

## EDITOR'S PAGE

# The Challenging Task of Assessment in Higher Education

**Ronald D. Simpson**

Consider briefly these three scenarios for our time. Scenario one: A professor in an allied health area confesses after many years of teaching that the criteria used for admission of students to her college have very little to do with their ultimate performance as practitioners in their field. "The students," she says, "learn how to play the game, how to make good grades, and how to look good on paper. But they aren't really motivated to become critical thinkers or independent problem-solvers. I could do a better job selecting committed, enthusiastic, and reflective practitioners than does our mechanical and number-based system."

Scenario two: Professor John Adams is a master at documenting everything he does. He keeps records of how much time he advises students, he organizes all his student comments, and he spends considerable time packaging his teaching materials in an attractive manner. Professor Adams, however, jealously guards his time and spends little unscheduled time in his departmental office. By contrast, Professor Jane Smith practices an open-door policy, spends liberal amounts of time with students who need help, and is known as an unselfish teacher who cares deeply about students. Less flamboyant and structured than John Adams, her teaching portfolios are less impressive—influenced in large measure by a lack of motivation to chronicle her daily activities. Professor Adams won the major teaching award at his university last year and has experienced several salary adjustments over the years due in part to his thorough record-keeping capabilities.

Scenario three: Central State University has adopted a post-tenure review policy. Each tenured faculty member is now required to undergo a review process that includes peer evaluation. While the president of Central State has publically stated that the new policy is designed to improve the quality of instruction on campus, the faculty is cynical about this mandate. The sentiment is that the policy is politically motivated and demonstrates a lack of support and trust

by the governing board of the institution. A program ostensibly designed to promote excellence in teaching is being received by faculty with suspicion and defensiveness.

In each of these cases serious questions are raised regarding assessment. The validity of selected performance indicators is called into question. Powerful political implications are raised. The lack of differentiation between formative (for improvement purposes) and summative (for decision-making purposes) evaluation emerges. In effect, the age of accountability has produced challenges, with both positive and negative ramifications, that effect how people behave and feel.

There are more questions than solutions to these challenges. While assessment is complex and multifaceted, it is, nevertheless, an enterprise based on substantial theory and practice. For example, we know that assessment ultimately has to flow from a set of implicit or explicit criteria. Criteria are usually expressed in the form of standards, guidelines, or goals. Assessment also involves observation and measurement. There are many useful methods for collecting the kind of information and data needed. The third element of the assessment process involves judgment, the placing of human value and worth on the information once we have it.

In scenario one about admissions a question of validity emerges. Are the criteria being used to select students congruent with the educational outcomes of the program? In scenario one, would it help us to produce more effective health practitioners if we changed the admission criteria? Or would it be appropriate to change the curriculum and the learning environment of these students so that critical thinking and problem-solving would be better served? In any kind of assessment procedure validity is the single most important factor.

In scenario two about teaching and record keeping there is also a validity problem compounded by a political problem. By most informal measures, faculty member B has a greater impact on the students and the total department than faculty member A; but because faculty member B does not spend substantial time collecting quantifiable data, the hard evidence of her impact on student learning is not as evident as it is with faculty member A. In part, the value of Jane Smith is minimized because the leadership and culture of her department may not be sensitive to many other qualitative variables that, when taken into account, would underscore the learning outcomes of her students.

Scenario three dealing with post-tenure review suggests several political and socio-emotional factors that go well beyond the mechanical process of assessment. When members of governing boards legislate policies that imply a lack of trust toward academicians or when faculty members become defensive when the sponsoring public speaks of accountability, the entire process of assessment changes. When this happens, the positive role that assessment can play toward improvement is diminished. When a lack of trust becomes the dominant mode of thinking and behaving within an organization, the system changes. When communication breaks down and the purposes of assessment are not made clear, the entire enterprise moves in directions that are at odds with why we should evaluate in the first place.

Assessment in higher education today represents perhaps our greatest challenge. It does so because establishing uniform criteria or standards that are acceptable to all the constituents of this huge enterprise is not totally possible. In other words, maximizing quality, access and cost variables, all at the same time, is not possible. Another challenge is associated with our methodology. In order to get closer to the truth (to improve validity) we must use multiple measures. This means that robust quantitative measures as well as thoughtful and sensitive qualitative measures should be used. This means that ultimately we need to look at impact over time and make sure we are not just focusing on the things which are the easiest to measure.

The most important aspect of assessment, however, is the interpretation and judgment that follows. And with this needs to come the wisdom and experience that must accompany any activity where human value and worth is paramount. The mistakes that can be made in the business of assessment are abundant. Obviously, if the appropriate criteria are not established at the beginning, the rest of the process is jeopardized. Furthermore, if the appropriate methods of measurement are not used or the correct data are not collected, the process becomes irrational. But the largest mistake of all is failing to apply sound judgment to the information we have gathered. It is possible to have sound criteria and reliable measures but fail to be thoughtful in how this information is used for improvement purposes or to make important academic and personnel decisions.

In all three of the scenarios there was a failure to consider things that were really important. For example, it is important to understand that great teachers touch deeply the lives of their students and that this influence potentially goes on for generations. Similarly,

what makes an excellent doctor, teacher, engineer, or political leader is based on a constellation of knowledge, skills, and attitudes that may have little to do with how well one performs on a given standardized test (for example, a test that measures how many vocabulary terms can be recognized and the speed with which one can solve mathematical problems). It is also very important to understand that meaningful and lasting change within an organization occurs best when social, emotional, and political factors are treated with the same care as the administrative policies that regulate the process.

One of my favorite educators, Lee Shulman of Stanford University, tells the story of how the Chicago Health Department used to visit his parents' delicatessen. They would look for clean counters, swept floors, grease-free utensils, and roaches, then put on the wall a number signifying the rating for that period. In effect, Lee explains, this rating if at an acceptable level meant that it was safe to eat there. What really is needed, Shulman states, if one wants to learn what a restaurant is *really* like, is to read one of the many guides to fine dining. This kind of information is not a health certificate but, rather, a measure of excellence. When assessment is done poorly, it often serves as little more than a health certificate. When it is done thoughtfully, it can serve as a sensitive, refined, and civil way of promoting excellence.