

Why Is Grading Difficult?

Lawrence M. Aleamoni

Abstract

Grading and more recently grade inflation have been of major concern for most educators. This paper begins by presenting the rationale and methods instructors typically use to determine student grades and concludes with a proposal for an alternate grading system. The alternate grading system proposes a more accurate method of reflecting the quality and quantity of student learning according to instructor and course expectations.

An instructor's decision to teach implies a commitment accurately to evaluate and report the level and quality of student learning. This is perhaps the most demanding task in the teaching-learning process. Grades are typically used as a basis for evaluating student achievement rather than as a means of describing such evaluation. The reason for this confusion is that grades are considered to be: (a) an appropriate means of providing feedback to the student on how he/she is progressing, (b) a fairly reliable and valid index of academic achievement, (c) commonly interpreted and therefore allowing the interchange of relatively standardized information about students between schools, (d) a means of motivating the students to learn, and (e) a basis of sorting and certifying students (Pascal and Geis, 1974).

Grades are judgments that reflect a set of complex and sometimes irrelevant and subjective variables such as attendance, sex, neatness. What goes into a grade varies from one instructor to another as evidenced by the fact that instructors of multi-section courses who may agree on standards of achievement will still differ widely in their distribution of grades (Ericksen and Bluestone, 1971).

To gain some insight into why grading is so difficult for instructors, let us consider the methods they use in assigning grades.

While course grades are sometimes still based exclusively on a single end-of-course examination, by far the more common practice is to combine the grades students earn in several different course activities (e.g., term papers, book reports, oral reports, discussion, quizzes, midterms, and finals).

To determine each student's final grade, the instructor generally performs the following tasks: (a) assesses and grades the student's performance in each of the several different activities, (b) assigns suitable weights to each separate grade, and (c) combines the separate grades into a single grade in a manner that recognizes

their varying weights, their variability, and the practices of the institution (Brown and Thornton, 1971).

To grade the student's performance in each of the several different course activities the instructor typically resorts to either "grading-on-the-curve" or "percent grading."

Grading-on-the-Curve Method

Instructors who employ this method are interested in indicating a student's achievement relative to that of his peers by using normal curve characteristics to determine the percentage of the group to be assigned each grade. One technique of grading-on-the-curve is to assume that the students performance in the classroom will be distributed like a normal curve so that there should be 3 percent A's, 13 percent B's, 68 percent C's, 13 percent D's, and 3 percent F's. Another technique would be to use the mean and standard deviation of the actual distribution of student classroom performance and specify that students with scores of more than 1.5 standard deviations above the class mean will receive A's, those between .5 and 1.5 standard deviations above the class mean will receive B's, and so on (Ebel, 1965). The major difference between these two techniques is that the first one predetermines the percentage of students receiving each grade, whereas the second one does not.

The alleged benefits of grading-on-the-curve (which could also be applied to grading in general) might be stated as follows: (a) since scores based on a "curve" are competitive, they help prepare students for the competition of life; (b) grades motivate the students to work and learn; and (c) grades are a fairly reliable and valid index of academic achievement.

The alleged defects of grading-on-the-curve (as well as grading in general) might also be stated as follows: (a) percentages are arbitrary; (b) different groups can be expected to depart from preconceived distributions; (c) grades are not standardized; (d) grades do not tell anyone (student, employer, etc.) specifically what the student has actually learned or not learned; and (e) grades as rewards promote "grade-getting" behavior.

Percent Grading Method

Instructors who employ the percent method of grading are interested in defining a student's achievement according to some absolute standard by identifying in advance percentage score ranges that will be associated with each grade. For example, a student with a score in the 90 to 100 percent range would qualify for a grade of A, and so on.

Aleamoni is Director of the Office of Instructional Research and Development at the University of Arizona, Tucson, AR 85721.

The alleged benefits of percent grading are similar to those of grading-on-the-curve with the following additions: (a) it clearly relates achievement to degree of mastery of what was set out to be learned; and (b) it provides fixed, standard measures of achievement so that the students are not competing against one another (Ebel, 1965).

Unfortunately, percent grading often fails to live up to its promise of providing truly meaningful and stable measures of achievement, thereby suggesting the following alleged defects: (a) performance standards are typically based on instructor observations of what students typically can do. This means that the standards are more relative than absolute; and (b) if no one falls in the highest score range, does that indicate a problem with the teaching, student learning, or grading system?

An Alternate Grading Method

A grading system cannot be all things to all people. A single symbol cannot represent low achievement from one point of view (i.e., actual degree of subject matter mastery) and high achievement from another (i.e., progress in relation to reasonable expectation). What it can and should have is one clearly defined and scrupulously guarded kind of meaning. College and department faculties have the opportunity and the obligation to establish and maintain clearly defined meanings for the symbols used in their grading systems.

Grades should represent the degree of achievement in the subject matter of the course. This implies an adequate sampling of subject matter on a competitive basis by means of tests or other valid appraisals. Describing student qualities, characteristics, or achievements in terms of various letter grades might be done somewhat as follows (Brown and Thornton, 1971):

- A - Signifies that both major and minor instructional objectives have been achieved and the work is of superior quality. This grade is reserved for outstanding students who are clearly capable of going on to do advanced work in the field.
- B - Major instructional objectives achieved with excellent, above average standards; some minor objectives not achieved; easily capable of doing the next stage of advanced work in the field.
- C - Major instructional objectives achieved with minimum acceptability; many minor objectives not achieved; work of average quality; minimally capable of doing advanced work in the field, with no major handicaps to performance.
- D - Most major objectives not achieved with even limited acceptability; below average work, but above failure.
- E - No major objectives achieved; work of unacceptable quality.

This is substantially more informative than the typical definition of grades which can be found in the latest University of Arizona General Catalog (1977):

A - Excellent B - Good C - Fair D - Poor E - Failure

The following points should be considered if instructors would like to reduce the difficulty encountered in grading:

1. The grading system should put the students in competition with well defined standards of excellence, rather than with each other.
2. The instructor can and should determine at the beginning of a course what he/she expects the students will know and can do when they have completed the course.
3. The instructor should predetermine the percent of material students must demonstrate mastery on for particular grades.
4. Grades ought to reflect how well students have acquired the knowledge they committed themselves to study.
5. Test scores or project scores should not be converted to letter grades before combining to determine the final grade. Such a procedure results in the loss of information. One should convert all scores to standard scores (e.g., T-scores) then weight them and sum them. Grades should then be assigned to the summed and weighted standard score distribution.
6. Do not use the natural breaks in score distributions as justification for assigning different grades because such breaks are due to **chance alone**.

To reduce the difficulty encountered in grading and create a system that accurately reflects the quantity and quality of student learning according to instructor and course expectations, instructors may need to gather and refine their evaluation information for at least one year before implementation. This means that the standards used to arrive at grades should be carefully gauged to the instructional expectations of student performance.

It should be apparent that if we as instructors take the necessary steps to make our grades an accurate representation of student learning outcomes, then such things as "grade inflation" should not occur. "Grade inflation" is largely the result of arbitrary, subjective grading systems easily influenced by irrelevant conditions.

References

- Brown, J. W., and Thornton, J.W. Jr. 1971. *College teaching: A systematic approach* (2nd ed.). New York: McGraw-Hill.
- Ebel, R. L. 1965. *Measuring educational achievement*. New Jersey: Prentice-Hall.
- Ericksen, S.C., and Bluestone, B.Z. 1971. *Grading evaluation* (Memo to the Faculty, No. 46). Ann Arbor, Michigan: University of Michigan, Center for Research on Learning and Teaching.
- Pascal, C.E., and Geis, G. J. March, 1974. *An outline of methods of grading student performance* (Learning and Development, Vol. 5, No. 5). Montreal, Quebec: McGill University, Centre for Learning and Development.