Using assessment-based methods for grading and for engaging students in computer-science courses

David Keil, Computer Science Department, Framingham State University

Abstract. Attention to objectives, and grading on objectives, may bring clarity about why we use certain methods of evaluation for grading. A grading system for computer-science courses that is 30-40% based on attainment of desired learning outcomes, rather than just on scores on evaluation, has shown advantages for grading and for student learning. Survey results using a small sample seem to indicate that a grading system based entirely on measured attainment of objectives (not limited to learning outcomes) may work well. There is reason to hope that by focusing grade-related interest in the direction of course content, as opposed to methods of evaluation, student engagement in course work can be encouraged.

1. Introduction

This paper presents a way that data derived from steps taken in assessment of learning outcomes in computer science courses have been used in grading student work in those courses. To do this, triple use has been made of some numbers generated in evaluating quiz questions: to assign quiz grades; to assess particular learning outcomes; and (indirectly) to generate some numbers usable in semester grading that depend exclusively on attainment of learning outcomes.
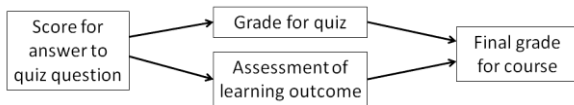
Technical tools available today support such double and triple uses of data. Electronic spreadsheet applications enable powerful and flexible uses of numbers generated in the grading process.

Section 2 describes this method, which has been presented to the Computer Science Department as part of the FSU assessment project.

In some computer-science courses, 30% to 40% of the semester grade is computed from aggregate figures for individual student attainment of stated learning objectives (desired learning outcomes). These objectives are listed in the course syllabi and are referenced in course material and classroom presentations.

In Section 3, we report some student responses as early results of a system that focuses students' grade concerns more toward learning objectives and less toward the vehicles (assignments, quizzes) used to assess them.

Based on tentative success with a system in which a large part of the final grade is based on assessment data, Section 4 describes an envisioned system in which the entire grade would be generated from evaluations of learning outcomes as well as other desirables, such as student participation, research, presentations, and steady activity keeping up with the course material and course schedule.

Section 5 offers some reflections on how methods of outcomes assessment and grading may be used to increase student engagement. The expanding use of metrics offers both opportunities and risks to learning, teaching, and academic freedom. This paper acknowledges the risks and focuses on the opportunities offered to the instructor to use grades to help win students' engagement with often difficult and technical course material.

2. A system of grading

I assess objectives using in-class problem-solving or long-answer quizzes given when topics are completed or at the end of the semester. Topic quizzes typically have two to four questions and may require fifteen to twenty minutes of students' time.

Each question is written so as to assess a course objective, such as "2b. Prove correctness of an algorithm, using invariants, preconditions, and postconditions" (CSCI 347 Analysis of Algorithms). The course objectives come from a list of desired outcomes for all core courses required in computer-science major concentrations, adopted by the Computer Science Department at Framingham State University in 2009. The lists vary in length from a handful to three dozen or so. Each objective describes a capacity or skill that a student is to be able to demonstrate by end of semester. Some objectives are course-wide; others are topic specific.

Questions are graded on a scale of 0 to 1.0, where the score maps naturally to a grade; e.g., a score of 0.85

multiplied by 100 yields 85, whose meaning is "B quality work" or "good quality answer."

Two significant advantages of assigning numbers on a 0 to 1 scale to each question in an assignment or quiz are that the meaning of any number is immediately available, and that a weighting system implemented by spreadsheet formula enables re-weighting of questions at any time, including after a quiz is administered. If the weights of quiz questions are in spreadsheet cells P1, Q1, and R1, then for a student whose scores are recorded in row 6, the formula "=P\$1*P6+Q\$1*Q6+R\$1*R6" gives the quiz score, out of 100, provided the sum of cells P1, Q1, and R1 is 100 and provided cells P6, Q6, and R6 contain values in the range of 0 to 1.

I use numbers from 0 to 1, rather than 0 to 100, because they seem to be easier to write and to type. Some students don't immediately understand that 0.6 means 60% and 0.9 means 90%. What seems most important is to have a uniform scale for evaluation of the quality of student answers.

If a question is graded under a rubric of multiple criteria, then the weights of grading criteria are used in a similar way to the weights of questions (above), and each criterion is evaluated in a range of 0 to 1.

If hardly any student attains a certain objective, I might scale up the scores for grading purposes or reduce the objective's grading weight. Certainly each objective needs a weight, chosen carefully.

In the past, I was used to allocating "points" to quiz and assignment problems, so that the grading decision was how many points out of, say fifteen, to give to a student's answer. A good answer might get twelve points. A good answer to a ten-point question might get eight points. What is the precise meaning of eight or twelve to the student? It is likely to be unclear.

For assessment data collection, scores on quiz questions may be collected in a series of spreadsheet columns, one per course objective. Averages for a given objective over all students yields assessment information on how well the students attained the objective. Averages for a given student over all objectives yield data on how well the student attained all objectives. Averages may be weighted for objectives that are more important than others.

Students may track scores for learning objectives posted online. Thus student attention is drawn not only to quiz and assignment scores, but to scores directly related to the stated objectives of the course.

If average scores on an objective are lower than those on another objective, I may give it more attention during the semester to present the material more thoroughly. Alternatively, the objective may be deemed hard to attain

in the time available, and the scores may be scaled up for fairness.

This system of grading may be extended to enable students to show attainment of objectives at several times in the semester, with at least three opportunities offered per objective. The MAX function in the Excel spreadsheet application supports this best-performance notion.

I have been allocating 30% to 40% of semester grades to objectives attainment. Other components of grades are classroom participation, quizzes, assignments, an exam, and research. Thus scores for quiz questions that assess learning outcomes may have two separate effects on final grades.

Course-wide objectives may include the following as examples:

- Solve a problem as part of a team
- Present a short talk in the classroom
- Write a short documented research paper

I find that it makes sense to include in my grading system a way to take into consideration that the assessment deadline to determine learning outcomes is end of emester, not the dates of quizzes or assignments. I encourage students to correct or make up their work for credit after submission and (possibly preliminary) grading. This departs from what is described as "most students' expectations of 'produce a product one time only and receive a grade' " (Taylor), which reflects the notion of the student as a vendor and the teacher as a customer.

3. Some preliminary results

I am happy with my system because, using the same numbers, it enables me to assess my students' learning as well as to generate fair grades, and it helps me to convey my course objectives to my students in a way that is meaningful to them, both numerically and by reference to the course objectives. Students pay attention to grading. Grading on objectives helps students pay attention to them.

Students value grading fairness very highly. Stating my objectives, and then assigning grades that correspond to performance on them, supports the claim of fairness. It also gets attention for what I value in my teaching, which is not quizzes or grades, but is summarized in my objectives.

There is reason to hope that by focusing grade-related interest in the direction of course content, as opposed to method of evaluation, student engagement in course work can be encouraged.

Objectivity is important both in assessment and in grading. Low student performance on a quiz may be in part a result of insufficient development of the method of

instruction or a result of using very challenging questions. If either is the case, then a negative objective assessment of poor learning results can be balanced by scaling up grades. A low assessment may indicate a need for a review of how course content was presented or a review of quiz questions.

At the end of the fall 2011 semester, I requested that students in two courses answer five questions on a survey about the use of objectives in grading the courses. One was Analysis of Algorithms, a third-year core course for computer-science majors; the other was Information Technology and Society, a first-year General Education course for students of diverse or undeclared majors. Seven Algorithms students and ten IT and Society students answered survey questions.

Reducing responses to yes/no or positive/neutral/negative, the survey results may be summarized as follows:

Question	Reply	3rd-year	1st-year
How would you describe the grading system for this class and your opinion of the grading system?	Positive Neutral Negative	5 0 2	4 2 4
Do you think that listing course objectives in the syllabus and referring to them during the semester helped you to know better what was expected of you?	Yes No	8	9 1
Do you believe that this course focused too much on the objectives, too little, or about the right amount? Please explain.	Too much About right Too little	1 6	6 4
Do you have any other classes that focus heavily on objectives?	Yes No	3 5	4 5
What would your thoughts be about a course that would be graded entirely on course objectives?	Positive Neutral Negative	1 2 5	1 4 5

Thus two thirds of the students rated the grading system positively and almost all appreciated the listing of objectives in the syllabus. Third-year students considered the degree of focus on objectives to be "about right"; more first-year students responded "too much" than "about right."

More students reacted negatively than positively to the idea of grading a course entirely on objectives; a third were neutral. Overall, it seemed that students appreciated knowing the course objectives, but were both somewhat open-minded and somewhat wary about grading based on attainment of objectives.

4. A plan for a grading system based entirely on objectives

Attention to objectives, and grading on objectives, may bring clarity about why we use certain methods of evaluation for grading. For example, why do we assign several assignments, rather than one paper, and administer several quizzes, rather than just a midterm and a final exam? It is surely in part because one of our objectives is to support students staying at a pace with us and with their classmates in their studies throughout the semester, rather than each working separately and at a separate pace. We balance this concern with the acknowledgement that everyone learns at a different pace.

Our grading systems reflect what we value and what we want to encourage. Even though what we value for grading purposes (such as keeping pace) is not only learning outcomes, it is fair to call keeping pace an "objective," Especially since it supports collaborative learning.

Rather than assigning quiz and assignment components of the final grade, it may provide clarity to assign a grading weight to the evidence that a student has kept up with the pace of the classroom presentation and has taken part in classroom discussion. Attendance records and written work may be evidence of this that is attainable, though indirect. A page at Carnegie Mellon University's web site contrasts attainment of learning outcomes with other tasks, activities, and behaviors that may be part of a semester grade (Eberly Center, undated).

Our pedagogical objectives do not necessarily include at all the mere physical presence of students in the classroom. Many first-year students at FSU assume that physical presence is equivalent to success. Grading schemes send messages about these matters. I use "Participation" to label intellectual presence and participation, however measured – by taking attendance, by estimating students' active participation in classroom discussion, or by going to the blackboard.

When a student submits an assignment or takes a quiz, he or she partially satisfies my concern that everyone participate. Quizzes or assignments aren't values per se, but rather are *strategies* for realizing my values that the students attain desired learning objectives and that they participate together throughout the semester.

For spring 2012 courses, I'm taking my grading all the way to base it on assessment objectives. Here is a tentative grade breakdown for a Spring 2012 course:

Grading component	%	Instruments
Application of capabilities and knowledge		Problem quizzes
core objectives	25	
other objectives	25	
Knowledge of concepts and facts	15	Mult-choice quizzes
Independent inquiry	10	Research paper
Presenting results in person	10	Rubric
Participation	15	Recorded participation

The following modified diagram indicates that a quiz problem, like any instrument of evaluation, may feed into a framework of assessment data, from which a grade may be generated as well.

Score for answer to quiz question → Assessment of learning outcome → Final grade for course
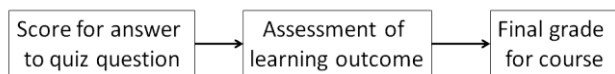
5. Reflections and conclusion

I am changing how I see myself: from customer, with the message, "Bring me good work and I'll compensate you with good grades," to provider of learning support services, with the message, "You have a choice to engage with my course or not. Here is what it has to offer you if you choose to join our effort."

Good teaching entails gaining the students' voluntary attention, such as by "beginning a lecture with a provocative question or problem that raises issues in ways that students had never thought of before" (Bain, 2004). Thus we need to *invite* our students to engage actively with what we are presenting.

If we have some learning outcomes in mind, then one way to offer transparency and to win buy-in is to place these desired outcomes before the students in a way that will gain their attention: as part of the grading scheme. Since what we value is the outcomes, our grading schemes may appear as more authentic to students if they reflect the outcomes and other concerns that matter to us.

Student awareness of intended learning outcomes is reportedly inadequate. Administrators report common sets of intended learning outcomes for all undergraduates at their institutions, but also report that students aren't aware of goals (Hart Research Associates, 2009).

Assessment in education and other fields occurs within a globalized economy, enabled by a global information infrastructure that elevates the role of quantitative data. One may refer to an "assessment revolution" and an "assessment society"; data collection for summative assessment purposes has become "a major instrument of social control" (Broadfoot, Black, 2004).

Formative assessment may facilitate learning with understanding (deep learning). Some researchers report, however, that "assessment fails to have a truly formative role in learning" (Harlen, James, 1997). Others note,"the majority of the existing literature is centred on summative assessment" (Rushton, 2005).

Researchers may be saying that if grades and assessment were to serve chiefly to help shape learning and teaching (with a formative intention), rather than chiefly to evaluate student work and teaching effectiveness (summative intention), then they would be more likely to improve student engagement and deep learning.

From the student's point of view, grading may play two divergent roles: as a means of social control and selection, or as a way to obtain self-insight and desired guidance and validation from an expert in the field. For many students, the second role may fade in importance. Moreover, research shows that if students perceive extrinsic motivation (grading) as manipulation of them, then the effect on learning is negative (Bain, 2004).

It seems to me that self-reflection, an activity of deep learning, may be informed, however, by familiarity with the goals of the instructor, provided the student is aware of these goals and provided the student embraces them.

Assessment, including assessment based on grading, contains potential risks to academic freedom. There is concern that metrics processed by spreadsheets may impose external control within the zone where teacher judgment and decision are required. Every teacher needs the freedom to creatively shape one's own objectives and methods within broad boundaries.

(Broadfoot, Black, 2004) envision assessment principles "born of educational, rather than measurement, priorities." That is the goal of a possible reframing of the notion of grading. I am choosing to reframe grades as part of a process that aims at formative assessment, student engagement, and deep learning.

References

Ken Bain. *What the Best College Teachers Do.* Harvard University Press, 2004.

Patricia Broadfoot, Paul Black. Redefining assessment? The first ten years of 'Assessment in education'. *Assessment in Education* 11(1): 7-27 (2004).

Eberly Center for Teaching Excellence. Grading Versus Assessment of learning outcomes: What's the difference? Carnegie Mellon University. *http://www.cmu.edu/teaching/ assessment*. Downloaded 1/12.

Wynne Harlen, Mary James. Assessment and Learning: differences and relationships between formative and summative assessment. Assessment in Education 4(3): 365-379 (1997).

Hart Research Associates. Learning and Assessment: Trends in Undergraduate Education, April 2009.

Alison Rushton. Formative assessment: A key to deep learning? *Medical Teacher*, 27(6): 509-513 (2005).

Peter Taylor. More assessment, less grading. *http://www.cct.umb.edu/ moreassessment.html*. Downloaded 1/12.