

MEASURE BY MEASURE

A new effort to determine how well schools teach

BY JAY MATHEWS

.....

Almost since *U.S. News & World Report* began publishing its list of "America's Best Colleges," in 1983, there have been complaints within academia that the magazine's rankings are distorted—that they largely measure how selective a school is, rather than how good an education it offers—and that they have undue influence among college-bound students and their families. Seven years ago Richard H. Hersh, then the president of Hobart and William Smith Colleges, tried to do something about it. At the annual meeting of the Annapolis Group, an organization of liberal arts colleges, Hersh proposed a radical idea to his fellow college presidents: "Why don't we just stop supplying them with our data?"

That particular plan went nowhere, but Hersh refused to relent. Leaving Hobart and William Smith in 1999, he set up shop in his white clapboard house in Hamden, Connecticut, and devoted himself to the question of how colleges could truly measure how much their students learned.

American higher education has been trying to answer this question for decades. One of the most recent efforts has been the National Survey of Student Engagement, launched in 1999. (See "[What Makes a College Good?](#)" by Nicholas Confessore, November 2003 *Atlantic* .) Used by more than 850 colleges and universities, the NSSE (pronounced "nessie") polls undergraduates on their collegiate experiences: what they think they gained from their classes, how much interaction they had with professors, and so forth. But Hersh wanted to go beyond this, to find a way of determining not just the conditions students learned under but *how much* they actually learned. When he discovered that Roger Benjamin, the president of the RAND Corporation's Council for Aid to Education, was looking at the same question, the two men put together a research group, based in New York City, with about a dozen employees and outside advisers. In 2002 the organization unveiled the fruit of its labors: a three-hour test called the Collegiate Learning Assessment, or CLA.

The purpose of the test is to measure not the particular facts students have memorized but, rather, how well they have learned to *think*. To accomplish this the CLA group settled on a written examination with two principal sections: one is made up of two essays, one arguing a point of view on a particular subject and the other critiquing an existing argument; another consists of a longer, "critical thinking" essay. A sample set of instructions for the latter began,

You are the assistant to Pat Williams, the president of DynaTech, a company that makes precision electronic instruments and navigational equipment. Sally

Evans, a member of DynaTech's sales force, recommended that DynaTech buy a small private plane (a SwiftAir 235) that she and other members of the sales force could use to visit customers. Pat was about to approve the purchase when there was an accident involving a SwiftAir 235.

The test taker was given newspaper articles about the accident, a federal report on small-plane in-flight breakups, charts on the performance of the SwiftAir 235, and a memo from Pat asking for recommendations on how to proceed. The instructions continued,

Please prepare a memo that addresses the questions in Pat's memo to you. Be sure to describe the data that support or refute the claim that the type of wing on the SwiftAir 235 leads to more in-flight breakups, as well as the factors that may have contributed to the accident and should be taken into account. Please also make an overall recommendation about whether DynaTech should purchase the plane and cite your reasons for this conclusion.

For the initial trials of the CLA, in 2002, fourteen unidentified colleges of various sizes supplied 1,365 student test takers who had been lured with payments of \$20 to \$25 an hour. In a series of reports available on the Council for Aid to Education Web site (www.cae.org), the CLA researchers say the test worked. College seniors had significantly higher CLA scores than freshmen with comparable SAT scores, suggesting that something that improved with college teaching had been measured; some colleges with similar SAT averages had significantly different CLA averages, suggesting that the results had something to do with the nature of education at the school. So far about fifty colleges and universities are planning to take the test this fall; the group hopes to get that number up to eighty by next spring.

The fact that the CLA shows some colleges doing better than others is, educators say, both encouraging and dangerous. CLA officials frown at the thought of ranking schools the way *U.S. News* does, even though they consider their criteria more compelling. For the moment the CLA is providing results only to the individual colleges and universities, for internal use. For example, the schools can use the results to determine how much students benefited from particular school policies. If a school could show that students' scores increased markedly during their time within its walls, it might even use the data to assist in recruiting or fundraising.

Once the CLA begins to show how much students are really learning, there may be one more job for it to do. A little paragraph tossed off at the end of a technical review of the data says that CLA scores correlated more strongly with college grades than did SAT scores. If the CLA proves successful, it's not out of the question that it could be administered to high school students and perhaps even begin to replace the SAT. If it does, you can be sure that test-preparation companies will be quick to figure out how much they can charge for ten weeks of lessons on writing Pat a dynamite memo about the company plane. —*JAY MATHEWS*