STUDENT LEARNING ASSESSMENT

Options and Resources

SECOND EDITION
There are many ways to approach the evaluation of student learning. The characteristics of good evidence of student learning include considerations of direct and indirect methods for gathering evidence of student learning, the appropriate use of quantitative and qualitative evidence, and other methodological considerations. First, however, it is important to understand the fundamental assessment concepts of formative and summative assessment and benchmarking.

Formative and Summative Assessment

Formative assessment is ongoing assessment that is intended to improve an individual student’s performance, student learning outcomes at the course or program level, or overall institutional effectiveness. By its nature, formative assessment is used internally, primarily by those responsible for teaching a course or developing a program. Ideally, formative assessment allows a professor, professional staff member, or program director to act quickly to adjust the contents or approach of a course or program. For instance, a faculty member might revise his or her next unit after reviewing students’ performance on an examination at the end of the first unit, rather than simply forging ahead with the pre-designated contents of the course. Many instructors also solicit repeated brief evaluations of their teaching, and the data gleaned from these can be used to make adjustments that may improve learning, such as the introduction of more discussion into a class.

In contrast, summative assessment occurs at the end of a unit, course, or program. The purposes of this type of assessment are to determine whether or not overall goals have been achieved and to provide information on performance for an individual student or statistics about a course or program for internal or external accountability purposes. Grades are the most common form of summative assessment.

Goals for student learning will be expressed summatively when faculty members are describing what they expect students to be able to do or what skills they expect students to achieve when they complete a course or a program or when they graduate from the institution.

Formative and summative assessment work together to improve learning. They should be central components of assessment at the course level, and where appropriate, at the program level.

Benchmarking

The term benchmarking is now common in assessment plans and conversations about assessment. Originally, benchmarking was a term used in the corporate environment to define a set of external standards against which an organization could measure itself. The organization identifies comparable, peer, or “reach” organizations and systematically compares its practices or achievements against those of the other organization.

In higher education settings, a university might use benchmarking techniques to define its comparison group—its peer institutions—and to compare its
own outcomes to theirs. This benchmarking could be based, for example, on retention rates, five-year graduation rates, admissions yield data (the number of enrollees as a function of the number of students accepted), employment and graduate school placement rates, and performance on national or professional examinations. Theoretically, any outcome for which there are data from peer institutions and programs can be compared in a benchmarking study.

Two other related forms of benchmarking are used in higher education settings. A college or university can compare itself to a national norm by reviewing the data from a published test or survey such as the National Survey of Student Engagement (NSSE). Alternatively or in addition, an institution can set for itself the goals or benchmarks that it hopes to achieve within a specified time period (e.g., to increase job placement rates from 70% to 90% in five years).

The benefit of inter-institutional comparison is that it can flag problem areas to investigate the causes of results that differ from the norm. For example, two comparable liberal arts colleges with similar selectivity, similar student preparedness, similar socioeconomic profiles for their students, and similar science curricula, may discover that proportionately more students are accepted to medical schools from one institution than from another. Further investigation may reveal that the excelling college requires a hospital internship for all of its pre-med students.

The discovery that an institution’s students are below the norm on a national metric (e.g., amount of time devoted to school work outside the classroom) challenges the institution to determine the reason for this result. Similarly, an institution that sets its own internal benchmarks must design and implement a program to achieve its goals.

Before beginning to articulate goals for student learning, program faculty and leaders of institutional assessment should consider how the use of benchmarks could enhance their institution’s ability to achieve its goals and whether useful measures from comparable peer institutions are available.

---

**Direct and Indirect Methods for Assessing Student Learning**

The concepts of direct and indirect methods of evaluating student learning are often confused with each other and with quantitative and qualitative forms of information. Each of these has its merits and drawbacks.

Direct and indirect methods of evaluating learning relate to whether or not the method provides evidence in the form of student products or performances. Such evidence demonstrates that actual learning has occurred relating to a specific content or skill. Indirect methods reveal characteristics associated with learning, but they only imply that learning has occurred. These characteristics may relate to the student, such as perceptions of student learning, or they may relate to the institution, such as graduation rates.

When a student completes a calculus problem correctly and shows her work, learning is demonstrated directly. When the same student describes her own calculus abilities as excellent, she is demonstrating indirectly that she has learned calculus. Both of these pieces of information about the student’s performance are important. For example, a student’s perception that she is doing poorly in calculus when she is actually doing well would provide important information to both the student and the professor. However, indirect evidence—in this case, a perception—is less meaningful without the associated direct and tangible evidence of learning.

Figure 5 includes examples of direct and indirect measures of student learning at the course, program, and institutional levels. Many of the examples presented in Figure 5 can be used as measures of student learning at more than one level. For example, portfolios of student work and student satisfaction surveys can be used at the course, program, and institutional level, and internship performance ratings could be used at the course or program level.
Figure 5
Examples of Direct and Indirect Measures of Student Learning (Course, Program, and Institutional Levels)

<table>
<thead>
<tr>
<th></th>
<th>Direct Measures</th>
<th>Indirect Measures</th>
</tr>
</thead>
</table>
| **Course**       | - Course and homework assignments  
- Examinations and quizzes  
- Standardized tests  
- Term papers and reports  
- Observations of field work, internship performance, service learning, or clinical experiences  
- Research projects  
- Class discussion participation  
- Case study analysis  
- Rubric (a criterion-based rating scale) scores for writing, oral presentations, and performances  
- Artistic performances and products  
- Grades that are based on explicit criteria related to clear learning goals | - Course evaluations  
- Test blueprints (outlines of the concepts and skills covered on tests)  
- Percent of class time spent in active learning  
- Number of student hours spent on service learning  
- Number of student hours spent on homework  
- Number of student hours spent at intellectual or cultural activities related to the course  
- Grades that are not based on explicit criteria related to clear learning goals |
| **Program**      | - Capstone projects, senior theses, exhibits, or performances  
- Pass rates or scores on licensure, certification, or subject area tests  
- Student publications or conference presentations  
- Employer and internship supervisor ratings of students’ performance | - Focus group interviews with students, faculty members, or employers  
- Registration or course enrollment information  
- Department or program review data  
- Job placement  
- Employer or alumni surveys  
- Student perception surveys  
- Proportion of upper-level courses compared to the same program at other institutions  
- Graduate school placement rates |
| **Institutional**| - Performance on tests of writing, critical thinking, or general knowledge  
- Rubric (criterion-based rating scale) scores for class assignments in General Education, interdisciplinary core courses, or other courses required of all students  
- Performance on achievement tests  
- Explicit self-reflections on what students have learned related to institutional programs such as service learning (e.g., asking students to name the three most important things they have learned in a program) | - Locally-developed, commercial, or national surveys of student perceptions or self-report of activities (e.g., National Survey of Student Engagement)  
- Transcript studies that examine patterns and trends of course selection and grading  
- Annual reports including institutional benchmarks, such as graduation and retention rates, grade point averages of graduates, etc. |