September 19, 2010

**A Measure of Education Is Put to the Test**

**Results of national exam will go public in 2012**

Wyatt McSpadden for the Chronicle

The Collegiate Learning Assessment measures "the skills that you want every undergraduate to walk away with," says Pedro Reyes, associate vice chancellor for academic planning and assessment at the U. of Texas.

By David Glenn

You have 90 minutes to complete this test.

Here is your scenario: You are the assistant to a provost who wants to measure the quality of your university's general-education program. Your boss is considering adopting the Collegiate Learning Assessment, or CLA, a national test that asks students to demonstrate their ability to synthesize evidence and write persuasively.

The CLA is used at more than 400 colleges. Since its debut a decade ago, it has been widely praised as a sophisticated alternative to multiple-choice tests. At some colleges, its use has helped spark sweeping changes in instruction and curriculum. And soon, many more of the scores will be made public.

But skeptics say the test is too detached from the substantive knowledge that students are actually expected to acquire. Others say those who take the test have little motivation to do well, which makes it tough to draw conclusions from their performance.

You may review the following documents:

* Graphs of Collegiate Learning Assessment scores on the University of Texas system's campuses over a four-year period.
* [An essay](http://www.aacu.org/liberaleducation/le-sp08/le-sp08_Hawthorne.cfm) in which an assistant provost at a flagship campus describes her "grave concerns" about using CLA scores to compare different colleges.
* A report in which the CLA's creators reply to their critics.

Your task: Write a two-page memorandum to your boss that describes and analyzes the major arguments for and against adopting the CLA. When you have finished, please hand your materials to the proctor and leave the room quietly.

It is easy to see why the test format that you just tasted has been so appealing to many people in higher education. The CLA is a direct measure of skills, in contrast to surveys about how much time students spend studying or how much they believe they have learned. And unlike multiple-choice-based measures of learning, the CLA aspires to capture a student's ability to make an argument and to interpret multiple types of evidence. Those skills are close to the heart of a liberal-arts education.

"Everything that No Child Left Behind signified during the Bush administration—we operate 180 degrees away from that," says Roger Benjamin, president of the Council for Aid to Education, which developed and promotes the CLA. "We don't want this to be a high-stakes test. We're putting a stake in the ground on classic liberal-arts issues. I'm willing to rest my oar there. These core abilities, these higher-order skills, are very important, and they're even more important in a knowledge economy where everyone needs to deal with a surplus of information." Only an essay test, like the CLA, he says, can really get at those skills.

Richard J. Shavelson, an educational psychologist at Stanford University and one of the CLA's creators, makes a similar point in his recent book, *Measuring College Learning Responsibly: Accountability in a New Era* (Stanford University Press). "If you want to find out not only whether a person knows the laws governing driving but also whether she can actually drive a car," he writes, "don't judge her performance solely with a multiple-choice test. Rather, also administer a behind-the-wheel driving test."

"The CLA is really an authentic assessment process," says Pedro Reyes, associate vice chancellor for academic planning and assessment at the University of Texas system. "The Board of Regents here saw that it would be an important test because it measures analytical ability, problem-solving ability, critical thinking, and communication. Those are the skills that you want every undergraduate to walk away with." (Other large systems that have embraced the CLA include California State University and the West Virginia system.)

One feature that appealed to Mr. Reyes and his colleagues is that the CLA typically reports scores on a "value added" basis, controlling for the scores that students earned on the SAT or ACT while in high school. In raw terms, the highest scores in the Texas system are at Austin and Dallas, the most-selective campuses. But in value-added terms, it appears that students at San Antonio and El Paso make stronger gains between their freshman and senior years.

The CLA's overseers, however, say they do not want colleges to become overly concerned with bean-counting and comparing public scores. Instead, they emphasize the ways in which colleges can use their own CLA scores to experiment with improved models of instruction. Since 2007, Mr. Benjamin's organization has invested heavily in "performance-task academies," which encourage colleges to add CLA-style assignments to their liberal-arts courses.

One campus that has gone down that road is the University of Evansville, where first-year-experience courses have begun to ask students to do performance tasks.

"We began by administering a retired CLA question, a task that had to do with analyzing crime-reduction strategies," says Brian R. Ernsting, an associate professor of biology at Evansville. "We talked with the students about the modes of thinking that were involved there, how to distinguish correlation from causation and anecdotes from data."

Similar things are happening at Pacific Lutheran University. "Our psychology department is working on a performance task that mirrors the CLA, but that also incorporates disciplinary content in psychology," says Karen E. McConnell, director of assessment. "They're planning to make that part of their senior capstone course."

**How to Interpret the Scores?**

Mr. Ernsting and Ms. McConnell are perfectly sincere about using CLA-style tasks to improve instruction on their campuses. But at the same time, colleges have a less high-minded motive for familiarizing students with the CLA style: It just might improve their scores when it comes time to take the actual test.

And that matters, in turn, because by 2012, the CLA scores of more than 100 colleges will be posted, for all the world to see, on the "College Portrait" Web site of the Voluntary System of Accountability, an effort by more than 300 public colleges and universities to provide information about life and learning on their campuses. (Not all of the colleges have adopted the CLA. Some use the Educational Testing Service's "Proficiency Profile," and others use the ACT's Collegiate Assessment of Academic Proficiency.)

A few dozen colleges in the voluntary project, including those in the Texas system, have already made their test scores public. But for most, the 2012 unveiling will be a first.

"If a college pays attention to learning and helps students develop their skills—whether they do that by participating in our programs or by doing things on their own—they probably *should* do better on the CLA," says Marc Chun, a research scientist at the Council for Aid to Education. Such improvements, he says, are the main point of the project.

But that still raises a question: If familiarizing students with CLA-style tasks does raise their scores, then the CLA might not be a pure, unmediated reflection of the full range of liberal-arts skills. How exactly should the public interpret the scores of colleges that do not use such training exercises?

Trudy W. Banta, a professor of higher education and senior adviser to the chancellor for academic planning and evaluation at Indiana University-Purdue University at Indianapolis, believes it is a serious mistake to publicly release and compare scores on the test. There is too much risk, she says, that policy makers and the public will misinterpret the numbers.

"Standardized tests of generic skills—I'm not talking about testing in the major—are so much a measure of what students bring to college with them that there is very little variance left out of which we might tease the effects of college," says Ms. Banta, who is a longtime critic of the CLA. "There's just not enough variance there to make comparative judgments about the comparative quality of institutions."

Compounding that problem, she says, is the fact that most colleges do not use a true longitudinal model: That is, the students who take the CLA in their first year do not take it again in their senior year. The test's value-added model is therefore based on a potentially apples-and-oranges comparison.

The test's creators reply that they have solved that problem by doing separate controls for the baseline skills of freshman test-takers and senior test-takers. That is, the freshman test-takers' scores are assessed relative to their SAT and ACT scores, and so are senior test-takers' scores. For that reason, colleges cannot game the test by recruiting an academically weak pool of freshmen and a strong pool of seniors.

Another concern is that students do not always have much motivation to take the test seriously. That problem is especially challenging with seniors, who are typically recruited to take the CLA toward the end of their final semester, when they can already taste the graduation champagne. Who at that stage of college wants to carefully write a 90-minute essay that isn't required for any course?

For that reason, many colleges have had to come up with elaborate incentives to get students to take the test at all. (See the graphic below.) A recent study at Central Connecticut State University found that students' scores were highly correlated with how long they had spent writing their essays.

**Take My Test — Please**

The Collegiate Learning Assessment has been widely praised. But it involves an arduous 90 minutes of essay writing. As a result, many colleges have resorted to incentives and requirements to get students to take the test, and to take it seriously.

View [Take My Test—Please](http://maps.google.com/maps/ms?hl=en&ie=UTF8&t=p&msa=0&msid=104356845934162344685.00049078047e4ec01f90b&ll=38.959409,-98.789062&spn=27.242942,53.085938&z=4&source=embed) in a larger map

As of last week, there were some significant bugs in the presentation of CLA scores on the College Portrait Web site. Of the few dozen universities that had already chosen to publish CLA data on that site, roughly a quarter of the reports appeared to include erroneous descriptions of the year-to-year value-added scores. In some cases, the errors made the universities' gains appear better than they actually were. In other cases, they made them seem worse.

Seniors at California State University at Bakersfield, for example, had CLA scores that were 155 points higher than freshmen's, while the two cohorts' SAT scores were similar. The College Portrait site said that the university's score gains were "below what would be expected." The University of Missouri at St. Louis, meanwhile, had senior scores that were only 64 points higher than those of freshmen, and those two cohorts had identical ACT scores. But those score gains were reported as "well above what would be expected."

"It doesn't make sense, what's presented here," said Stephen Klein, the CLA's director of research and development, when *The Chronicle* pointed out such discrepancies. "This doesn't look like something we would produce." Another official at the Council for Aid to Education confirmed that at least three of the College Portrait reports were incorrect, and said there appeared to be systematic problems with the site's presentation of the data.

As *The Chronicle* went to press, the Voluntary System of Accountability's executive director, Christine M. Keller, said her office would identify and fix any errors. The forms that institutions fill out for the College Portrait, she said, might be confusing for administrators because they do not always mirror the way the CLA itself (and the Collegiate Assessment of Academic Proficiency and ETS's Proficiency Profile) present their official data. In any case, Ms. Keller said, a revised version of the College Portrait site is scheduled to go online in December.

It is clear that CLA scores do reflect some broad properties of a college education. In a study for their forthcoming book, *Academically Adrift: Limited Learning on College Campuses* (University of Chicago Press), the sociologists Richard Arum and Josipa Roksa asked students at 24 colleges to take the CLA during their first semester and then again during their fourth. Their study was conducted before any significant number of colleges began to consciously use CLA-style exercises in the classroom.

The two authors found one clear pattern: Students' CLA scores improved if they took courses that required a substantial amount of reading and writing. Many students didn't take such courses, and their CLA scores tended to stay flat.

The pattern was consistent across the ability spectrum: Regardless of whether a student's CLA scores were generally low or high, their scores were more likely to improve if they had taken demanding college courses.

So there is at least one positive message in Mr. Arum and Ms. Roksa's generally gloomy book. Colleges that make demands on students can actually develop their skills on the kinds of things measured by the CLA.

"We found that students in traditional liberal-arts fields performed and improved more over time on the CLA," says Mr. Arum, a professor at New York University. "In other fields, in education, business, and social work, they didn't do so well. Some of that gap we can trace back to time spent studying. That doesn't mean that students in education and business aren't acquiring some very valuable skills. But at the same time, the communication and reasoning skills measured by the CLA really are important to everyone."

**Dueling Purposes**

For more than a century, scholars have had grand visions of building national tests for measuring college-level learning. Mr. Shavelson, of Stanford, sketches several of those efforts in his book, including a 1930s experiment that tested thousands of students at colleges throughout Pennsylvania. (Sample question: "Of Corneille's plays, 1. *Polyeucte,* 2. *Horace,* 3. *Cinna,* 4. *Le Cid* shows least the influence of classical restraint.")

Mr. Shavelson believes the CLA's essays and "performance tasks" offer an unusually sophisticated way of measuring what colleges do, without relying too heavily on factual knowledge from any one academic field. But in his book he also notes the tension between the two basic uses of nationally normed tests: Sometimes they're used for internal improvements, and sometimes they're used as benchmarks for external comparisons. Those two uses don't always sit easily together. Politicians and consumers want easily interpretable scores, while colleges need subtler and more detailed data to make internal improvements.

Can the CLA fill both of those roles? That is the experiment that will play out as more colleges unveil their scores.