GRADING SYSTEMS

The two most common types of grading systems used at the university level are norm-referenced and criterion-referenced. Many professors combine elements of each of these systems for determining student grades by using a system of anchoring or by presetting grading criterion which is later adjusted based on actual student performance.

1. Norm-Referenced Systems:
Definition: In norm-referenced systems students are evaluated in relationship to one another (e.g., the top 10% of students receive an A, the next 30% a B, etc.). This grading system rests on the assumption that the level of student performance will not vary much from class to class. In this system the instructor usually determines the percentage of students assigned each grade, although this percentage may be determined (or at least influenced) by departmental expectations and policy.

Advantages:
♦ Norm-referenced systems are very easy to use.
♦ They work well in situations requiring rigid differentiation among students where, for example, program size restrictions may limit the number of students advancing to higher level courses.
♦ They are generally appropriate in large courses that do not encourage cooperation among students but generally stress individual achievement.

Disadvantages:
♦ One objection to norm-referenced systems is that an individual's grade is determined not only by his/her achievements, but also by the achievements of others. This may be true in a large non-selective lecture class, where we can be fairly confident that the class is representative of the student population; but in small classes (under 40) the group may not be a representative sample. One student may get an A in a low-achieving section while a fellow student with the same score in a higher-achieving section gets a B.

♦ A second objection to norm-referenced grading is that it promotes competition rather than cooperation. When students are pitted against each other for the few As to be given out, they're less likely to be helpful to each other.

Possible modification:
When using a norm-referenced system in a small class, the allocation of grades can be modified according to the caliber of students in the class. One method of modifying a norm-referenced system is anchoring. Jacobs and Chase in Developing and Using Tests Effectively: A Guide for Faculty, 1992, describe the following ways to use an anchor:

"If instructors have taught a class several times and have used the same or an equivalent exam, then the distribution of test scores accumulated over many classes can serve as the anchor. The present class is compared with this cumulative distribution to judge the ability level of the group and the appropriate allocation of grades. Anchoring also works well in multi-section courses where the same text, same syllabus, and same examinations are used. The common examination can be

used to reveal whether and how the class groups differ in achievement and the grade in the individual sections can be adjusted accordingly... If an instructor is teaching a class for the first time and has no other scores for comparison, a relevant and well-constructed teacher-made pretest may be used as an anchor."

Modifying the norm-referenced system by anchoring also helps mitigate feelings of competition among students as they are not as directly in competition with each other.

Before Deciding on a Norm-Reference System, Consider:
♦ What is the expected class size? If it is smaller than 40, do not use a norm-referenced system unless we use anchoring to modify the system.

♦ Is it important for students to work cooperatively in this class (e.g., do we ask them to form study groups, or work on projects as a group)? If the answer is yes, a norm-referenced system is not appropriate for the class.

2. Criterion-Referenced System
Definition: In criterion-referenced systems students are evaluated against an absolute scale (e.g. 95-100 = A, 88-94 = B, etc.). Normally the criteria are a set number of points or a percentage of the total. Since the standard is absolute, it is possible that all students could get As or all students could get Ds.

Advantages:
♦ Students are not competing with each other and are thus more likely to actively help each other learn. A student's grade is not influenced by the caliber of the class.

Disadvantages:
♦ It is difficult to set reasonable criteria for the students without a fair amount of teaching experience. Most experienced faculty set these criteria based on their knowledge of how students usually perform (thus making it fairly similar to the norm referenced system).

Possible modifications:
♦ Instructors sometimes choose to maintain some flexibility in their criteria by telling the class in advance that the criteria may be lowered if it seems appropriate, e.g., the 95% cut off for an A may be lowered to 93%. This way if a first exam was more difficult for students than the instructor imagined, s/he can lower the grading criteria rather than trying to compensate for the difficulty of the first exam with an easy second exam. Raising the criteria because too many students achieved As, however, is never advisable.

♦ Another way of doing criterion-referenced grading is by listing objectives and assigning grades based on the extent the student achieved the class objectives (e.g., A = Student has achieved all major and minor objectives of the course. B = Student has achieved all major objectives and several minor objectives, etc.).

Before Deciding on a Criterion-Reference System, Consider:
♦ How will we determine reasonable criteria for students? When teaching the class for the first time, maintain some flexibility.

3. Other Systems: Some alternate systems of grading include contract grading, peer grading, and self-evaluation by students.

♦ In contract grading instructors list activities students can participate in or objectives they can achieve, usually attaching a specified number of points for each activity (e.g. book report = 30 points, term paper = 60 points). Students select the activities and/or objectives that will give them the grade they want and a contract is signed. It is advisable to have qualitative criteria stated in the contract in addition to listing the activities.

♦ In some classes, a portion of a student's grade is determined by peers' evaluation of his/her performance. If students are told what to look for and how to grade, they generally can do a good job. Agreement between peer and instructor rating is about 80%. Peer grading is often used in composition classes and speech classes. It can also be a useful source of information for evaluating group work; knowing that group members have the opportunity to evaluate each other's work can go a long way in motivating peers to pull their weight on a project and to reassure group members that their contributions will be recognized. If used, peer evaluation should always be done anonymously.

♦ Students can also be asked to assess their own work in the class and their assessment can be a portion of the final grade. This method has educational value as learning to assess one's own progress contributes to the university's goal of preparing our students to be life-long learners. A research analysis found that the percentages of self-assessors whose grades agree with those of faculty graders vary from 33% to 99%. Experienced students tend to rate themselves quite similarly to the faculty while less experienced students generally give themselves higher grades than a faculty grader. Students in science classes also produced self-assessments that closely matched faculty assessment. Not surprisingly, student and instructor assessments are more likely to agree if the criteria for assessment have been clearly articulated. Without these shared understandings, students, for example, don't know whether to assess themselves on the amount of work they put into a course, on the improvement they've seen in certain skills, or on their final level of achievement. If self-assessment is used, the instructor and student should meet to discuss the student's achievement before the self-evaluation is made.