Testing and Assessment: Looking in the Wrong Places

A bunch of guys had a late afternoon game of touch football on a field. As it started getting dark and the players moved toward their cars, one realized he didn't have his car keys and must've lost them while playing. Several guys walked back to the field to search. One guy was searching at the corner, under the street lamp. They called out to him, "Hey, we were playing over here, not over there." He astutely responded, "Yeah, but the light's so much better over here."

Testing and Assessment

With a few exceptions (like applying for a driver's license), we rarely encounter paper-and-pencil tests of our knowledge once we leave school. Yet, quizzes and exams are the default method used throughout our educational career. Why do we continue to use something found nowhere else in nature as the means of measuring what our students know?

The obvious answer is because of its simplicity. Like the guy looking for the car keys, we prefer looking where it's easier rather than in the place likely to yield the best results. Someone teaching western civilization to 500 undergraduates certainly can't be expected to give an oral exam to that many students, or grade 500 oral presentations or papers, or even 125 group reports. We have a system that makes evaluation much more efficient, even though it may not be measuring exactly what we want it to measure (known by social scientists as validity). There is no shortage of research that demonstrates that the person with the highest test score may not "know" the material better than someone else who has better comprehension but suffers from text anxiety (a question of reliability for a social scientist).

Certainly, many university courses are relying on more than paper-and-pencil tests (an antiquated term, actually. These are far more likely to be completed on an electronic device), but one wonders why we do it at all. If our goal is to prepare students for the world of work, nothing they will be asked to do on the job resembles taking a test. If we have a grander purpose of educating our students for "life" (a grandiose goal that I have yet to see measured), there is no greater likelihood that they will encounter quizzes and exams. Furthermore, given the plethora of research that shows our students likely to "cram" for tests, and that studying in this way results in less long-term learning, no proponent of college as a way to grow intellectually can endorse the use of tests and exams.

As challenging as it might be, could higher education do away with quizzes and exams, at least in classes of fewer than 40 students, and survive? The simplest answer may be that we are most likely to teach the way we were taught. Most college faculty received training in research methods, but significantly fewer had any study or practice in teaching. It is logical that we fall back on the ways we were taught, largely because it's what we know. We were tested, so we must test.

After all, college admissions are substantially influenced by standardized tests such as the SAT and ACT, so testing must be the appropriate way of determining what students know, right? Graduate and professional schools have specialized tests such as the GRE, LSAT, and MCAT, so surely testing is good, right? It's not as easy as that: the use of those tests for admission has been questioned for a variety of reasons. There are arguments on both sides, but it is a fact that their relative importance in admissions decisions is less than it was in previous years. But there is an even stronger reason for testing's continued use.

As Shakespeare had Cassius say, "The fault, dear Brutus, is not in our stars, but in ourselves." Unfortunately, some in the academy would have us believe that classes without exams lack rigour. I have heard colleagues express that very sentiment. That attitude is known by pre-tenure faculty members who give exams to be positively evaluated for promotion and tenure. After seven years of giving exams, newly-tenured faculty are unlikely to make major changes to their courses, relying instead on the "time tested" techniques they have been using. After all, they were successful using those techniques, so there would be no reason to change. Note that in this example, "success" is measured by their achievement of tenure and not by the students' mastery of the content, skills, or future employment.