Developing assessment methods at classroom, unit, and university-wide levels

Professor Trudy W Banta, Indiana University-Purdue University Indianapolis, USA

Definitions of outcomes assessment

In the early 1980s the term assessment, or more accurately, outcomes assessment, was adopted in the United States to refer to information obtained from students, graduates, and other stakeholders that may be used to improve academic programs and student services within universities. In many other countries, this process is called evaluation, or program evaluation. The term assessment is preferred in the US to distinguish the process designed to improve programs and services from evaluation, a process designed to gauge the achievements of academic staff for purposes of awarding promotions, tenure, and merit pay.

This author views outcomes assessment as a prudent step in a process that begins with planning what we wish to do. Plans are implemented and simultaneously appropriate data can be collected for use in assessing progress. If assessment findings are used to improve our processes, our plans may be adjusted, and the cycle of planning, implementing, assessing, and improving begins anew. Assessment in this context may be defined as a process of providing credible evidence of the outcomes of higher education that is undertaken for the purpose of improving programs and services within an institution. A second, simpler definition focuses squarely on the paramount college outcome, student learning. Former vice president of the American Association for Higher Education, Theodore Marchese, calls assessment “a rich conversation about student learning informed by data” (personal communication, January 7, 2004). This definition may provide the best context for the study of assessment currently underway in Scotland.

Assessment of individuals and groups

When academic staff hear the term assessment, they think most often in terms of assessing individual student development. They assess basic skills such as the ability to write, communicate orally, or use mathematics, for the purpose of advising students about appropriate placement in courses. They review student performance in their classes or modules using assignments, papers, and projects. And as students complete some programs, they are given comprehensive written and/or oral exams that test what they have learned throughout their years of study. Important outcomes of assessing individual student development include the following: (1) faculty can assign marks or grades to students, (2) students learn about their own strengths and weaknesses so that they can correct them and improve their future performance, and (3) students acquire skills in self-assessment that they can use throughout their lives. Assessment of individual student development is a critically important component of the higher education experience.

For purposes of conducting outcomes assessment, we need a second look: at aggregated student work in a class or module, in sections of the same class, and even across classes in a curriculum. Looking at student work collectively, we can tell where learning is satisfactory and where gaps in learning exist. We may also obtain some clues about which approaches to instruction produce the most learning for which students. These group assessment activities consist of classroom assignments, tests, and projects—all the same sorts of measures that are used to assess individual student development. But with group assessment we can add a variety of other measures, such as questionnaires for students, graduates, and employers. Interviews and focus groups yield helpful data. We can look at program completion data to see how many students complete our courses and curricula and how long it takes them. We can look at the placement of students in further education or careers. By tracking our graduates, we can see how successful they are in postgraduate programmes or on the job and if they have received awards...
or recognition for their performance. Finally, we can use the results of group, or outcomes, assessment to improve our programs and to demonstrate accountability to external stakeholders.

To summarise, assessment of individual student development can assist students in mastering content as well as in learning to assess their own strengths. Group, or outcomes, assessment can help faculty improve instruction and enable institutions to demonstrate their accountability.

Good assessment, or evaluation as many call it, embodies the same principles as does good research. In both we pose an important question, determine an appropriate approach to answering the question, collect data, analyse the findings, and issue a report. Assessment goes a step farther in that the findings are utilised to improve instruction in individual classrooms as well as entire academic programs and university-wide services.

Preparing academic staff to conduct assessment

Since most academic staff are not trained as teachers, faculty development is an important prerequisite for conducting good assessment. Faculty development can help instructors:

• write clear objectives for student learning in modules and curricula,
• individualise instruction using a variety of methods and materials, and
• develop assessment tools that test higher order intellectual skills.

In determining appropriate approaches to assessment, it is very helpful to write goals and objectives for student learning using action verbs. For instance, if we want students to improve their writing skills, an appropriate assessment of their progress would be a written assignment. If we want them to develop skills in locating reliable information, we could give them a project incorporating the use of such skills in order to assess their Internet search and analysis strategies.

Bloom’s Taxonomy of Educational Objectives (Bloom, 1956) consists of six increasingly complex categories that describe what Bloom has called the cognitive domain. These extend from knowledge and comprehension at the lowest level of complexity through application, analysis, synthesis, and evaluation. Action verbs may be associated with each of these levels of the domain. For instance, if we develop an objective for students using a verb such as identify, define, or describe, this learning objective is at the knowledge level. If we ask them to demonstrate, compute, or solve, students will be performing at the application level. If we expect them to criticise, compare, or conclude, the students will be developing skills at the evaluation level. In faculty development, discussing the use of verbs from the various levels of Bloom’s Taxonomy can be a helpful step in developing the ability to assess learning outcomes.

The use of action verbs in learning objectives may be illustrated more specifically as follows: If we ask a student in an English course to demonstrate how language influences intellectual and emotional responses, we are testing the student’s application skills. Synthesis skills would be illustrated in the following objective: Synthesize diverse issues and responses raised in collaborative discussions of texts. Learning outcomes in science might include the following: Define and explain basic principles, concepts, and theories of science (knowledge level); solve theoretical and experimental problems in science (application level); and evaluate scientific arguments at a level encountered by informed citizens (evaluation level).

A matrix can be useful in a number of ways in promoting conceptual thinking about assessment. A matrix format with six columns that has been used successfully at many colleges and universities in the United States is one that has as a heading for the first column, “What general outcome are you seeking (eg critical thinking)?” The second column is headed “How would you know it (the outcome) if you saw it—that is, what would the student know or be able to do?” The third column heading is “How will you help students learn the concept, in class or out of class?” And the fourth heading is “How could you measure each of the desired behaviours listed in
column 2?" The fifth column heading reads “What are the assessment findings?” And the sixth asks “What improvements are or might be based on assessment findings?” Completing such a matrix can enable faculty to explain to students and other stakeholders (1) specific learning outcomes of a module or a course of study, (2) collective student outcomes, and (3) actions undertaken to improve student learning based on assessment findings.

**Classroom, unit, and university-wide levels of assessment**

Outcomes assessment occurs at a number of levels. It begins with the individual student in a classroom. Aggregating the work of all students in a classroom will provide information to inform classroom assessment. Aggregating student work across various classes or modules can provide assessment (evaluation) of the impact on learning of an entire course of study. Looking at student products across the disciplines in a college provides assessment at that level. Assessment findings from various academic units within a university can provide a measure of institutional effectiveness that can be used to demonstrate accountability at the state, regional, or national level.

A distinction must be drawn between direct and indirect measures of student learning. Direct measures are those assignments, exams, projects, and papers that enable us to see what students actually know and can do. Indirect measures include questionnaires, interviews, and focus groups that enable us to assess the process of learning or other aspects of the student experience. Direct measures of learning are critical if we are to assess acquisition of knowledge and skills. But no test score will tell us why certain components of students’ knowledge are strong or weak. Thus indirect measures are needed to help us understand why weaknesses are occurring and what might be done to address them. Good assessment includes both direct and indirect measures.

Citing some examples of assessment at various levels may add clarity to this concept. Fast feedback, or classroom assessment, can be used at the individual classroom level. Students are asked during the last five minutes of a classroom session to state the most important thing they learned in the class that day and to tell the instructor what is still unclear. Then they may be asked about the helpfulness of the advance reading assignments for the day’s work. Finally, they may be asked for suggestions for improving the class and/or the assignments. In an illustration from the Graduate School of Business at the University of Chicago, students responded to the last question in that sequence by suggesting the following improvements: (1) install a portable microphone, (2) increase the type size on transparencies, (3) leave lights on when using a projector, (4) don’t cover the assigned reading in great detail, but instead (6) provide more examples from actual practice in class lectures and discussion (Bateman and Roberts, 1993).

We can adapt the typical course evaluation to include questions about the student experience. Are students encountering in the course principles of good practice in undergraduate education (Chickering and Gamson, 1987)? We might ask, for instance, if in a given module or in an entire curriculum (1) learners held high expectations for one another, (2) learners interacted frequently with academic staff in and outside class, (3) learners participated in learning teams, (4) learners respected diverse talents and ways of learning (Cournoyer, 2001).

**Primary trait scoring**

Primary trait scoring is an assessment method that can be used in both direct and indirect measures, and at all levels (Walvoord and Anderson, 1998). Instructors identify the traits or attributes that are necessary for success in an assignment, then compose a scale or rubric that gives clear definition to each point, and finally evaluate student work according to the rubric. For example, a project that involves developing and presenting a research paper encompasses at least the following primary traits: (1) an appropriately narrow topic or purpose, (2) a bibliography, (3) an outline, (4) a first draft, (5) a final draft, and (6) an oral defence. For each of the traits of
this assignment we might develop a three-point rubric, defining each point carefully and explicitly. The bibliography, for instance, might be assessed as follows:

3  (Outstanding) References current, appropriately cited, representative and relevant
2  (Acceptable) References mostly current, few citation errors, coverage adequate, mostly relevant
1  (Unacceptable) No references or containing many errors in citation format, inadequate coverage, or irrelevant

If one creates a matrix containing the primary traits of an assignment as row titles and the levels of each rubric as column headings, such a matrix can serve three purposes. First, it can be shared with students prior to an assignment so that they will understand the criteria being used to judge their work. Second, it can be completed for each student on the basis of the work submitted and thus provide detailed feedback when returned to the student. Third, if the instructor places a check mark in the appropriate box of the matrix for every mark assigned in evaluating the work of all students, the matrix can indicate to the instructor where there are weaknesses in student learning and suggest what changes may need to be made to enable every student to reach the desired learning outcomes.

Another matrix might list principal outcomes as row titles and courses in a curriculum as column headings. Placing check marks in the matrix to demonstrate which outcomes each course addresses will help students understand where they will learn specified knowledge and will assist instructors in spotting gaps in the curriculum.

Primary trait scoring can be used in virtually any field. For instance, at Ball State University in Indiana, sophomore competence in mathematics was tested. Students were asked to turn in their supporting work in connection with their item responses on a math test. Then instructors used a four-point scale to score responses in terms of conceptual understanding, consistent notation, logical formulation, and completeness of the solution (Emert and Parish, 1996).

At North Dakota State University faculty in sociology and anthropology developed scenarios appropriate to the discipline, then asked graduating students to respond to the scenarios in groups (Murphy and Gerst, 1997). A faculty facilitator asked questions related to outcomes faculty had identified in three areas—concepts, theory, and methods. Then two faculty observing the group work used a 0-3 scale to rate each student on each question. Looking at aggregate scores across all student groups enabled faculty working together to ascertain strengths and weaknesses of their curriculum.

Group interaction also can be assessed using primary traits and scoring rubrics. Faculty at the Purdue University College of Pharmacy in Indiana developed a five-point scale ranging from 5 = consistently excellent to 1 = inconsistent and/or inappropriate to judge the performance of students working in groups (Chalmers and Mason, 1994). The characteristics faculty were observing included the following:

- listened to others
- actively contributed to discussion
- challenged others effectively
- was willing to alter own opinion
- effectively explained concepts/insights
- summarised proposed solutions
Involving stakeholders in assessment

In a comprehensive assessment program, it is important to involve as many stakeholders as possible. Such groups include students, faculty, student affairs professionals, administrators, graduates, and employers.

An example of involving students may be drawn from the experience of the Department of Psychology at Montevallo University (Judith Rogers, personal communication, May 4, 1994). Students were asked to form an advisory council and to provide continuous assessment of the psychology major. Students made a number of important recommendations for improving their program. They asked that a statement of expected ethical behaviors for students be drawn up and volunteered to do this themselves. They suggested that a second research course with a laboratory be added to the curriculum because they felt they needed more research experience. They asked for more comparative psychology; additional terminals for the statistics lab; and more opportunities in all their courses to write, make oral presentations, and conduct research projects.

At Harvard University, Richard Light (1992) has involved students in teams to evaluate both courses and programs. In addition to the good information about the strengths and weaknesses of courses and curricula that the students provide to academic staff, the students experience learning benefits as they engage in the assessment process. As they interact with their peers, they strengthen their communication skills and develop competence in working in a team. Their self-confidence increases and they improve their understanding of others’ perspectives. Finally, student involvement in learning increases.

At Indiana University Purdue University Indianapolis, faculty across the institution have agreed on student outcomes in six areas, including:

1. core communication and quantitative skills
2. critical thinking
3. integration and application of knowledge
4. intellectual depth, breadth, and adaptiveness
5. understanding society and culture
6. values and ethics.

Efforts are underway to develop an electronic portfolio that will give students the responsibility for demonstrating their skills in connection with each of the six Principles of Undergraduate Learning. The electronic format will give students the opportunity to use multiple media to demonstrate unique individual skills and achievements. Written reflections on the material they put into the portfolio will enable students to enhance their metacognitive skills.

Involving student services professionals with faculty in the assessment of learning is illustrated in an experience at Virginia Commonwealth University (Fuhrmann, 1995). There first-year students in English composition wrote a theme each week for 15 weeks. The subject of each theme was the student’s experience with, and opinions of, some aspect of campus life, such as the admissions process, the new-student orientation program, or academic advising. Teams of faculty and student affairs professionals read and evaluated the essays. Faculty from the Department of English assessed the student work in terms of the quality of the writing, while student affairs professionals gained valuable information about students’ perceptions of the quality of various campus programs and services. As a result of this team approach to assessment, a number of changes have been made at Virginia Commonwealth, including provision of a new central advising center, a revised advising handbook, and enhanced multicultural workshops.

Kretovics and McCambridge (1999) at Colorado State University provide an example of involving employers in assessment. They have used a combination of surveys and focus groups for employers of business graduates to identify skills, knowledge, and personality attributes that are
sought by employers. The findings yielded by these assessment techniques have encouraged faculty to make curriculum changes; motivated students to develop needed skills; and strengthened ties among faculty, students, and employers. Faculty have added a credit to the business communications course that increases time for students to work in teams and make more oral presentations. In addition, business ethics and social responsibility are now discussed in introductory courses, and a new Introduction to Business course emphasizes career decision-making.

**Using assessment findings**

The time and energy devoted to outcomes assessment can only be justified if the findings are used to improve student learning and development. A few examples illustrate such uses. At Columbia College, academic staff in social work wanted students to apply critical thinking skills in their clinical practice settings (Baskind, Shank, and Ferraro, 2001). More specifically, faculty specified that students would use statistical analysis to inform practice decisions. Laboratory exercises were used to assess the skill, and faculty set as a standard that 80% of all students would earn a score of at least 3.0 on a 4-point scale on their lab exercises. Columbia College faculty found that only 64% of the social work students scored 3.0 or better. By asking students what would help them most, faculty learned that students needed more time in the laboratory—more time to practice their skills. In response, faculty made the learning laboratory available to students five days a week.

At Southern Illinois University-Edwardsville, for more than a decade faculty have required every student to undertake and pass a "senior assignment" in their discipline (Eder, 2001). In business, the senior assignment is a case study analysis. In education, students must complete a professional portfolio. In psychology, students must develop a poster presentation based on a research project. In engineering, there is a senior design project, and in nursing, students must develop a plan of care for a patient. At the end of each academic year, faculty meet to go over the senior assignments for that year and draw conclusions about strengths and weaknesses in students' knowledge and skills. Over the years, each discipline has made changes based on the evidence provided in the student work. In business, more case studies and research projects have been required at earlier levels of the curriculum. In education, more practice in classroom management has been offered to improve student learning. In psychology the statistics requirement has been changed. In engineering, students have been given more opportunities to practice their writing and speaking skills throughout the curriculum. And in nursing, an expensive simulation lab with computerized patients has been purchased so that students may practice their clinical and diagnostic skills without harming a patient!

At Eastern New Mexico University, portfolio review and/or an audition is conducted for every fine arts student every semester by a panel that includes faculty, students, community representatives, and professional staff or faculty outside fine arts (Alec Testa, personal communication, April 27, 2004). As a result of this intensive assessment process, changes have been made in instruction and programming and now it is possible to see some results of those changes. Faculty have observed that students' creativity, conceptualization, and technique have improved.

**Characteristics of effective assessment**

Over the years since 1980 assessment practitioners have developed a number of characteristics that define effective assessment (Palomba and Banta, 1999). First, there should be an overall plan for assessment, suggesting where and when students will be evaluated, the evaluation instruments that will be used, and what will be done to respond to the data collected. As indicated earlier, all stakeholders in higher education should be drawn into the assessment process. For instance, students and employers, as well as faculty, can help to set goals and objectives for courses and curricula, select or design assessment methods, and even collect
some of the data. Assessment findings then should be reported to all stakeholders and their use encouraged. Finally, the assessment program itself should be evaluated periodically, preferably through peer review.

Successful assessment initiatives are led by committed individuals and include collaboration between faculty and student affairs leaders. Assessment depends upon effective teamwork in planning, implementation, and utilization of findings. Assessment thrives in a supportive campus climate where faculty and administrators demonstrate their concern for students and their commitment to continuous improvement. Once assessment results are reported, campus leaders must follow up periodically to make sure that the findings are being used to improve practice.

If the experience of the last quarter-century is any guide, the need for providing evidence of accountability in higher education will not diminish, but rather will increase, as it has virtually every year since 1980. As scholarship reveals how assessment can improve instruction and student learning, more faculty will realize its benefits. Additional electronic assessment methods will be developed—already this is occurring at a rapid pace. More sharing of assessment methods will take place among groups of faculty. And ultimately, assessment will enable faculty to gain a deeper understanding of student learning and student learning will improve as assessment findings are used to effect continuous improvement of the environment for learning.

References


IUPUI Principles of Undergraduate Learning http://www.iport.iupui.edu