

JOAN HAWTHORNE

Accountability & Comparability

PERSPECTIVES

IN MID-NOVEMBER, the National Association of State Universities and Land-Grant Colleges (NASULGC) announced its official approval of the Voluntary System of Accountability (VSA). For those of us who have been hiding under rocks during the past few years (not such an odd place to find academics, whose interests often run more to details of Shakespearean tragedy or, indeed, rocks, than to higher education policy), here's why that matters.

General education is a poor measure of institutional effectiveness and a poor basis for institutional comparability

Background

The VSA is an indirect outgrowth of the work of the Spellings Commission, a group that served at the behest of Secretary of Education Margaret Spellings and that was charged with examining problems in higher education. Accountability was one of those problems. After extensive research and debate, the Spellings Commission, under the leadership of chair Charles Miller, former regent for the University of Texas system (and the same person who led development of a K–12 accountability system that eventually served as a model for No Child Left Behind), issued a series of recommendations.

At an early stage in the drafting of those recommendations, Miller lobbied hard to mandate standardized value-added testing of general education skills, using a measure like the Collegiate Learning Assessment (CLA), as a basis for accountability and comparability among institutions. In response to push-back from higher education, the mandate was dropped from the final version—but only after

JOAN HAWTHORNE is assistant provost at the University of North Dakota, where she is charged with responsibility for assessment of student learning.

What's Wrong with the VSA Approach?

the commission received assurances from organizations like NASULGC and the American Association of State Colleges and Universities (AASCU) that they themselves would take the lead in developing a rigorous system of assessment and accountability. NASULGC and AASCU appointed a committee to develop such a system, which they promised would fulfill both the letter and spirit of the Spellings Commission's intent.

The result of this committee's work is the VSA. Unfortunately, the VSA seems predicated on the assumption that the Spellings Commission approach is inevitable. Therefore, the VSA proposes mechanisms to control implementation of value-added testing (for example, via a phase-in, with decision points for campuses along the way), while failing to challenge the premise that standardized testing of purposefully generic (and thus theoretically broadly applicable) general education outcomes is the best means of encouraging and measuring institutional quality.

Further, the VSA, like institutional accreditation itself, is described as "voluntary." But the intent, clearly, is that participation in something like the VSA will quickly become expected and, de facto, an essential practice for legitimate institutions of higher education.

What's wrong with general education skills as the unit of comparability?

Those of us who take general education seriously might at first imagine the potential benefits of this nationwide focus on general



University of
North Dakota

education outcomes. There would be satisfaction in seeing general education, often completed in the first year or two of college and described disparagingly to students as something to be “gotten out of the way,” more publicly valued. But however important general education should be to institutions, it is a poor measure of institutional effectiveness and a poor basis for institutional comparability. Here’s why.

First and most important, there’s the issue of student mobility. While courses in the major usually are taken primarily at the degree-granting institution, many students (and at some colleges and universities, most students) take some or all of their general education curricula at other schools. Given the political and pragmatic pressures for convenient portability of credits, especially general education credits, it is very difficult for an institution to control that portion of a graduate’s curriculum. Comparing institutional outcomes on skills primarily taught in the general education curriculum is therefore likely to tell us very little about the graduating institution.

Even if we argue that a degree-granting institution must accept responsibility for the

Comparing institutional outcomes on skills primarily taught in the general education curriculum is likely to tell us very little about the graduating institution

effectiveness of courses and curricula students have completed elsewhere, it is important to note that the last half of a student’s curriculum is typically taken within a major. Near the time of graduation, it is the major with which students likely identify. Learning that occurs within the major courses can therefore

be expected to strongly influence student performance on tests administered shortly before graduation. Although at first blush, general education skills like critical thinking and written communication may appear generic, the fact is that these skills, as enacted within the disciplines, can be surprisingly variable. As almost-graduates, both English majors and engineering majors may be highly experienced critical thinkers, but the generic skills learned in general education courses will have been largely retooled to reflect the thinking and reasoning processes most valued in each student’s own discipline.

Measuring critical thinking on a standardized national test requires ignoring these disciplinary differences. If the test-makers’ definition of critical thinking aligns well with skills valued within an institution’s most popular majors, scores will likely be good. But if not, well, someone within the university will be searching for a different test. Little information of value—to the public or to the institution itself—will have been gained.

Standardized tests of general education outcomes among senior students are therefore of questionable value as measures of institutional quality (although administration of a carefully selected test that aligns well with institutionally valued outcomes may be another matter). As the primary measure of comparability and accountability, such tests must be deeply suspect.

Standardized testing for comparability and accountability

Flaws in existing tests and testing methodologies make matters worse. The Collegiate Learning Assessment (CLA), often described as one of the best of a new generation of standardized assessments of general education outcomes, demonstrates just one of the problems with such tests. The CLA attempts to measure value-added gains in skills like critical thinking,

Voluntary System of Accountability (VSA)

Through “College Portrait,” an online reporting template, the Voluntary System of Accountability (VSA) provides key higher education stakeholders with a consistent framework for information on the undergraduate student experience. The template is five pages in length, and the data elements are organized into three sections: (1) student and family information, (2) student experiences and perceptions, and (3) student learning outcomes.

The VSA project is the result of a partnership between the American Association of State Colleges and Universities and the National Association of State Universities and Land-Grant Colleges. The boards of both associations approved the VSA in November 2007. The template and explanations are available online at www.voluntarysystem.org.



written communication, analytical thinking, and problem solving, a process that requires testing both freshmen and seniors. Getting freshmen to take such a test seriously (despite the 180-minute testing period required for a longitudinal administration, 90 minutes for cross-sectional) is challenging but usually manageable. Getting similarly committed participation from savvy college seniors, for whom the test is unlikely to “count,” is often exceedingly difficult. And even if seniors can be persuaded to take the test, can they be counted on to take it seriously? If they are about to graduate as part of a class of five or twenty thousand, can they be expected to take the test as conscientiously as students graduating from a small private school where institutional loyalty has been carefully nurtured over four years? And, if not, will the scores be “comparable” in any meaningful way?

There are other concerns about general education outcomes testing as well. Questions have been raised about what the tests actually measure, whether it is reasonable to assume “content neutrality” can or should be achieved in tests, whether what gets measured in a skills test is really the skill that universities intend to teach, whether it is theoretically possible to segment out performance on “generic” general

education skills to the exclusion of years spent in a nongeneric major, etc. Each question represents a major issue regarding the value of the test and a reason for caution about widespread implementation for institutional comparability.

Other (better) options

Taken in concert, these questions must raise grave concerns about the standardized testing of general education outcomes for institutional comparability and accountability. Other approaches, including other testing options, would likely be more meaningful and more useful—and some of these other approaches are already available.

For example, a number of disciplines already use standardized testing for professional licensure or certification. At least one discipline within the traditional arts and sciences core (chemistry) has available a similarly well-regarded test of professional knowledge. For schools that offer majors where standardized exit testing already occurs and is well regarded, scores from such tests could be a useful comparability measure. If this approach were taken, additional discipline-specific tests might be developed, extending the reach of the measures, or relevant subject-area GRE

The College of
William and Mary



Valid Assessment of Learning in Undergraduate Education (VALUE)

In fall 2007, the Association of American Colleges and Universities (AAC&U) was awarded a grant from the U.S. Department of Education's Fund for the Improvement of Postsecondary Education to support an initiative on student learning assessment. Titled *Rising to the Challenge: Meaningful Assessment of Student Learning*, the initiative establishes a consortium among AAC&U, the American Association of State Colleges and Universities, and the National Association of State Universities and Land-Grant Colleges to build campus leadership and capacity to implement meaningful student learning assessment approaches and use assessment results to improve levels of student achievement.

For its part of the project, titled *Valid Assessment of Learning in Undergraduate Education (VALUE)* and initially funded by the State Farm Industries Foundation, AAC&U is leading an effort to develop an e-portfolio framework for assessing a wider array of learning outcomes than those measured by current tests. This part of the project will foreground practices that base assessments on authentic examples of student work collected over time in an e-portfolio. This research and development effort will collect and synthesize best practices in faculty-developed rubrics to highlight commonalities of outcomes and expectations of achievement levels across institutions. AAC&U will also develop models and templates through which e-portfolios can be used to demonstrate, share, and assess student accomplishment of advanced and integrative learning outcomes. For more information on the VALUE project, visit www.aacu.org/value.

exams might be considered for use as a proxy. If "passing" such a test were a route to licensure or certification in the field, or if employers learned to ask students to report their exit test scores in the same way employers now ask about GPA, the problem of student compliance would be solved.

Of course, this still raises the question of what the public might learn from such scores that would be useful when comparing institutions. But the point is that there are many approaches to assessment of learning within an

institution, and there are many purposes such assessment can serve. Once a particular system is implemented, even partially and voluntarily, it may rapidly become inevitable in the same way that No Child Left Behind (NCLB) now seems inevitable (or in the same way that ACT, SAT, and GRE tests became inevitable some years ago). However flawed the NCLB system, the question most often asked about it today is how Congress might tinker with it at the margins—not whether it has proven beneficial to student learning. Yet that remains a significant question.

I believe the current proposal to implement the VSA and its "College Portrait" scheme is too hasty. It's like a book that's been rushed to print to capitalize on the death of a celebrity—there just hasn't been time to get it right.

For too long, colleges and universities have resisted pressures for greater accountability for student learning, taking a head-in-the-sand, "if we don't acknowledge it, maybe it'll go away" approach. That's been a mistake. But I am equally suspect of knee-jerk reactions, and there is danger of that now, in the wake of the report from the Spellings Commission.

Furthermore, this may be the worst time, in some ways, to plunge into a testing approach to assessment. Implementing the VSA will focus institutional and national attention on assessment as a measure of institutional value rather than as a tool for improvement. Institutional resources for assessment will get poured into testing, which would be clearly beneficial to testing companies but less clearly beneficial to student learning. Students' assessment participation will be largely spent on particular measures of still-questionable utility. Just as faculty are beginning to ask serious questions and generate useful answers about student learning, their energies will be diverted into outcomes testing—an approach that will simply reinforce suspicions that there was always a secret agenda behind talk about assessment. We'll generate scores for convenient (if not necessarily meaningful) comparison. But will education be any better as a result?

Surely we can do better. We owe it to our students and our institutions to make the effort. □

To respond to this article, e-mail liberaled@aacu.org, with the author's name on the subject line.