Assessment and evaluation in higher education: Some concepts and principles

The ability to engage in high-quality assessment has become a sine qua non for the college-level educator. But, effective assessment requires mastering the professional knowledge and skills involved. The field of assessment and evaluation, like all other specialized disciplines, has developed many important concepts, principles, and methods to guide practice. A few of these are briefly discussed here. They should become part of the academician's professional armamentarium.

Assessment and evaluation

Assessment is a process of determining "what is." Assessment provides faculty members, administrators, trustees, and others with evidence, numerical or otherwise, from which they can develop useful information about their students, institutions, programs, and courses and also about themselves. This information can help them make effectual decisions about student learning and development, professional effectiveness, and program quality. **Evaluation** uses information based on the credible evidence generated through assessment to make judgments of relative value: the acceptability of the conditions described through assessment.

The statement "If you don't have any goals, you don't have anything to assess" expresses the close relationship between goals and effective assessment. It is goal achievement that effective assessment is generally designed to detect. An effective assessment program helps a college's or university's administrators and faculty members understand the outcomes – the results – their efforts are producing and the specific ways in which these efforts are having their effects.

Types of assessment

What is being assessed and evaluated determines the appropriate type of assessment and evaluation. For purposes of planning, desired **outcomes** (the ultimate results desired or actually achieved) as well as **processes** (the programs, services, and activities developed to produce the desired outcomes) and **inputs** (the resources: students, faculty and staff members, buildings, psychological climate) are all articulated in terms of goals and objectives. Thus, one can distinguish among outcome goals and objectives and outcome assessment and evaluation, process goals and objectives and process assessment and evaluation, input goals and objectives and input assessment and evaluation. Because it is results that count for most of higher education's stakeholders and critics, the emphasis today is on outcome assessment and evaluation.

Causation of outcomes

Outcome assessment, however, does not by itself produce enough evidence to permit thorough understanding of the behavior of an educational system. Outcome assessment indicates what results have been produced and how much of them. Clarifying causation – determining why the results were achieved – is the task of process assessment. Improving the quality of results depends upon improving the quality of processes. Thus, outcome assessment is not enough. In the case of learning and student development, a detailed understanding of the functioning of orientation, curriculum, instruction, academic advising, and other key educational processes is necessary for maximal improvement of institutional results. In other words, the results of both outcome and process assessment are needed to improve the quality of outcomes. The findings of process assessment research are interpreted in the light of empirically based higher-education theory to determine whether the processes being used can be expected to produce the outcomes desired with any particular set of students.

Input assessment is also necessary. It helps us understand our students. For example, input assessment can describe the characteristics of entering students: their various abilities as judged by placement testing and, among other important variables, the approaches they take to learning, their capacity for abstract reasoning and critical thinking, and their levels of epistemological and moral judgment development. This information gives faculty members, administrators, and others crucial information for designing programs appropriate to the developmental needs of specific kinds of students and of individual students.

Value added

Impressive assessed outcomes at the end of a curriculum or course cannot necessarily be attributed to the effects of these experiences. Pre-program student input characteristics must first be removed from student characteristics at program's end before the value added to students by the program can be determined. Moreover, there are other potential causative variables such as students' natural biological maturation and off-campus experiences that need to be considered when interpreting the results of outcome assessment. In the end, what counts is whether the program itself makes a difference and what and how much of a difference

that is.

Methodology

The methodology employed in assessment in higher education is diverse and must be appropriate to the purposes for which the assessment is being used. The student characteristics and other college inputs, the educational and other processes we use, and the multifarious outcomes we seek all require equally varied, sensitive, and appropriate methods if the results produced by assessment are to be credible and useful. Qualitative methods such as surveys, focus groups, portfolios, and direct observations are at least as important as quantitative methods, those that use and produce numerical data; and multiple rather than single methods of assessment often provide a richer and mutually corroborative array of evidence.

Whatever the methods used, two important technical qualities characterize useful assessment. **Validity** describes a condition where an assessment method, such as a paper-and-pencil test, assesses what it claims to assess and thus produces results that can lead to valid inferences usable in decision making. For example, a classroom test claiming to measure higher-order thinking skill but actually assessing only memorized knowledge lacks validity. Inferences and decisions concerning students' thinking skills cannot justifiably be made from evidence produced by such an assessment.

Reliability is the capacity of an assessment method to perform in a consistent, stable fashion during successive uses. Reliability is a prerequisite for validity. An unreliable indicator cannot produce trustworthy results.

The need for systemic thinking and for using the results of assessment

All institutional components and functions are interrelated. For example, the efficacy of a curriculum depends upon the quality of instruction in courses, and the effectiveness of both curriculum and instruction depends in turn on the quality of academic advising and the campus psychological climate. Likewise, all three types of assessment – outcome, process, input – are naturally connected to each other. Each serves a critical role in understanding certain aspects of the institution (college, program, course, students) as parts of a system. In practice, however, input assessment is often scanty, and outcome assessment is by no means always connected to process assessment such that institutional performance and quality can be understood and improved. For that matter, the results of assessment of any kind – when assessment occurs – may not be used, thus wasting valuable resources, failing to enhance institutional quality, reducing staff and student morale, and weakening institutional credibility with off-campus constituencies.

Powerful methods for understanding our students and our impact on them now exist. We must use them.

Resources for further learning

Publications

- Palomba, Catherine A., and Trudy W. Banta. 1999. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. San Francisco: Jossey-Bass. [405 pp.]
    - A general guide to assessment theory and practice for institutions and their programs.

- Gardiner, Lion F. 2000. "Monitoring and Improving Educational Quality in the Academic Department." In *Leading Academic Change: Essential Roles for Department Chairs*, Ann F. Lucas and Associates, 165-194. San Francisco: Jossey-Bass.
    - An overview of assessment and its role in leadership for the academic department.

- Gardiner, Lion F., Caitlin Anderson, and Barbara L. Cambridge, eds. 1997. *Learning through Assessment: A Resource Guide for Higher Education*. Washington, D. C.: American Association for Higher Education Assessment Forum. [116 pp.]
    - The most comprehensive directory available of assessment resources of all types.

Conference

The American Association for Higher Education presents each June what probably is the largest and most diverse conference on assessment. For further information see http://www.aahe.org.

The National Academy for Academic Leadership

Because of its essential role in ensuring academic quality in institutions, assessment occupies a central place in the curriculum of The National Academy. Learn more about the academy's programs and see additional resources on assessment.

–– Lion F. Gardiner, Rutgers University, gardiner@andromeda.rutgers.edu