

Document source:

http://www.umdnj.edu/idsweb/idst5350/assess_eval_test_grade.htm

Assessment, Evaluation, Testing and Grading

Craig L. Scanlan, EdD, RRT

What is Assessment?

To many teachers (and students), “assessment” simply means giving students tests and assigning them grades. This conception of assessment is not only limited, but also limiting (see section below on Assessment versus grading). It fails to take into account both the utility of assessment and its importance in the teaching/learning process.

In the most general sense, *assessment* is the process of making a judgment or measurement of worth of an entity (e.g., person, process, or program). *Educational assessment* involves gathering and evaluating data evolving from planned learning activities or programs. This form of assessment is often referred to as *evaluation* (see section below on Assessment versus Evaluation). *Learner assessment* represents a particular type of educational assessment normally conducted by teachers and designed to serve several related purpose (Brissenden and Slater, n.d.). These purposes include:

- motivating and directing learning
- providing feedback to student on their performance
- providing feedback on instruction and/or the curriculum
- ensuring standards of progression are met

Learner assessment is best conceived as a form of two-way communication in which feedback on the educational process or product is provided to its key stakeholders (McAlpine, 2002). Specifically, learner assessment involves communication to **teachers** (feedback on teaching); **students** (feedback on learning); **curriculum designers** (feedback on curriculum) and **administrators** (feedback on use of resources).

For teachers and curriculum/course designers, carefully constructed learner assessment techniques can help determining whether or not the stated goals are being achieved. According to Brissenden and Slater (n.d.), classroom assessment can help teachers answer the following specific questions:

- To what extent are my students achieving the stated goals?
- How should I allocate class time for the current topic?
- Can I teach this topic in a more efficient or effective way?
- What parts of this course/unit are my students finding most valuable?
- How will I change this course/unit the next time I teach it?
- Which grades do I assign my students?

For students, learner assessment answers a different set of questions (Brissenden and Slater, n.d.):

- Do I know what my instructor thinks is most important?
- Am I mastering the course content?
- How can I improve the way I study in this course?
- What grade am I earning in this course?

Why Assessment is Important

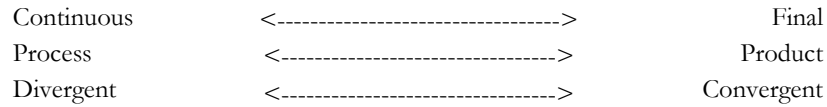
First and foremost, *assessment is important because it drives students learning* (Brissenden and Slater, n.d.). Whether we like it or not, most students tend to focus their energies on the best or most expeditious way to pass their ‘tests.’ Based on this knowledge, we can use our assessment strategies to manipulate the kinds of learning that takes place. For example, assessment strategies that focus predominantly on recall of knowledge will likely promote superficial learning. On the other hand, if we choose assessment strategies that demand critical thinking or creative problem-solving, we are likely to realize a higher level of student performance or achievement. In addition, good assessment can help students become more effective self-directed learners (Angelo and Cross, 1993).

As indicated above, motivating and directing learning is only one purpose of assessment. Well-designed assessment strategies also play a critical role in educational decision-making and are a vital component of ongoing quality improvement processes at the lesson, course and/or curriculum level.

Types and Approaches to Assessment

Numerous terms are used to describe different types and approaches to learner assessment. Although somewhat arbitrary, it is useful to these various terms as representing dichotomous poles (McAlpine, 2002).

Formative	<----->	Summative
Informal	<----->	Formal



Formative vs. Summative Assessment

Formative assessment is designed to assist the learning process by providing feedback to the learner, which can be used to identify strengths and weakness and hence improve future performance. Formative assessment is most appropriate where the results are to be used internally by those involved in the learning process (students, teachers, curriculum developers).

Summative assessment is used primarily to make decisions for grading or determine readiness for progression. Typically summative assessment occurs at the end of an educational activity and is designed to judge the learner’s overall performance. In addition to providing the basis for grade assignment, summative assessment is used to communicate students’ abilities to external stakeholders, e.g., administrators and employers.

Informal vs. Formal Assessment

With informal assessment, the judgments are integrated with other tasks, e.g., lecturer feedback on the answer to a question or preceptor feedback provided while performing a bedside procedure. Informal assessment is most often used to provide formative feedback. As such, it tends to be less threatening and thus less stressful to the student. However, informal feedback is prone to high subjectivity or bias.

Formal assessment occurs when students are aware that the task that they are doing is for assessment purposes, e.g., a written examination or OSCE. Most formal assessments also are summative in nature and thus tend to have greater motivation impact and are associated with increased stress. Given their role in decision-making, formal assessments should be held to higher standards of reliability and validity than informal assessments.

Continuous vs. Final Assessment

Continuous assessment occurs throughout a learning experience (intermittent is probably a more realistic term). Continuous assessment is most appropriate when student and/or instructor knowledge of progress or achievement is needed to determine the subsequent progression or sequence of activities. Continuous assessment provides both students and teachers with the information needed to improve teaching and learning *in process*. Obviously, continuous assessment involves increased effort for both teacher and student.

Final (or terminal) assessment is that which takes place only at the end of a learning activity. It is most appropriate when learning can only be assessed as a complete whole rather than as constituent parts. Typically, final assessment is used for summative decision-making. Obviously, due to its timing, final assessment cannot be used for formative purposes.

Process vs. Product Assessment

Process assessment focuses on the steps or procedures underlying a particular ability or task, i.e., the cognitive steps in performing a mathematical operation or the procedure involved in analyzing a blood sample. Because it provides more detailed information, process assessment is most useful when a student is learning a new skill and for providing formative feedback to assist in improving performance.

Product assessment focuses on evaluating the result or outcome of a process. Using the above examples, we would focus on the answer to the math computation or the accuracy of the blood test results. Product assessment is most appropriate for documenting proficiency or competency in a given skill, i.e., for summative purposes. In general, product assessments are easier to create than process assessments, requiring only a specification of the attributes of the final product.

Divergent vs. Convergent Assessment

Divergent assessments are those for which a range of answers or solutions might be considered correct. Examples include essay tests, and solutions to the typical types of indeterminate problems posed in PBL. Divergent assessments tend to be more authentic and most appropriate in evaluating higher cognitive skills. However, these types of assessment are often time consuming to evaluate and the resulting judgments often exhibit poor reliability.

A convergent assessment has only one correct response (per item). Objective test items are the best example and demonstrate the value of this approach in assessing knowledge. Obviously, convergent assessments are easier to evaluate or score than divergent assessments. Unfortunately, this “ease of use” often leads to their widespread application of this approach even when contrary to good assessment practices. Specifically, the familiarity and ease with which convergent assessment tools can be applied leads to two common evaluation fallacies: the Fallacy of False Quantification (the tendency to focus on what’s easiest to measure) and the Law of the Instrument Fallacy (molding the evaluation problem to fit the tool).

Assessment versus Evaluation

Depending on the authority or dictionary consulted, assessment and evaluation may be treated as synonyms or as distinctly different concepts. As noted above, if a distinction exists, it probably involves what is being measured and why and how the measurements are made. In terms of what, it is often said that we assess students and we evaluate instruction. This distinction derives from the use of *evaluation research* methods to make judgments about the worth of educational activities. Moreover, it emphasizes an individual focus of assessment, i.e., using information to help identify a learner's needs and document his or her progress toward meeting goals.

In terms of why and how the measurements are made, the following table (Apple & Krumsieg, 1998) compares and contrasts assessment and evaluation on several important dimension, some of which were previously defined.

<i>Dimension</i>	<i>Assessment</i>	<i>Evaluation</i>
Timing	Formative	Summative
Focus of Measurement	Process-Oriented	Product-Oriented
Relationship Between Administrator and Recipient	Reflective	Prescriptive
Findings and Uses	Diagnostic	Judgmental
Modifiability of Criteria, Measures	Flexible	Fixed
Standards of Measurement	Absolute (Individual)	Comparative
Relation Between Objects of A/E	Cooperative	Competitive

From: Apple, D.K. & Krumsieg, K. (1998). *Process education teaching institute handbook*. Pacific Crest

The bottom line? Given the different meaning ascribed to these terms by some educators, it is probably best that whenever you use these terms, you make your definitions clear.

Assessment versus Grading

Based on the above discussion, grading could be considered a component of assessment, i.e., a formal, summative, final and product-oriented judgment of overall quality of worth of a student's performance or achievement in a particular educational activity, e.g., a course. Generally, grading also employs a comparative standard of measurement and sets up a competitive relationship between those receiving the grades. Most proponents of assessment, however, would argue that grading and assessment are two different things, or at least opposite pole on the evaluation spectrum. For them, assessment measures student growth and progress on an individual basis, emphasizing informal, formative, process-oriented reflective feedback and communication between student and teacher. Ultimately, which conception you supports probably depends more on your teaching philosophy than anything else.

Bibliography

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. San Francisco: Jossey-Bass.

Apple, D.K. & Krumsieg, K. (1998). *Process education teaching institute handbook*. Corvallis, OR: Pacific Crest Software.

Brissenden, G., & Slater, T. Assessment primer. In College Level One (CL-1) Team. *Field-tested learning assessment guide*. Available at <http://www.flaguide.org>.

Linn, R. L. (1995). *Measurement and assessment in teaching* (7th ed.). Englewood Cliffs, NJ: Merrill.

McAlpine, M. (2002). *Principles of assessment*. Glasgow: University of Glasgow, Robert Clark Center for Technological Education. Available at: <http://www.caacentre.ac.uk/dldocs/Bluepaper1.pdf>

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey-Bass.

Wass, V., Van der Vleuten, J., & Shatzer, R.J. (2001). Assessment of clinical competence. *The Lancet*, 357, 945-949.