Linking course objectives to learning outcome assessment efforts in course design.

**Grading By Objectives: A Matrix Method for Course Assessment**

*Ido Millet and Suzanne Weinstein*

**Abstract**

This article describes a method for linking course assessments to learning objectives. This method allows instructors to see the relative weight and performance of each learning objective, as reflected by course assignments and exams. While designing the course, instructors can use this information to ensure the relative weights are aligned with the relative importance of the learning objectives. When the course is completed, instructors can see, at a glance, which objectives students mastered and which ones they did not. This information can be used to modify the course prior to the next offering. Furthermore, this information may be utilized for learning outcomes assessment efforts. At our business school, this method was implemented via a spreadsheet and used for several years by a faculty member. We propose integrating the methodology into learning management systems.

**Keywords**

Assessment, Learning Objectives, Grades

**Introduction**

According to Frazer (1992), the basis for quality in higher education is self-evaluation, and a “mirror” is required for teachers and universities to become “self-critical and reflective” (p. 18). In this article we describe a method that allows instructors to see the extent to which assessments are aligned with learning objectives, as well as how students have performed on specific learning objectives. This allows instructors to adjust assignments and tests to better reflect the desired balance across learning objectives. Because learning objectives with poor student performance become visible, this reporting system can also lead to beneficial adjustments to teaching strategies. For course objectives that reflect program-level objectives, the information generated by this system may also contribute to program-level assessment.

Graded assignments and exams are one of the most important features of any course because they provide the opportunity for both students and instructors to assess how well students have learned the course content. The educational assessment process begins with the development of learning objectives, which define what we expect students to know or be able to do following the course or program (Biggs, 1999; Fink, 2003; Suskie, 2009; Walvoord, 2004; Walvoord & Anderson, 1998; Wiggins & McTighe, 2005). According to Wiggins and McTighe (2005), after learning objectives are developed, assessments are designed to inform the instructor and the student about the extent to which the student has met those objectives. When done effectively, this process results in course assignments, projects, and tests that are closely aligned with each learning objective. Research has shown that such alignment results in powerful effects on student learning (Cohen, 1987).

After the assessments are designed, the instructor plans the teaching strategies that will prepare students to perform well on these tasks. However, even if faculty members design their courses in this way, they may not take the time to evaluate the extent to which their assessments match their objectives or how well students performed on each objective. Instructors are typically more concerned with how well students performed on the combination of assessments, which is how course grades are determined (Weinstein, Ching, Shapiro, & Martin, 2010). Furthermore, a single assignment, project, or exam frequently...
assesses multiple course objectives, making it difficult to map students’ performance back to individual course objectives.

How Course Design Strategy Enhances Assessment

In this article we will describe a matrix method for assessment that is easy to implement and can help instructors improve the design of their courses while also contributing to program-level assessment.

Courses are commonly designed around topics. For example, an introductory psychology course may cover topics such as the biological underpinnings of behavior, memory, and abnormal behavior. In contrast, instructional design experts advocate that courses be designed around learning objectives, which are statements that delineate what students will know or be able to do after taking the course (Wiggins & McTighe, 2005). In this process, called constructive alignment (Biggs, 1999) or backward design (Wiggins & McTighe, 2005), the instructor begins with the desired results before developing assessments and teaching strategies.

Using the introductory psychology course as an example, the faculty member may state that he/she wants students to be able to compare and contrast the different theories of learning. He/she would then design an assessment, perhaps an essay question on a test, which aligns with that objective. The next step involves determining what activities students should engage in so that they are prepared to answer the essay question. For example, they may first read about the theories and then engage in a class discussion. When course objectives drive course design, both students and instructors are clear about what students will be able to know and do after completing the course.

Many instructors include instructional objectives in their syllabi and course design process, but how can we provide evidence that the course actually achieves its stated objectives? Aligning learning objectives with assignments and tests can help answer this basic learning assessment question (Diamond, 1989; Fink, 2003; Huba & Freed, 2000; Nitko, 1996; Suskie, 2009; Walvoord & Anderson, 1998; Walvoord, 2004; Wiggins & McTighe, 2005).

Explicit links between course objectives and assessments offer benefits beyond the course design process. Such links can ensure that an appropriate percentage of course assessments address each learning objective. Such links can also help measure students’ performance on each learning objective. The instructor can then use this information to improve the course in appropriate ways. For example, the information may prompt the instructor to add assignments and test questions linked to a relatively neglected course objective. Similarly, a course objective with relatively poor performance may prompt the instructor to change the course design to address the deficiency. Likely causes of low performance on a particular learning objective include problems with the objective itself, the assessments used to evaluate the objective, or the teaching strategies used to prepare students for the assessment (Suskie, 2012).

Beyond the contribution to course design and evaluation, linking assessments to course objectives may also benefit program-level assessment. If some course-level objectives address program-level objectives, the evidence of student performance on these objectives can be incorporated into the program assessment materials. This “embedded assessment” strategy can save time for instructors (Weinstein et al., 2010).

What follows is a method for linking course objectives to assessments using computer software. Once these links are established, no extra effort (beyond the usual grading of assignments and tests) is needed to generate the information described above.

A Matrix Method for Grading By Objectives

The core idea behind the proposed grading by objectives (GBO) method is that each graded task (assignment, exam question, quiz, or project) should be linked back to course objectives via a matrix. Figure 1 shows a simple case where a matrix links two graded tasks with two course learning objectives (LO).

<table>
<thead>
<tr>
<th>Objective</th>
<th>Matrix</th>
<th>Derived weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO one</td>
<td>100%</td>
<td>Max points</td>
</tr>
<tr>
<td>LO two</td>
<td>100%</td>
<td>Weight</td>
</tr>
</tbody>
</table>

Max points: 70

Figure 1: A Simplified Matrix That Associates Tasks With Learning Objectives

Although this scenario, in which each task is linked to only one objective, is overly simplified, it serves to demonstrate how useful information can be generated with minimal input from the instructor. For each graded task, the instructor simply needs to specify the relative extent to which it evaluates each of the course learning objectives. Given the relative weights (or points) assigned to each task, this allows us to compute the relative grading weight assigned to each learning objective. For example, the maximum points derived for objective one are 100% x 70 (from task one) and 0% x 30 (from task two) for a total of 70 points, or 70% of the grading in this course. If the instructor believes the intended relative importance of a learning objective does not match its actual derived grading weight, an obvious next step would call for changing the composition or relative weights of the tasks to close the gap.
When Tasks Are Not Specific to Learning Objectives

The case above was extreme in the sense that each task was tied specifically to only one learning objective. Figure 2 shows the other possible extreme where each task is associated equally with all learning objectives.

When tasks are not weighted toward specific learning objectives, we cannot derive relative weights or relative performance for learning objectives. Because of the centrality of the alignment between assessments and learning objectives for good course design, it can be argued that a course without this alignment may require reconsideration of the learning objectives. If the evaluation of one learning objective always entails equally weighted evaluation of all other objectives, we should probably rethink our course design.

For example, consider a Genetics 101 course with the expectation that, upon completing the course, students should be able to:

- Describe natural selection mechanisms and implications for disease resistance in humans.
- Describe natural selection mechanisms and implications for disease resistance in primates.

Since these knowledge areas overlap, there is a good chance that graded tasks in this course would have very low specificity (similar to Figure 2). This may prompt us to rearrange the course objectives so that, upon completing the course, students should be able to:

- Describe natural selection mechanisms in primates.
- Describe natural selection implications for disease resistance in primates.

This would probably yield much higher task specificity and, we believe, better learning objectives.

Mixed-Specificity Case

Our experience has been that even for well-designed course objectives, some tasks may not be 100% specific to a single learning objective. Figure 3 depicts such a realistic scenario.

In this particular case, the points allocated to each task are split, in the proportions specified by the matrix, across the objectives. Learning objective one receives 63 points (90% x 70 points) from task one and six points (20% x 30 points) from task two for a total of 69 points, or 69% of grading in this course. This demonstrates that even when tasks are not 100% specific, the results can still be quite useful.

Note that we can derive grading weights for the learning objectives even before the course has begun. This allows instructors to modify the course design by adjusting the mix of tasks to better reflect the relative importance of learning objectives.

Using Rubrics and Subtasks to Align Assessment With Learning Objectives

Even when a task as a whole is not specific to learning objectives, a rubric may provide separate evaluation criteria that, when aligned with learning objectives, can significantly increase assessment specificity (Suskie, 2009; Walvoord & Anderson, 1998). For example, a rubric for evaluating a paper may provide separate grades for writing, critical thinking, and knowledge of ethical principles. Similarly, although a final exam as a whole may not be specific to learning objectives, each question within the exam may be quite specific.

When a single task generates separate grades for different criteria or subtasks, we should record and treat the grade for each criterion or subtask as a separate assessment with its own weight. This would preserve useful information and increase our ability to align assessments with specific learning objectives.

From Task Grades to Learning Objective Performance Scores

Although we can derive grading weights for the learning objectives even before the course has begun, tasks must be graded before we can assess how well our students performed on each learning objective. Figure 4 shows how task grades are transformed...
through the GBO matrix into performance scores for the learning objectives. 

On average, students in this course scored 60% on task one and 90% on task two. This means that out of a maximum of 70 points, students averaged 42 points on task one, and out of a maximum of 30 points, students averaged 27 points on task two. Multiplying these task performance points (TPPs) by the allocation percentages in the GBO matrix allows us to split and recombine these points into learning objective performance points (LOPP). For example, learning objective one accumulates 90% of the 42 TPPs from task one and 20% of the 27 TPPs from task two for a total of 43.2 LOPPs. Given that the maximum LOPPs for the first objective is 69, we can compute an overall performance score of 63% for learning objective one. Similarly, learning objective two accumulates 10% of the 42 TPPs from task one and 80% of the 27 TPPs from task two for a total of 25.8 LOPPs. Given that the maximum LOPPs for the second objective is 31, we can compute a performance score of 83% for learning objective two.

Even though the tasks are not 100% specific to single objectives, this procedure provides useful information. We would be justified in concluding that students are struggling to attain learning objective one but are doing quite well with learning objective two. The instructor may then investigate the reasons for the poor performance on learning objective one and change the design of the course to address these deficiencies.

**Summative Versus Formative Tasks**

According to Allen (2005), the validity of a measure of student performance (e.g., a grade) is diminished if measures of behavior, effort, or practice are included. Thus, when measuring academic performance using the GBO method, we recommend focusing only on summative assessments scores, such as exams, end-of-topic assignments, and final papers because these types of assessments are designed to determine the level at which students achieved the learning outcomes at the end of a unit or course. In such a case, we should exclude formative assessments, such as practice quizzes or early paper drafts, which are designed to provide feedback and help students improve.

Yet, when we move from the measurement of academic performance to a broader objective of course diagnostics, we may include metrics for formative assessments. Keeping both types of graded tasks in the matrix (and classifying each as summative or formative) would provide useful information such as the relative assessment attention each learning objective receives in terms of summative assessments, formative assessments, or both. For example, the scatter chart in Figure 5 highlights a divergence between the summative and formative assessment weights for two out of four learning objectives. Learning objective one receives high summative but low formative assessment attention, while learning objective two receives low summative but high formative attention. These disparities may or may not be appropriate for these learning objectives. In any case, the GBO method would help make such disparities visible to the instructor, who can then make modifications if necessary.

**Ancillary and Mixed Grades**

Just as formative grades should be excluded when assessing the achievement of learning objectives, so should grades dealing with non-academic performance. Allen (2005, p. 220) states “grades should not be a hodgepodge of factors such as student’s level of effort, innate aptitude, compliance to rules, attendance,
social behaviors, attitudes, or other nonachievement measures." However, Allen (2005, p. 119) also claims that “although ancillary information such as effort and attitude could be part of an overall student report, they should not be part of a grade that represents academic achievement” (Tombari & Borich, 1999). Thus, instructors may choose to award points to non-academic behaviors for motivational purposes, but these points should be excluded from the GBO matrix if it is to represent a valid measure of student performance.

**A Case in Point**

For several semesters, one of the authors has used the GBO technique in a senior-level undergraduate course on business intelligence. A sanitized version (no student names) of the grading spreadsheet with the integrated GBO technique is available for download from: https://dl.dropboxusercontent.com/u/38773963/Grading_By_Objectives_Sample.xlsx.

The spreadsheet starts with a row-wise listing of the learning objectives. The intersection of learning objective rows with assessment columns (cells J7 to AS9) is used to indicate how each assessment is allocated across the learning objectives. Since each assessment must distribute all its weight across the learning objectives, each column in that range adds up to 100% (J10:AS10). The maximum points value for each learning objective is computed by multiplying the assessment points for each assessment (J14:AS14) by the percent allocated for that learning objective and summing across all assessments. The weight % column is then computed by dividing the maximum points value for each learning objective by the sum of maximum points across all objectives. In Figure 6, the weight % column shows that 50% of the assessment in this course is allocated to the third learning objective.

The class performance on each assignment is converted to percent scores (J1:AS1) by dividing the average score for each assessment by its maximum points. Multiplying these actual scores by the allocation percentages for each learning objective and summing across all assessment provides the actual points (G7:G9) scored for each learning objective. Finally, dividing actual points by maximum points provides the percentage of maximum metrics (H7:H7) and reflects how well the class performed on each objective. Similar logic is applied to compute how well individual students performed across the learning objectives. In Figure 6, the percentage of maximum column shows that the class performed on the first learning objective was lower (75%) compared to the other two learning objectives (89% and 88%).

When the method was first used, the second learning objective (suggest and design improvements) had a grading weight of just 17%. This revelation was a surprise to the instructor who then added two new assignments and several exam questions to bring the grading weight for that objective to its current value of 31%. This supports and demonstrates Frazer’s (1992) claim that self-evaluation is necessary to achieve quality in higher education, and that a mirror is required for teachers to become “self-critical and reflective” (p. 18).

In this particular course, there were more than 20 graded tasks. This explains how even an experienced instructor might be unaware that an important learning objective is not subject to sufficient assessment. We believe that the GBO method becomes even more helpful when a course has many graded tasks.

**Integrating Grading by Objectives Into Learning Management Systems**

The increasing popularity of learning management systems provides an opportunity to embed the computational logic of the GBO method in the software already used by instructors to set up tasks, assign relative points to these tasks, and record grades. The GBO method would simply require that instructors specify learning objectives, allocate each task across these learning objectives, and classify each task as formative or summative.

Once these aspects are embedded within a learning management system, we estimate the extra input required from the instructor would demand no more than 20 minutes per course. This does not count the time to interpret and act upon the reports generated by the method. However, since this feedback would help instructors improve their course designs, we believe most instructors would welcome it. In cases where the same course is taught by different instructors, reports from the system can highlight metrics with significant differences. For example, an instructor whose students are struggling with a particular learning objective would be able to seek advice from the instructor whose students are performing best on that particular objective.

**Using Grading by Objectives for Program-Level Assessment**

Although course grades are not appropriate metrics for learning outcomes assessment, scores on specific tasks within a course that align with program-level objectives are appropriate methods for providing evidence that students are meeting the program learning outcomes.
objectives. Thus, the GBO method provides an added benefit for instructors teaching courses in which assessments for one or more learning objectives will be used as evidence that students have met program-level objectives. The embedded formulas will automatically generate a percentage that represents the extent to which students have met a particular objective. This result can then be included in the program assessment report and discussed by program faculty as part of the process.

Limitations

The proposed methodology assumes proper learning objectives can be identified for courses. Yet, several researchers cast doubt on the ease and advisability of establishing such objectives. A good review of such objections is provided by James (2005). We need to exercise care and proper balance in establishing learning objectives:

*If learning outcomes are defined too broadly, they lose their capacity to assist in comparison across cases and over time. Defined too narrowly, they become impotent, in the sense that they refer to so little of a de facto learning process that they are simply uninformative, powerless to generate or even signal improvement.* (James 2005, p. 90).

Although instructors may welcome access to GBO metrics and reports, one sensitive consequence of this method is that it would make problem areas visible to administrators and, possibly, to other instructors. This tension between "external control and internal improvement" (Padró, p. 2) is an issue for any assessment initiative. However, since the GBO method uses grades as input, it raises the threat that instructors might be tempted to assign higher grades to escape negative attention from administrators and peers. To avoid such unintended consequences, it may be wise to restrict detailed feedback to constructive use by the instructor.

Allen (2005) warns “grading systems used by teachers vary widely and unpredictably and often have low levels of validity due to the inclusion of nonacademic criteria used in the calculation of grades.” The GBO method strives to remove non-academic criteria by exclusively using summative grades for assessing the achievement of learning objectives. Still, there is a remaining concern about the reliability and consistency of summative grades. To reduce possible instructor bias, subjectively scored tasks, such as essays, papers, or presentations, should be scored using well-developed rubrics, which make scoring more accurate, unbiased, and consistent (Suskie, 2009). Close attention should also be paid to creating reliable objective tests such as multiple choice tests, which requires significant effort (Suskie, 2009). Also, a grade lift reporting system (Miller, 2010) may promote grading consistency across faculty members.

Future Research

Future research may investigate the impact of using the proposed GBO methodology on teaching practices, grades, academic performance, and student satisfaction. It would also be important to collect feedback from instructors who are early adopters of the technique. Such feedback may include overall satisfaction, suggestions for improvements, and level of impact on course and assessments designs.

As mentioned earlier, establishing proper learning objectives is an essential yet challenging aspect of any learning assessment effort. Future research is needed to establish guidelines for the creation of effective learning objectives for various educational contingencies.

Another interesting question relates to the proper balance between formative and summative assessment. As depicted in Figure 5, the GBO methodology provides descriptive information about that balance as reflected by graded tasks. However, we lack prescriptive insight. What type of balance is conducive to achieving different types of learning goals in different situations? For example, do undergraduate students require a greater proportion of formative tasks? Do students benefit from a greater proportion of formative tasks for learning objectives at higher levels of Bloom’s taxonomy, such as analysis or evaluation (Bloom, 1956)? Do courses with a higher proportion of formative tasks lead to better long-term knowledge retention? What is the impact on student engagement and performance when formative task grades are downplayed in computing final grades? Answers to these and other questions associated with the impact of formative assessment on student performance would benefit our educational systems.

References:


I do Millet, Ph.D. is a professor of management information systems at the Sam and Irene Black School of Business, The Pennsylvania State University-Erie. His research interests include the analytic hierarchy process, online reverse auctions, business intelligence, and use of academic data to support faculty and students. Millet’s industrial experience includes systems analysis, project management, consulting, and software development. His business intelligence software packages have been purchased by more than 5,600 organizations. For more information, contact him via email at ixm7@psu.edu.

Suzanne Weinstein, Ph.D. is director of instructional consulting, assessment, and research at the Schreyer Institute for Teaching Excellence at Penn State University. She also holds a courtesy appointment in the department of psychology. Weinstein joined the teaching and learning support community at Penn State in 2002 as an assessment specialist. Her assessment expertise ranges from single courses to multi-course projects to university-wide learning outcomes assessment. Contact her at swd107@psu.edu.

The Quality Approaches in Higher Education editors will announce an annual best paper award to the author(s) of a paper published in Quality Approaches in Higher Education. The award will be announced in January of each year for the best paper from the issues of the previous year and will be based on the largest single contribution made to the development or application of quality approaches in higher education. There is no nomination form for this award.

Visit our website at asq.org/edu/quality-information/journals/today!