
Some Pertinent Questions About Grading

Marilla D. Svinicki, Ph.D.
Center for Teaching Effectiveness
The University of Texas at Austin

The grading system an instructor selects reflects his or her educational philosophy. There are no right or wrong systems, only systems which accomplish different objectives. The following are questions which an instructor may want to answer when choosing what will go into a student's grade.

1. Should grades reflect absolute achievement level or achievement relative to others in the same class?

This is often referred to as the controversy between norm-referenced versus criterion-referenced grading. In norm-referenced grading systems the letter grade a student receives is based on his or her standing in a class. A certain percentage of those at the top receive A's, a specified percent of the next highest grades receive B's and so on. Thus an outside person, looking at the grades, can decide which student in that group performed best under those circumstances. Such a system also takes into account circumstances beyond the students' control which might adversely affect grades, such as poor teaching, bad tests or unexpected problems arising for the entire class. Presumably, these would affect all the students equally, so all performance would drop but the relative standing would stay the same.

On the other hand, under such a system, an outside evaluator has little additional information about what a student actually knows since that will vary with the class. A student who has learned an average amount in a class of geniuses will probably know more than a student who is average in a class of low ability. Unless the instructor provides more information than just the grade, the external user of the grade is poorly informed.

The system also assumes sufficient variability among student performances that the difference in learning between them justifies giving different grades. This may be true in large beginning classes, but is a shaky assumption where the student population is homogeneous such as in upper division classes.

The other most common grading system is the criterion-referenced system. In this case the instructor sets a standard of performance against which the students' actual performance is measured. All students achieving a given level receive the grade assigned to that level regardless of how many in the class receive the same grade. An outside evaluator, looking at the grade, knows only that the student has reached a certain level or set of objectives. The useful-

ness of that information to the outsider will depend on how much information he or she is given on what behavior is represented by that grade. The grade, however, will always mean the same thing and will not vary from class to class. A possible problem with this is that outside factors such as those discussed under norm-referenced grading might influence the entire class and performance may drop. In such a case all the students would receive lower grades unless the instructor made special allowances for the circumstances.

A second problem is that criterion-referenced grading does not provide "selection" information. There is no way to tell from the grading who the "best" students are, only that certain students have achieved certain levels. Whether one views this as positive or negative will depend on one's individual philosophy.

An advantage of this system is that the criteria for various grades are known from the beginning. This allows the student to take some responsibility for the level at which he or she is going to perform. Although this might result in some students working below their potential, it usually inspires students to work for a high grade. The instructor is then faced with the dilemma of a lot of students receiving high grades. Some people view this as a problem.

A positive aspect of this foreknowledge is that much of the uncertainty which often accompanies grading for students is eliminated. Since they can plot their own progress toward the desired grade, the students have little uncertainty about where they stand.

2. Should grades reflect achievement only or nonacademic components such as attitude, speed and diligence?

It is a very common practice to incorporate such things as turning in assignments on time into the overall grade in a course, primarily because the need to motivate students to get their work done is a real problem for instructors. Also it may be appropriate to the selection function of grading that such values as timeliness and diligence be reflected in the grades. External users of the grades may be interpreting the mark to include such factors as attitude and compliance in addition to competence in the material.

The primary problem with such inclusion is that it makes grades even more ambiguous than they already are. It is very difficult to assess these nebulous traits accurately or consistently. Instructors must use real caution when incorporating such value judgments into final grade assignment. Two steps instructors

should take are (1) to make students aware of this possibility well in advance of grade assignment and (2) to make clear what behavior is included in such qualities as prompt completion of work and neatness or completeness.

3. *Should grades report status achieved or amount of growth?*

This is a particularly difficult question to answer. In many beginning classes, the background of the students is so varied that some students can achieve the end objectives with little or no trouble while others with weak backgrounds will work twice as hard and still achieve only half as much. This dilemma results from the same problem as the previous question, that is, the feeling that we should be rewarding or punishing effort or attitude as well as knowledge gained.

A positive aspect of this foreknowledge is that much of the uncertainty which often accompanies grading for students is eliminated. Since they can plot their own progress toward the desired grade, the students have little uncertainty about where they stand.

There are many problems with "growth" measures as a basis for change, most of them being related to statistical artifacts. In some cases the ability to accurately measure entering and exiting levels is shaky enough to argue against change as a basis for grading. Also many courses are prerequisite to later courses and, therefore, are intended to provide the foundation for those courses. "Growth" scores in this case would be disastrous.

Nevertheless, there is much to be said in favor of "growth" as a component in grading. We would like to encourage hard work and effort and to acknowledge the existence of different abilities. Unfortunately, there is no easy answer to this question. Each instructor must review his or her own philosophy and content to determine if such factors are valid components of the grade.

4. *How can several grades on diverse skills combine to give a single mark?*

The basic answer is that they can't really. The results of instruction are so varied that the single mark is really a "Rube Goldberg" as far as indicating what a student has achieved. It would be most desirable to be able to give multiple marks, one for each of the variety of skills which are learned. There are, of course, many problems with such a proposal. It would complicate an already complicated task. There might not be enough evidence to reliably grade any one skill. The "halo" effect of good performance in one area could spill over into others. And finally, most outsiders are looking for only one overall classification of each person so that they can choose the "best." Our system requires that we produce one mark. Therefore, it is worth our while to see

how that can be done even though currently the system does not lend itself to any satisfactory answers.

In Summary

The process of deciding on a grading system is a very complex one. The problems faced by an instructor who tries to design a system which will be accurate and fair are common to any manager attempting to evaluate those for whom he or she is responsible. The problems of teachers and students with regard to grading are almost identical to those of administrators and faculty with regard to evaluation for promotion and tenure. The need for completeness and objectivity felt by teachers and administrators must be balanced against the need for fairness and clarity felt by students and faculty in their respective situations. The fact that the faculty member finds himself or herself in both the position of evaluator and evaluated should help to make him or her more thoughtful about the needs of each position.
