

Evaluating and Grading Undergraduates: Policies and Best Practices

ABSTRACT

Faculty choices on evaluating and grading students have a profound impact on learning. Assessment structures impact not just what a student learns but also their attitudes and motivation to learn. This paper discusses the Undergraduate Division's position on grading systems and assessment methods, presents information about the advantages and disadvantages of the two systems used, and offers guiding questions to help faculty decide and articulate assessment methods that are congruent with teaching philosophies and course objectives. Information about mid-semester academic warnings and final grade submission is also included.

INTRODUCTION

Faculty teaching undergraduate courses have discretion in setting grading standards and evaluation methods as they see fit. Unlike the Wharton MBA Program, the Wharton Undergraduate Division does not require a set average GPA or a particular grade distribution for a class. However, we do encourage faculty to carefully consider the methodology used in assessing work and assigning grades.

Criterion grading systems place the burden on faculty to design consistent evaluations. Normative grading systems place the burden on students by encouraging conflict through competition, though research shows that these systems tend not to motivate students to learn as well as criterion systems do. Whichever system you choose, we recommend that you communicate your standards and expectations to students at the beginning of the course.

We recognize that the choice of grading system and evaluation method has an impact not just on student learning, but also on student wellness. The Undergraduate Division strives to support faculty in creating classroom environments where both learning and wellness are optimized. Pedagogical choices that align with teaching philosophy and course objectives are appreciated by students, and contribute positively to their undergraduate experience.

GRADING SYSTEMS

Wharton faculty use both normative and criterion systems. The Wharton Dean's Advisory Board (WAB) student group recently reviewed the grading structures used by instructors of undergraduate business fundamentals ("core") classes in academic year 2014-15. Normative grading was used by 7 of the 11 sections in the review, and criterion grading was used by 4 of the 11. Normative grading standards vary from course to course. In academic year 2012-13, for example, FNCE 101 courses issued 43% As, 36% Bs, and 20% Cs (n = 587); OPIM 101 courses issued 33% As, 57% Bs, and 10% Cs (n = 541). Applying the more steeply-curved OPIM 101 distribution to FNCE 101, we would see significant differences in the grades disbursed: 193 students would have received As in FNCE 101 instead of the 252 that actually did. Criterion grading standards also vary, though the courses in the study tended to establish the A range as 90% or above, B range as 80% to 89.9%, and C range as 70% to 79.9%.

Though there is more evidence substantiating the benefits of criterion grading in terms of enhancing student engagement and motivation to learn, all grading schemes have their respective benefits and drawbacks. We encourage you to be aware of all the ramifications of whatever grading system you choose to use, and take reasonable measures to minimize any unintended consequences. We also encourage you to consider the point of view of the student as learner as you develop your evaluation schemes. Given the objectives you establish for your class, does your grading system match those objectives? Are you seeking to discover and reward a set number of high performers for a particular purpose? Are you seeking to encourage the attainment of mastery from most if not all students? How does your grading system help you achieve what you want to know about students' learning?

Normative grading is a relative grading system familiar to, if criticized by, college students. It is most common in core classes with large enrollment. Performance is measured relative to other students, and grades are determined according to one's relative rank in performance, which is then transferred onto a distribution that is determined by the professor. Mythology about a "Wharton curve" abounds even though no uniform standard is imposed or expected. Students are keenly aware that their effort and performance in a class with normative grading may not directly reflect in the final grade they receive, and often lament that the difficulty of an exam and arbitrary delineation of the distribution confounds their understanding of the quality of their work. Many students focus on the drawbacks of normative grading, especially to the extent that they are interested in knowing whether they have achieved mastery of the material. While normative grading allows for standardization within a class over time, it does not necessarily ensure standardization from the student's perspective. Normative grading tends to encourage more questions from students about their standing in the class, and can leave them uncertain about how to prioritize their efforts especially during final exam preparation. This uncertainty can then lead to stress that worsens performance. In courses that require collaboration, normative grading can discourage students from working well together, knowing that their teammate on a project is also competing with them for a spot amongst the set number of As made available to students. In general, motivation to learn is reduced within normative systems.

Criterion grading or mastery-based grading is more common in seminars, upper-level electives with fewer students to evaluate, and a growing number of core classes. Performance is measured by whether or not a student achieves a certain level of mastery for given criteria. Students have increasingly called for more criterion-referenced evaluation methods, seeking to reduce the real or perceived competitiveness as well as advocating for a more pure measure of what they are asked to learn. However, this system is also not without its challenges for faculty. Rubrics need to be well designed so that students have a clear sense of the exact criteria they are expected to achieve. Students may also expect that grades will be higher in a criterion-referenced course, since it is not a foregone conclusion that a predetermined number of lower grades will be issued. Therefore, one key to successful criterion grading is to set standards that are challenging but achievable—this in and of itself is a difficult task and may require frequent reevaluation of the rubric.

Some faculty employ a hybrid of normative and criterion grading systems. Normative grading might be utilized on exams, while criterion grading might be applied to homework assignments and projects. This type of "best of both worlds" approach can enhance collaboration and learning while minimizing competitiveness, intellectual disengagement, and stress.

ASSESSMENT METHODS

The WAB study of 2014-15 core classes reveals a range of assessment methods used and their weight relative to the final grade. Most classes employ a mix of assessment methods: homework (including cases, problem sets, lab assignments, and papers); group projects (including case presentations and simulation); quizzes; and formal exams (usually one or two midterm exams and a final exam). There is also variation on the weight of these assessment methods. For example, formal exams comprise at much as 90% of a student's final grade or as little as 25%.

As with the type of grading system employed, the Undergraduate Division does not mandate the assessment methods used and the weights assigned. However, we do again recommend that faculty consider how these and other evaluation methods and weights contribute to optimal student learning, and develop a philosophy and practice accordingly. This includes consideration of the unintended negative impact of such pedagogical choices. WAB found that most students prefer courses with a range of evaluation methods and multiple chances to demonstrate their mastery of the material.

By offering a *variety of tasks*, students have an opportunity to perform in a medium that might best fit their learning style. For example, kinesthetic learners, students with a high degree of creativity, and extroverted personalities may excel in a presentation format. Formal exams may better suit students who are “test wise,” more conventional or concrete thinkers, and visual learners. Variation of task also more closely mimics the way that evaluations are conducted in the work world, where students will be expected to demonstrate their knowledge in myriad ways. Variations in learning tasks also enhance learning by reinforcing deep-level understanding of the material. Asking students to apply knowledge through different means – case analysis, simulation, group discussion, quiz, etc. – requires more sophisticated learning than might occur with a single method such as a formal exam.

Multiple evaluations also help to ensure more student engagement in the class, protecting against “slacking” in between longer assessment intervals and the subsequent last-minute “cramming.” Learning theory recognizes that students more deeply encode learning as they have more opportunities to engage with the target material. In general, more exposure is better than less exposure, and offering several different evaluations rather than one or two during a semester is more conducive to learning. Students can also experience extraordinary stress when an evaluation accounts for a significant portion of their final grade. While some students' stress during finals can be as much a function of their own ineffective study habits, faculty can discourage procrastination and poor performance by reducing the stakes of a given exam.

We recognize that adding variety and quantity of evaluation methods is not without cost: one exam is logistically easier to write and administer than two or three would be, and pausing too many times to evaluate student learning interrupts the flow of instruction and, ultimately, limits learning. What we encourage you to consider is the benefits and the costs of the evaluation methods you use, and how those methods contribute to the kind of learning you want to see from your students.

COMMUNICATING POLICIES AND EXPECTATIONS

The course syllabus is the best way to document the grading structure and evaluations methods you choose to use. Undergraduate students in particular are eager to know the relative weight of assignments on final grades; details on how they will be assessed and what the feedback mechanisms are; and the quality of work that differentiates grades. By articulating your teaching and grading philosophy in your course syllabus, you can set a

clear message from the beginning and head off unwanted questions and negotiations that might otherwise surface later.

The Center for Teaching and Learning at Penn offers detailed advice on establishing a [grading policy](#) as well as [sample syllabus language](#) to communicate your policy.

COURSE PROBLEM NOTICES

The [Course Problem Notice \(CPN\)](#) application is Penn's early academic warning system. Faculty use the CPN system to inform students and their academic advisor about performance issues at any point during the semester. When an instructor creates a notice, an email is sent directly to the student; a copy of the notice is placed in the student's electronic record; and a copy of the notice is sent to the student's advisor or home school advising office. In the case of Wharton undergraduates, the student's assigned academic advisor follows up with him or her to ensure reported issues and problems are addressed according to the instructor's recommendations.

Early identification and intervention helps prevent problems from escalating. Students who receive CPNs mid-semester have a significantly better chance of remedying their problems, and we encourage faculty to document academic concerns as soon as possible. The Undergraduate Division has developed [guidelines](#) to help you decide when and for what reasons to use CPN. If in doubt about whether or not to document your concern, err on the side of sending the notice.

REPORTING GRADES

Grades are reported for each course at the end of the term via [Instruction Center](#). Students must obtain a grade of D or better to receive credit in any course. All Wharton courses are graded on a plus/minus system, from A+ to F. Wharton students who take courses in any other school of the University are subject to that school's grading system for those courses.

- *Failing grades.* An F in a course will remain permanently on the student's transcript and is factored in when calculating a student's cumulative GPA. If a student receives an F in a required course, the course must be retaken. If a student receives an F in a non-required course, she or he may either repeat the course or substitute another course. If a course is retaken, the new grade will not replace the original F, the new grade will be counted toward the cumulative GPA, and the student will receive credit for the repeated course.
- *Incompletes.* In extenuating circumstances, students may be granted an extension of time by an instructor to complete course requirements, including make-up exams. In this case, the grade of I (incomplete) is recorded. All work must be completed for the course within the first five weeks of the following semester, or the I will automatically convert to an F and may impact a student's academic standing. The F will remain on the transcript until the work has been completed and the instructor has submitted a change of grade for the course. According to Wharton's [policy on incomplete grades](#), students must resolve any incomplete within two full semesters (excluding summer) after the term in which the course was taken.