

Should College Students Be Tested to Hold Institutions Accountable for Student Learning?

Poor student mastery of basic skills shows higher education must maintain higher standards.

BY STEVE UHLFELDER

BEFORE THEY AWARD diplomas, colleges and universities must more effectively measure and improve students' performance in basic skills such as writing and critical thinking. Nothing makes clearer the need for change than a recent survey of college graduates, which reveals that a large number simply are not mastering the basic skills they need to compete.

An American Institutes for Research (AIR) survey found a majority of participating seniors at 80 randomly selected colleges and universities were unable to perform some common tasks that one would expect college graduates to have mastered: They could not calculate the cost per ounce of a food item. They could not compare or understand the viewpoints of two conflicting editorials. Even worse, 80 percent lacked basic quantitative literacy skills.

Students' actual coursework may be one contributor to these results. The AIR study found that literacy levels are significantly higher among students who say their coursework places a strong emphasis on applying theories or concepts to practical problems, in comparison with students who say their coursework rarely

touches on these skills. Frequently, students have too much autonomy when choosing their courses. Many large universities allow professors to use too many multiple-choice tests instead of essay exams and research papers. This may be exacerbated in part by underfunding, which creates larger class sizes.

To deal with such problems, the U.S. Secretary of Education's Commission on the Future of Higher Education is considering, among other measures, standardized testing. This controversial idea, which some members floated and then pulled away from, would be one way of prodding colleges to do something new and better.

Changing the prevailing attitudes at universities often can be as slow as melting an iceberg. Well-intentioned and dedicated faculties strongly object to any encroachment on their freedom or alternative measurements of educational objectives. But the results clearly indicate that something needs to change if our graduates are going to be prepared for the new and complex world marketplace.

Florida is one of a few states taking serious steps to deal with these problems. In 2004, when I was a member of Florida higher education's board of governors, I initially proposed some form of standardized testing as an accountability step to measure critical thinking and writing before graduation. As one might expect, the idea was unpopular. Eventually, the board reached a compromise and

decided to mandate that state universities adopt "Academic Learning Compacts" for their baccalaureate degree programs. The compacts require universities to establish simple, clear expectations for every student who enrolls and to determine whether each graduating student meets those standards with respect to content-area knowledge, critical thinking, and communication skills.

Our state's Academic Learning Compacts will make college degrees more meaningful and comparable. More important, the compacts will help ensure that graduates are capable of the work expected of them. According to the new policy, the compacts can use multiple assessment measures to certify academic competency. These could include standardized tests, essays, licensure exams, or a combination of these and other methods. However, the ultimate test of effectiveness will be what students' future employers think of the talents of each institution's undergraduates.

There is no silver bullet for assessing and improving higher education. It has taken us a while to get here, and the problems will not be solved overnight. However, real accountability measures will give us better benchmarks of undergraduates' skills.

Steve Uhlfelder (steve@sulaw.net) is a Tallahassee lawyer, a former chair of the Florida Board of Regents, and a former member of the Florida Board of Governors.

Colleges are looking at new ways to assess student learning and institutional effectiveness.

BY STEPHEN KLEIN

THE FEDERAL “No Child Left Behind” law (NCLB) requires that states test all their third through eighth graders, as well as students in at least one high school grade, annually in reading and mathematics. Testing these students makes sense because of the critical importance of these subjects and because virtually all students receive instruction in them.

In that broad sense, there is a common core curriculum applicable to every school’s mission. States reinforce this common core in several ways, such as specifying statewide curriculum goals for each subject and grade level, selecting certain tests for statewide use, and approving certain textbooks for statewide adoption.

But postsecondary education is different. Colleges and universities have varying missions. Students choose differing academic majors. Even within an institution, faculty members often disagree on what should be taught in their field and how learning should be measured. In short, one size does not fit all.

Hence, the NCLB approach is not applicable to colleges. However, the increasing push for accountability by parents, trustees, and elected officials has led accrediting agencies and the public to put more pressure on colleges and universities to provide information about educational outcomes. This pressure has led some to

suggest looking at how well students perform on admissions tests for graduate and professional schools.

While this approach helps solve the problem of student motivation, it has serious flaws. Different admissions tests measure different abilities. The abilities these tests measure may not be central to a school’s mission, and only a small percentage of graduates take such exams. And those who do are unlikely to be representative of a school’s graduates. Using graduate admissions tests to assess an institution’s effectiveness is therefore analogous to measuring the oral communication skills of a school’s graduates on the basis of the performance of its debate team.

NCLB and graduate school admissions tests rely almost exclusively on multiple-choice questions. This reliance stems from statistical, logistical, and cost considerations. For example, it is much cheaper to process several hundred machine-scored answer sheets than it is to grade a student’s essay. Although multiple-choice tests are effective in measuring many things, they are not appropriate for evaluating the higher order abilities we expect of college students. For example, open-ended tests are needed to assess a student’s ability to spot flaws and logical inconsistencies in someone else’s arguments and to integrate information from various sources to effectively form, present, and justify their own conclusions.

Some colleges have responded to accountability pressures by leaving it up to faculty in each department to develop their own indicators of student learning that are tailored to their specific goals. While this strategy has

some positive features, it cannot assess whether the improvements that do occur at an institution (say, between freshmen and senior years) are exemplary, adequate, or below par.

To make that type of judgment, we need benchmarks that can be used across institutions. Only then can we assess whether the gains at a school are more or less than what would be expected given the incoming abilities of its students.

These considerations have led the Council for Aid to Education to develop a set of open-ended measures called the Collegiate Learning Assessment (CLA) that are administered and scored online. These measures assess some of the important abilities that are mentioned in almost every college’s mission statement, such as written communication skills, problem solving, critical thinking, and analytic reasoning. CLA results are reported in terms of whether a school’s gains are more or less than what would be expected given (1) its students’ entering abilities (as measured by their SAT and ACT scores) and (2) the gains of comparable students at other schools.

Under NCLB, each state selects its own tests and passing scores; thus, there is no uniformity in standards. In contrast, the CLA uses the same value-added metric for all schools, encourages high standards, and emphasizes improvement. I see this as the more promising option for assessing learning.

Stephen Klein (klein@rand.org) is a senior research scientist at the RAND Corp. in Santa Monica, Calif., and director of research at the Council for Aid to Education in New York City.