities, another in social sciences, and another in natural sciences, separate from the student's major and each at least the equivalent of a semester's work, and perhaps each student should take a two-year sequence of courses in the intellectual history of the world.

Outcomes assessment is an odd business. It is not to the credit of higher education that we have tolerated this external assault on our work. Its 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.

John W. Powell is professor of philosophy at Humboldt State University in Arcata, California. His education has included a one-room school in the Ozarks of Missouri, public schools in Salt Lake City and Cabool, Missouri, a B.A. and graduate work in English at the University of Missouri, work as a conscientious objector medic in the US Army, and a Master's and PhD in philosophy at the University of Oregon. Current interests include Plato, Wittgenstein, and methods. He is currently General Faculty President at HSU. An article related to this one was published in the NEA's Thought and Action in Fall 1998 on the topic of what education is for.

Annotated Bibliography

Baldrige, J. V., et al. 1999. *Baldrige National Quality Program: Educational Criteria for Performance*

Excellence. Washington, DC: National Institute of Standards and Technology.

Barzun, Jacques. 1959. *House of Intellect*. (New York: Harper and Bros).

— — —. 1968. *The American University*. (New York: Harper and Row).

Best, Joel. 2007. "From Fad to Worse." *Chronicle of Higher Education*, April 14. Revised from his *Flavor of the Month: Why Smart People Fall for Fads* (Berkeley: University of California Press, 2007). Best offers explanations for the prevalence of fads, but in his focus on why people choose to endorse fads he neglects the powerful role of true believers among administrators, who require implementation from those who do not endorse the fad in question. This is conspicuously the case with outcomes assessment, where accrediting agencies, U.S. secretaries of education, trustees, presidents, and

provosts have decided in favor and so drag universities into compliance. Though Veblen found grim hope, with business taking over higher education, in the fact that businessmen don't really understand higher education and so are generally impotent to effect change, that hope may not apply here.

Birnbaum, Robert. 2000. *Management Fads in Higher Education*. San Francisco: Jossey-Bass. Birnbaum, who says, "I was once myself a votary," mostly focuses on the structures and life cycles of fads, but he nevertheless offers a veritable cold shower of examples. With many of these examples we get a brief history and genealogy, enough to get a sense of a taxonomy. The vocabulary in his analysis shows the bad seed also visible in outcomes assessment. Birnbaum also comments in his preface on the narrow perspective involved in asking, "Why can't a college be more like a business?" (one hears Rex Harrison as Professor Higgins, about to burst into song), and contrasts that with a better question, "Why can't a business be more like a college?"

Bracey, Gerald W. 1995. *Final Exam: A Study of the Perpetual Scrutiny of American Education*. Bloomington, IN: TECHNOS. This work is a model of critical thought based on wide-ranging historical foundations.

Buckman, Ken. 2007. "What Counts as Assessment in the Twenty-First Century?" *Thought and Action*, Fall, 29–38.

California State University Institute for Teaching and Learning. 1992. *Student Outcomes Assessment: What Makes It Work?* Long Beach: California State University. For considering what counts as research, see especially M. Riggs and J. Worthley, "Evaluation of Student Outcomes Assessment Pilot Projects in the California State University" (1–22), and the same authors' "Lessons from Pilot Projects" (23–30). Of particular concern is the degree to which the support and acquiescence of participants from the faculty and administration are identified as predictors of success. For true believers confirmation is easy to find.

Chickering, A., and Gamson, Z. 1987. "Seven Principles for Good Practice in Undergraduate Education." *Wingspread Journal* 9.2: 1. Available online at

http://www.johnsonfdn.org/Publications/ConferenceReports/SevenPrinciples/SevenPrinciples_pdf.pdf.

This very widely cited reference is interestingly different from how it is characterized by the outcomes assessment movement, which sees the seven principles as amenable to treatment as outcomes. For instance, see also the citation of the above and its characterization in Darrell W.

- Krueger, "Total Quality Management," in *Making a Difference: Outcomes of a Decade of Assessment in Higher Education* (San Francisco: Jossey-Bass, 1993), 269–78.
- Collingwood, R. G. 1939. *Autobiography*. London: Oxford University Press.
- Conant, John. 1959. *The American High School Today: A First Report to Interested Citizens*. New York: McGraw-Hill.
- Dickeson, Robert C. 1999. *Prioritizing Academic Programs and Services*. San Francisco: Jossey-Bass.
- Drucker, Peter F. 1954. *The Practice of Management*. New York: Harper/Collins.
- Evans-Pritchard, E. E. 1937. *Witchcraft, Oracles and Magic among the Azande*. Oxford: Oxford University Press. Evans-Pritchard provides an account of the education of medicine men (202–50). A briefer and more generalized account is in John S. Mbiti, *African Religions and Philosophy* (Nairobi: Heinemann, 1969), 166–71. There is, of course, a crying need for implementation of programs of outcomes assessment in these educations.
- Ferguson, Ronald F. 2006. *Toward Excellence with Equity: An Emerging Vision for Closing the Achievement Gap*. Cambridge: Harvard Educational.
- Friedland, E. N. 1974. *Introduction to the Concept of Rationality in Political Science*. Morristown, NJ: General Learning.
- Gatto, John Taylor. 2008. *Weapons of Mass Instruction*. Gabriola Island, BC: New Society.
- Herndon, James. 1968. *The Way It S'posed to Be*. New York: Simon & Schuster.
- — —. 1971. *How to Survive in Your Native Land*. New York: Simon & Schuster.
- Hoffer, Eric. 1951. *The True Believer: Thoughts on the Nature of Mass Movements*. **New York: Harper-Collins.**
- Hofstadter, Richard. 1962. *Anti-intellectualism in American Life*. New York: Random House.
- hooks, bell. *Teaching to Transgress*. New York: Routledge, 1994.
- Katz, Stanley N. 2008. "Taking the True Measure of a Liberal Education." *Chronicle of Higher Education*, May 23.
- Kim Abboud, Soo, and Jane Y. Kim-Hall. 2006. *Top of the Class: How Asian Parents Raise High Achievers — And How You Can Too*. New York: Berkley/Penguin.
- Newman, John Henry. 1947 [1852]. *On the Idea of a University*. Edited by Charles F. Harrold. New York: Longmans, Green.
- Pirsig, Robert. 1974. *Zen and the Art of Motorcycle Maintenance: An Inquiry into Values*. **New York: Wm. Morrow & Co.**
- Postman, Neil. 1995. *The End of Education: Redefining the Value of School*. New York: Knopf.

Powell, John W. 1998. "Productivity in Academe: What's Education For?" *Thought and Action: The NEA*

Higher Education Journal 5.12: 109–16. Available online at

http://www.nea.org/assets/img/PubThoughtAndAction/TAA_98Fal_12.pdf.

Ruben, Brent D., ed. 2004. *Pursuing Excellence in Higher Education*. San Francisco: Jossey-Bass. This

substantial volume gives a good snapshot of more or less current work written by major luminaries, informed by what Ludwig Wittgenstein suggested is a recurrent source of pseudoproblems in theory, namely, thinking that for every substantive (e.g., *excellence*) there must be a substance it names.

Perhaps because many of the companies turned out not to be so excellent after all, and because of a lack of interest in history, no mention is made of this book's ancestry going back to T. J. Peters and R. H. J. Waterman's bestseller *In Search of Excellence: Lessons from America's Best-Run Companies* (New York: Harper-Collins, 1982).

Seymour, D. T. 1992. *On Q: Causing Quality in Higher Education*. New York: Macmillan. See especially the summary of eleven different elements (one might say meanings) of *quality* (19–21) and the explicit relation to corporate models for education (128).

Sundre, Donna L. 2011. Personal communication. Sundre is chair of the PhD Program in Assessment and Measurement and executive director of the Center for Assessment and Research Studies at James Madison University.

Veblen, Thorstein. 1899. "The Higher Learning as an Expression of the Pecuniary Culture." In *The Theory of the Leisure Class*. New York: Macmillan.

— — —. 1918. *The Higher Learning in America: A Memorandum on the Conduct of Universities by Business Men*. New York: Macmillan.

Wingspread Group on Higher Education. 1993. *An American Imperative: Higher Expectations for Higher Education*. Racine, WI: Johnson Foundation.