Grading vs. Assessment of Learning Outcomes: What's the difference?

There is often confusion over the difference between grades and learning assessment, with some believing that they are totally unrelated and others thinking they are one and the same. The truth is, it depends. Grades are often based on more than learning outcomes. Instructors' grading criteria often include behaviors or activities that are not measures of learning outcomes, such as attendance, participation, improvement, or effort. Although these may be correlated with learning outcomes, and can be valued aspects of the course, typically they are not measures of learning outcomes themselves.[1]

However, assessment of learning can and should rely on or relate to grades, and so far as they do, grades can be a major source of data for assessment. To use grades as the basis for learning outcomes, grades would first have to be decomposed into the components that are indicators of learning outcomes and those that are indicators of other behaviors. Second, grades would have to be based on clearly articulated criteria that are consistently applied. Third, separate grades or subscores would have to be computed for the major components of knowledge and skills so that evidence of students' specific areas of strength and weakness could be identified. For example, although 30% of a class may receive a grade of B, the group may all have shown a very high level of competence on one skill set but only moderate achievement in another. This kind of strength and weakness assessment provides feedback that is useful to students because it can guide and focus their practice, to the instructor, because it can reveal topics and skills that require further instructional activities, and to the department, because it can guide potential changes in curriculum to appropriately address areas of strength and weakness.

This kind of analysis is not the same as producing sub scores for different course activities, such as a score for homework, one for exams, and another for projects. These are different methods of assessment, and each of them may assess multiple skills and abilities and may overlap with each other in terms of what knowledge and skills they assess. To accurately assess learning outcomes, each type of assessment (i.e., exam, project, programming assignment, etc), would need to be analyzed in terms of the different skills it addresses and scores across the various types of assessment activity would have to be compiled and assigned for each of the skills.

For example:

Items 1, 4, 5 and 9 on an exam and homework 2 might all deal with the ability to identify the appropriate strategy to apply in a given problem context. The combined score from those items would comprise the "identify solution strategy" score.

Many instructors already have this information to some degree but discard it when computing overall grades. Questions or problems on exams or homework are individually scored already. To then turn these scores into an assessment of student learning one would only have to classifying the problem

according to the skill (or learning objective) it addresses, and then compute separate totals for each different category.

Skill/ objective	Identify solution strategy (Exam 1) 15pt	Identify solution strategy (HW2) 20pt	Implement solution strategy (Exam 2) 10 pt	Implement solution strategy (HW 3) 10 pt	Total Identify strategy 35 pt	Total Implement strategy 20 pt	Total Score/ Grade 55pts
Std 1	10	15	6	9	25	15	40=C
Stdt 2	14	18	10	8	32	18	50=B
Std 3	14	17	5	5	31	10	41=C
Class Average	**12.67**	**16.67**	**7**	**7.33**	**29.34**	**14.33**	

In this example, two skills have been scored separately on four assessments: Exam 1, HW2, Exam2 and HW3. According to grade assignment, Student 1 and Student 3 obtain the same grade but in terms of their learning outcomes, it is clear that the students learning outcomes are very different, with student 1 demonstrating weakness in "identification" and "implementation" and student 3 demonstrating strength in "identification" and serious weakness in "implementation". Thus the grade alone does not identify for the student or the instructor which component skills the student has mastered. Furthermore, overall grades would not provide an instructor with feedback on which skills the class overall found difficult. The above representation would provide class level analysis of strengths and weaknesses. This kind of feedback could be used by individual instructors to target changes in instruction or by departments for curriculum revision. By continually monitoring learning outcomes the instructor could then track the impact of instructional or curricular changes on specific learning outcomes.

This model of assessment puts the responsibility for the design, implementation, and interpretation of the assessment in the hands of the instructors or faculty members, since they alone have the content expertise as well as the knowledge of the course learning goals and assessment methods and materials. Furthermore, only the course instructors, or the department faculty as a group, can decide on the appropriate standards or criteria to classify a learning outcome as Exceptional, Very good, Good, or Substandard. However, when such an assessment of student learning is combined with an explicit statement of the learning objectives and the standards of performance, then external groups (e.g., college deans, provost, external accreditation agencies, similar departments in other institutions, employers, etc.) would be able to evaluate the degree to which students, both individually and as group, have achieved the desired learning outcomes.

Assessment, Evaluation, Testing and Grading

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
Continuous	<-------------------------------->	Final
Process	<-------------------------------->	Product
Divergent	<-------------------------------->	Convergent

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 product 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 and Grading

Assessment Is More than Just Grading

The words "assessment" and "grading" are sometimes used interchangeably, but it is helpful to distinguish between them.

Assessment is something you do every day as you gauge where students are in the learning process. You are assessing your students when you ask them questions, read their homework, and listen to their mathematical conversations. These assessments guide your instructional decisions regarding pacing, teaching strategies, and "where to go from here." Getting as accurate a reading as possible requires that students be observed and assessed in real situations; hence the term authentic assessment, which is used frequently in educational reform.

Assessment should be part of the ongoing educational process and should enhance learning. Unlike the standardized tests, which create a break in learning in order to take a measurement, assessment should be part of the natural flow of the classroom. When the curriculum provides a window into a student's thinking, that is a natural time to assess that student. Such an assessment need not be something you assign a specific grade to--it may be simply for informational purposes, both for you and for the student.

How Does Grading Fit In?

In a sense, grading is one of the by-products of assessment. As teachers, we have the responsibility of assigning each student a grade periodically throughout the year. Somehow, you must determine a grade--usually a single letter or number--to reflect all of a student's performance in one lump sum. Not a simple task!

The first step in deciding of how to grade your IMP [Interactive Math Program] students is to sit down and decide what you really value in your IMP classroom. Some of the following may come to mind:

- Completion of homework
- Group and class participation
- Progress in the concepts and skills of the unit
- Mathematical communication through written work and oral presentations

These are broad goals; your task is to construct a grading scheme that reflects your priorities.

[...]

Whatever system you use, it is vital that your students be informed about the grading process. Students should know where their grade is coming from. They should know what is valued and should have ways to participate in the process.

Steps to Developing an Assessment Plan
http://www4.wccnet.edu/departments/curriculum/assessment.php?levelone=steps

There are several steps to follow in creating a new or improved assessment plan:

1. *Examine the learning outcomes* that have been outlined by the department for the course or program to be assessed (these should taken directly from the master syllabus or program approval documents).

2. *Identify those learning outcomes that will be assessed.* Rather than attempting to assess all the learning outcomes on the course syllabus, choose those that seem most critical to the overall goals of the course which can be meaningfully measured.

3. *Select your assessment tool(s): methods or instruments for gathering evidence to show whether students have achieved the expected learning outcomes.* Determine if there are existing data sources or tools that may be used or if new tools must be selected or developed. Select those tools that seem most appropriate to the learning objectives and student population being assessed. For sample tools, see Sample Assessment Methods.

4. *Select the student population to be assessed* (e.g. random sample of sections, graduating student, etc.).

5. *Specify procedures for analyzing and interpreting the evidence gathered in assessment.* Prior to administering assessments, create a scoring rubric or other method of evaluating results, and determine the departmental standard for performance expectations (e.g. success equals 75% of students meeting outcomes). Determine if the assessment will be episodic (a snapshot of student performance at one point in time) or ongoing (a recurring, consistent, and comparative assessment of student achievement over time).

6. *Determine how the information that results from assessment can be used for decision making, planning, and course/program evaluation and improvement.* Develop means whereby involved faculty can review the data, make recommendations for change as appropriate, and incorporate such changes in the unit's planning cycle. Report findings to the area Dean and the Assessment Steering Committee, and include them when submitting curricular changes to Curriculum Committee.

Sample Assessment Methods

The following are just some of the potential assessment methods that might be used in assessing student academic achievement in courses and/or programs. It is always recommended that you choose the assessment method(s) that seem most appropriate for the learning outcomes you wish to assess.

Capstone experience
A project or activity in which students demonstrate achievement of comprehensive learning outcomes that is usually completed at the end of a course or program. In a program, there may be a capstone course that includes the pertinent learning outcomes for the program.

Departmental exam
A common exam developed collaboratively by a department used in all sections being assessed; may be part of a graded final that is evaluated separately using a scoring rubric.

External certification/licensure exam
Exams developed by regional or national accrediting or licensing organizations to evaluate students on specific skills usually related to an occupational area, such as nursing or automotive technology.

Externally evaluated job performance
Evaluation of student competence, knowledge and skills by an employer in and internship, coop, or job placement. Useful for program assessment in occupational areas.

Externally evaluated performance or exhibit
Useful in the visual and performing arts, a performance or exhibit that is evaluated or judged by experts in the field other than the instructor for the assessed course/program. The external evaluator may be an instructor at WCC who teaches a different course/section.

Portfolio
A compilation of student work, including perhaps projects, artwork or writing samples, demonstrating achievement of multiple learning outcomes. May be in paper or electronic form, and may be used for course or program assessment. Portfolios are generally externally evaluated.

Pre- and post-test
A test or other assessment activity that is administered to students both at the beginning of a course or program and at the end, with the intention of demonstrating improved knowledge or skill upon completion.

Prompt
An assessment activity in which something such as a newspaper article, poem, or piece of art is presented to the student in order to prompt a specific response, usually written. Useful particularly in the arts and humanities.

Standardized Test
A test assessing academic achievement or of knowledge in a specific academic or vocational domain. Such tests are frequently objective (although some may be written tests with open-ended questions) and have scores referencing the scores of a norm group, providing comparative data. Standardized tests are generally commercial products and are useful in many areas. A current example at WCC is the use of CAAP tests to assess skills in the general education areas of math, writing, and natural science.

Surveys
Surveys may be used to evaluate perceptions of student achievement. Surveys of graduates, employers, or advisory committee members may help determine if program outcomes relating to employment and skill attainment have been met. Students may also be surveyed regarding self-perception of their success or, if administered as a pre-and post-test, of the improvement following completion of a course or program. Because surveys are indirect measures of student academic achievement, they are ideally used in combination with more direct measures.

Transfer follow-up
In courses or programs that have a high degree of transferability to other institutions, it may be useful to examine student success in subsequent courses at the receiving institutions. WCC has a large database of transfer follow-up information from Eastern Michigan University for assessment purposes.