

The Contamination of Student Assessment

IHE insidehighered.com/views/2018/08/09/why-faculty-shouldnt-factor-class-attendance-and-participation-final-grades-opinion

As professors are consistently reminded, in a student's world of class rank, graduate school admissions and a highly competitive job market, grades rule. Given that, fairness and accuracy in the testing by which we measure student performance and assign grades is one of the foremost commandments of the professoriate.

Yet, despite the best of intentions, faculty members often violate the commandment. Turning a blind eye to the methodological rigor by which we'd conduct an experiment or a survey, we often give our students quizzes and tests that include a hodgepodge of non-performance-related variables that, although they may enhance the classroom experience, taint the validity of our tests as comparative measurements of performance.

At least three such variables are: extra credit for attendance at or participation in outside events, upgrades for class attendance and downgrades for poor preparation or disruptive behavior. Upgrades for class participation present a fuzzier case.

In the era of outcomes assessment, testing serves to measure, more than ever, whether students have assimilated particular knowledge and developed certain skills. A student's mere exposure to information and instruction in skills does not, in today's assessment regime, reflect a successful outcome. The assessment crowd wants proof that it sank in, and grades are the unit of measurement.

Accordingly, extra credit for attendance at, say, even the most erudite and inspiring guest lecture outside class corrupts grades as a pure measurement of performance.

Sure, the lecture can be of value, whether demonstrable or not, in the intellectual development of the student, and giving credit for going to it is an effective incentive to attend. Nonetheless, that type of extra credit contaminates grades as a measure of performance, as it can allow the grades of students who attend extra-credit events to leapfrog over the grades of those who outperformed them on the exam but did not attend.

It's easy to dismiss this concern as mathematically improbable. That's bad math, though. The small likelihood of the event multiplied by the number of courses taught at one's college becomes, over time, a virtual certainty to affect a sizable number of students.

And, of course, there's always the chance that the student who attended the inspiring lecture for extra credit sat, eyes glued on the electronic device in their lap, lost in cyberspace the whole time. In addition, many students may have pre-existing commitments that preclude their attendance at one of the wide variety of events for which professors hand out extra credit as a door prize.

The next contaminant of grades as measurements of performance is the upgrade for stellar attendance. There is, of course, no guarantee that the ubiquitous attendee wasn't an accomplished daydreamer or a back-row socialite. And if they were present and genuinely plugged in to each class, their diligence should show up in their exam performance, so extra credit merely gilds the lily.

Using the stick as well as the carrot, some professors do the opposite: they downgrade students for disruptive behavior or chronically poor preparation or attendance. Like a doctor using a hammer to anesthetize a patient, downgrades aimed at controlling behavior produce collateral damage. Colleges have better tools -- like meetings with the dean of students -- to address conduct-related problems.

Finally, we come to what may well be the most common nonperformance variable incorporated into grades: the class participation upgrade that so many of us rely on to break the deafening silence we'd otherwise encounter in casting pearls of wisdom upon the class. Class participation upgrades that recognize and reward the volume, rather than quality, of a student's classroom contributions pollute performance-based assessment.

Participation upgrades for remarks that consistently *advance* the class discussion is a more complex issue. On the one hand, such upgrades can stimulate and enrich the conversation, draw otherwise detached students into the debate and thereby enhance the value of the course. Moreover, it's clearly possible, outside the formal testing process, for a student to demonstrate the knowledge and skills that grades are expected to reflect, and, by definition, strong class participation does exactly this.

Is it as clean a measurement as a test? Is the teacher's valuation of the strong contributor more subjective and more open to implicit bias than the cold numerical scores on a test? While the grading of multiple-choice exams seems impervious to bias and subjectivity, and anonymously graded exams appear bias resistant (although not necessarily in content), other common test formats allow for the same types of abuse that can warp a professor's identification of strong class participation. The bottom line here is that the case against upgrades for high-quality participation is no different or stronger than the case against the various forms of testing (with the exception of multiple-choice exams), and we're not about to eliminate those forms out of an inordinate distrust of teachers.

Professor bias, though, is not the only potential pitfall with performance-based participation upgrades. The playing field must be level for all students, and it seems doubtful that it always is. Although women make up a majority of today's college students, the historic culture of male domination may still, in many classrooms, exert a pressure against women to consistently speak out or to challenge the remarks of a male classmate. Minorities in many classrooms may well share this experience vis-à-vis their white classmates.

The unseen cultural forces that reverberate through class discussion aren't the only problem with performance-based participation upgrades. The final nail in their coffin is this: the student's firm grasp of the material that their high-level classroom remarks reflect is bound to

show up on tests, and thus the upgrade results in a higher grade than the stellar classroom contributor actually merits.

One might interject here that higher education is more than a contest of right and wrong answers, and that, in optimizing the classroom experience and exposing students to events outside the classroom, the various types of non-performance-related upgrades expand the intellectual horizons of our students. While true, we often have, as I've noted, alternative means of accomplishing these laudable ends. Moreover, we mustn't conflate the broad objectives of higher education with the requirements of fair grading.

Accordingly, for grades to represent the unadulterated measurements of knowledge and skills that we represent them to be -- and that employers and graduate admissions committees rely on them to be -- we should dispense with the various upgrades we award. Which raises the question of whether a faculty- or administration-imposed ban on upgrades is a violation of academic freedom. Academic freedom, as we know, must be vigilantly guarded in a pluralistic society -- and never more so than in times when rights in general are threatened. Academic freedom, however, is not license to compromise fairness in grading, and merely because faculty members have been free to hand out upgrades in the past does not grandfather in the contamination of the grading process. Indeed, rather than an incursion into academic freedom, a ban on upgrades represents the recognition of the academic rights of students.

Professors, who were students once themselves, understand the inestimable role grades play in a person's ability to pursue the career of their choice. As we make or break career dreams through the grades we hand out, we must keep in mind that the purity of the process should be sacrosanct.