method which allows students to learn whenever they wish using mediated instruction such as audio tapes and slides. This approach has been used in biology instruction primarily, and is discussed in S.N. Postlewaite, J.D. Novak, and H. Murray, **An Integrated Experience Approach to Learning with Emphasis on Independent Study** (Minneapolis: Burgess Publishing Company, 1964). Many of the principles suggested in the above approaches can be modified to fit your particular situation.

Evaluation/Revision

We have ended where we started. The next step is to evaluate the impact of the new strategy in terms of student learning, student attitude, demands on your time, costs, and a range of variables which may reflect the degree to which your development effort has succeeded. These are the variables discussed in the first section. The aim of this activity is to suggest further ways in which instruction can be improved. In fact, this process should continue indefinitely! As evaluation points up new areas of revision, make those revisions and assess their effectiveness. As you assess the effectiveness of the revisions, more areas of improvement will be suggested. This process reflects the fact that teaching and instructional development are systematic processes which can be examined at critical points.

Issues in Formulating Course Grading Policies

David A. Frisbie

Abstract

Assigning course grades to students is a component of the instructional process which must be defensible to both students and colleagues. This paper identifies issues significant in grading and explores the rationale for accepting or rejecting each.

Evaluating student performance is a necessary component of the instructional process, and grading student performance is one widely-used method of providing students with feedback about their progress. If this feedback system is to operate successfully, students must know which aspects of their performance will be evaluated and what standards will be used to make the judgments. This important information must be communicated to students at the outset of the course. With these basic premises in mind, what should a faculty member consider in building a personal philosophical basis for grading and arriving at a sound methodological framework for translating that philosophy into action?

The issues which contribute to making the grading of student achievement a controversial topic are primarily philosophical in nature. There are no empirical studies that can answer questions like: What should an "A" grade mean? What percent of the students in my class should receive a "C"? Should spelling and grammar be judged in assigning a grade to a paper? What should a course grade represent? These "should" questions require value judgments rather than an interpretation of empirical data; the answer to each will vary from instructor to instructor. But each instructor must ask these questions and find acceptable answers to them in establishing his/her own grading philosophy. It is not sufficient to have a method of assigning grades; the method used must be defensible to the user in terms of her/his philosophical foundations.

Dr. Frisbie is supervisor of Placement and Proficiency Testing, Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign.

This paper identifies many of the significant issues in grading and explores the rationale for accepting or rejecting the pros and cons of each. Through this process an instructor might arrive at a systematic and cohesive set of values which could provide the foundation for using particular procedures in assigning course grades to students.

An instructor's philosophy is influenced by many factors; as these factors change there may be a corresponding change in philosophy. The type of instructional strategy used in teaching dictates, to some extent, the type of grading procedures to use. For example, a mastery learning approach to teaching is incongruent with a grading approach which is based on competition for an arbitrarily set number of "A" or "B" grades. Grading policies of the department, college, or campus may limit the procedures which can be used and force a basic grading philosophy on each instructor in that administrative unit. The recent response to grade inflation has caused some faculty, individually and collectively, to alter their philosophies and procedures. Pressure from colleagues to give lower or higher grades often causes some faculty members to operate in conflict with their own philosophies. Student grade expectations and the need for positive student evaluations of instruction probably both contribute to the shaping or altering of the grading philosophies of some faculty. The dissonance created by institutional restraints probably contributes to the wide-spread feeling that end-of-course grading is one of the most dreaded tasks facing a college instructor.

What Meaning Should a Course Grade Carry?

The meaning that a given grade represents seems to vary substantially between instructors within the same department, college, or institution. This variability tends to deflate the value of course grades as an index of quality of performance; the interpretation of a course grade or comparisons between grades is ambiguous. What does it

mean if a student receives a "B" in Soil Chemistry? How should one interpret a student's grade-point average of 3.8 in agricultural economics courses? This issue can be addressed by examining four possible meanings a grade might represent and by weighing the severity of the problems associated with using each.

1. **A grade might represent growth** — the knowledge and skill a student possessed at the end of Soil Chemistry compared to his level of achievement at the beginning of the course. Large gains are assigned high grades and small gains are represented by low grades. Students who enter the course with some knowledge of the reactions in soil development are obviously penalized; they have less to gain from the course than a relatively naive student. The post-test-pretest gain score is more error-laden, from a measurement perspective, than either of the scores from which it is derived. Though growth or amount of learning is important, it offers very little as a basis for determining course grades. The educational value of such grades in a college-level course is probably minimal.

2. **A course grade might represent the amount a student learned in a course relative to what we would expect him/her to learn, based on his/her measured academic ability.** Students with high ability scores (e.g., scores on the Scholastic Aptitude Test (SAT) or American College Test (ACT) would be expected to achieve higher final examination scores than those with lower ability scores. The terms "over-achiever" and "under-achiever" are derived from this ability-based representation of grades. An "over-achiever" and an "under-achiever" may receive the same grade in a Farm Taxation course, yet their levels of competence with respect to tax matters may be vastly different. What must be assessed, then, is what information the course grade should communicate. The first student might not be prepared for an additional course on taxation, but the second student may be. Even when the statistical and clerical burden of ability-based grading are set aside, the educational value seems not to warrant its use.

3. **Grades can take on a norm-referenced meaning.** By comparing a student's end-of-course performance with that of some relevant group of students, the instructor assigns a grade to show the student's level of achievement or standing within that group. An "A" in Livestock Management would mean that the student's performance was as high as the best students in the group with which the instructor made the comparison. But what group did the instructor use to make the normative comparison? This problem is probably the greatest source of misinterpretation of grades based on norm-referencing. A grade of "A" might not represent "adequate" knowledge and skill if the reference group is somewhat inept. The nature of the reference group is the key to interpreting such grades. The norm group might be all students enrolled in Livestock Management during a given semester, all students who enrolled in the course over the past two or three semesters, or all students who were enrolled since the instructor first began teaching the course. A norm-referenced grade remains ambiguous until the norm group has been sufficiently described by the instructor.

4. **Grades might be defined in terms of absolute rather than relative standards of performance.** There may be more educational value in indicating what domain of tasks a student has command of rather than what portion of a particular group a student has exceeded in his performance within that domain. Absolute standard grading requires that performance standards or criteria be set by the instructor for each grade category. A "C" in Introduction to Dairy Production might indicate that the student has only minimal (but passing) knowledge about breeds of cattle, herd management, and the scope of the dairy industry. A much higher achievement level would be required for a student to earn an "A". Note that a student's grade depends only on his level of content mastery; how well his/her classmates perform has no bearing on the course grade he/she earns. There are no quotas in each grade category because it could be possible that no students in a given class would receive an "A" or a "B", etc.

Grades which reflect growth or achievement relative to measured ability have little value in post-secondary education. Norm-referenced grading, currently the most frequently used, seems less appropriate for most educational contexts than it has in the past. Greater emphasis is being placed on helping students achieve **at least** a minimally acceptable level of competence with regard to a group of instructional goals. There is less attention paid to "insuring" that a certain portion of each class receive a "D" or "F" grade. Absolute standard grading removes the focus from inter-student competition to intra-student competition, a focus on the domain of skills or tasks to be mastered by the learner as determined by the instructor. (For additional reading on this topic see Ebel, pp. 328-31; Gronlund, pp. 525-7; Mehrens and Lehmann, pp. 591-4; Terwillinger, pp. 26-75.)

What Aspects of Performance Should be Evaluated?

A distinction should be made between factors which an instructor evaluates and factors which are used to determine course grades. Factors or variables which contribute to determining course grades should reflect each student's competence in the course content. The components of a grade should be academically oriented; they should not be tools of discipline or awards for pleasant personalities or merely "good" attitudes. A student who gets an "A" in a beef production course should have a firm grasp of feeding, management, and breeding principles. If he is merely marginal academically but very industrious and congenial, an "A" grade would mislead the student and render a blow to the motivation of the excellent students in the program. Instructors can give feedback to students on many traits or characteristics, but only academic performance factors should be used in determining course grades.

Some Evaluation Variables

Some potentially invalid grading components are considered below. Though some exceptions could be noted, the variables considered below generally should **not** be used to determine course grades.

Class attendance. Students should be encouraged to attend class meetings because it is assumed that the lectures, demonstrations, and discussion will facilitate their learning. If a student misses several classes then his performance on examinations, papers, and projects will likely suffer. If the instructor further reduces his/her course grade because of absence, the instructor is essentially submitting the student to "double jeopardy." An instructor may say that attendance counts ten percent, but for some students this may in effect amount to 20 percent. Teachers who experience a good deal of class "cutting" might examine their classroom environment and methods to determine if changes are needed.

Class participation. Obviously seminars and small classes depend on student participation to some degree for their success. When participation is important, it may be appropriate for the instructor to use participation grades. In such cases the instructor should keep notes regarding frequency and quality of participation; waiting until the end of the semester and relying on memory makes a relatively subjective task even more subjective. Participation should probably not be graded in most courses, however. Dominating or extroverted students tend to win and introverted or shy students tend to lose. Students should be graded in terms of their achievement level, not in terms of their personality type. Instructors can give feedback to students on many dimensions, but grading should not be the only means of doing so.

Mechanics. Neatness in written work, correctness in spelling and grammar, and organizational ability are all worthy traits. They are assets in most vocational endeavors. To this extent it seems appropriate that instructors evaluate these factors and give students feedback about them. However, unless the course objectives include instruction in these skills, students should not be graded on them in the course. A student's grade on an agronomy essay exam should not be influenced by his/her general spelling ability, neither should his/her course grade.

Personality Factors. Most of us are attracted to students who are agreeable, friendly, industrious, and kind; we tend to repell those with opposite characteristics. To the extent that certain personalities may interfere with class work or have limited chances for employment in their field of interest, constructive feedback from the instructor may be necessary. An argumentative student who earns a "C" in Principles of Nutrition should have minimal knowledge about the need for various nutrients at all stages of the life cycle. The nature of his/her personality should not have direct bearing on the course grade earned.

Instructors can and should evaluate many aspects of student performance in their course. Only the evaluation information which relates to course goals should be used to assign a course grade, however. Judgments about writing and speaking skills, personality traits, effort, and motivation should be communicated in some other form. Some faculty use brief conferences for this purpose, others communicate through comments written on papers or through the use of mock letters of recommendation.

Some Grading Variables

How an instructor evaluates the achievement of his/her students depends primarily on the nature of the course goals. Examinations (objective, essay, and performance), reports (oral or written), and projects (demonstrations, products, simulations, and model construction) are all typical forms of evaluating achievement. The usefulness of each form varies according to its fit with course objectives, efficiency for instructor and students, and accuracy.

Examinations. Though tests cannot be used to measure the achievement of all types of course objectives, they probably are the most appropriate, accurate, and efficient means of evaluating most forms of student performance. Objective tests can be used when students must recognize, identify, differentiate, or choose among specific alternatives. When students should describe, explain, construct, define, or list characteristics, essay or short answer tests seem more appropriate. Performance tests can be used if students must classify, identify, demonstrate, construct, or simulate. The wording of course objectives is probably the best indicator of the type of test or combination of tests needed for evaluation. Finally, it is important to recognize that somewhat different test construction procedures are required when grading is norm-referenced vs. criterion-referenced.

Papers and reports. Student achievement can be measured by evaluating papers and presentations which require students to integrate concepts and principles, summarize readings or research, or explain and diagram the functioning of equipment, mechanical systems, or production methods. A major evaluation problem of the instructor who requires papers and reports is the separation of content and mode of presentation. The mechanics of writing or speaking might be evaluated, but the content of the paper or report must be evaluated separately for grading purposes.

Projects. Some of the evaluation problems inherent in the use of papers and reports are associated with class projects as well. The need to separate content and quality of presentation is sometimes less severe with projects because the act of presenting is oftentimes a part of the instructional goal. The organization and mode of reporting a cost analysis of egg production may be nearly as important as the analytic procedures themselves. A model layout of a swine feeder lot cannot be evaluated only on principles of efficiency or economy; the quality of the construction of the components and the appearance contribute to the whole effect. Some instructors encourage students to blend their efforts in completing a group pro-

ject. Though group projects can often be productive, the additional burden on the instructor of separating individual contributions makes evaluation more complex. Perhaps the complexities involved in evaluating projects explain why project grades seem to be fairly homogeneous within a class.

Conclusions

Assigning course grades to students is a component of the instructional process which must be as defensible to students and colleagues as any other component. A philosophical, theoretical, and practical base must be established by the instructor so that his grading procedures have credibility. Without such a foundation, the various meanings which can be associated with a grade become muddled and communication fails.

Instructors must be able to separate behaviors to be judged for grading purposes from the set of behaviors they wish to evaluate. Students need feedback on many dimensions of their achievement, yet only those dimensions which relate closely to the course goals should be used as the basis for grading.

References

Ebel, Robert L. 1972. *Essentials of educational measurement.* Prentice-Hall, Englewood Cliffs, NJ.

Gronlund, Norman E. 1976. *Measurement and evaluation in teaching.* Macmillan Publishing, New York, NY.

Mehrens, William A. and Irvin J. Lehmann. 1973. *Measurement and evaluation in education and psychology.* Holt, Rinehart and Winston, New York, NY.

Terwillinger, James S. 1971. *Assigning grades to students.* Scott Foresman, Glenview, Ill.

what research says about LEARNING

Cameron Fincher

There is a current concern with the improvement of college instruction that has not been evident in the past. Most of this concern takes the form of evaluation and comes under the rubric of accountability — an old concept with a new harshness that has grown out of the public's vague but intense dissatisfaction with education. If the Puritans were the first to establish a legal requirement that children be taught to read, they were not the last to assign broad, extensive responsibilities to the public school and to express high expectations for the performance of teachers.

Yet, the notion that teaching is to be evaluated is an uncomfortable one in higher education, and somehow the notion persists that it cannot be evaluated. Unlike the physician whose mistakes are buried, the lawyer whose mistakes go to jail, or the minister whose failures become town gossip, the college instructor is thought to remain unaffected by the mistakes of his profession because the continued ignorance of his students is easily concealed. If exposed, there are always reasons why the student failed to gain the knowledge or skill the instructor presumably tried to instill. It is significant, therefore, that a recent Gallup Poll has shown that the public still regards the college teacher as having high standards. Only physicians and engineers were seen by more people as having high standards.

Neither accountability nor evaluation is the threat that some college professors perceive them to be. Accountability is a much broader concept then evaluation and should imply a larger sense of responsibility on the part of the teaching profession at all levels. Moreover, it should imply an acceptance and a responsiveness by the teaching faculty rather than a superimposition by the public or its representatives. In any event, it should imply a greater concern with specific, identifiable, constructive changes in the lives of students.

It is in the latter sense that converging concepts of accountability may be detected. There has been an obvious concern with: (1) behavioral objectives in instruction with measurable outcomes, (2) criterion-referenced measurement as opposed to the more traditional norm-referenced tests, and (3) a general systems approach to collegiate instruction at the two-year and four-year levels. These suggest, in turn, a pervasive interest in making college instruction both more generally effective and more demonstrable. In other words, there is not only an increasingly serious attempt to improve college instruction but an effort to demonstrate that effectiveness.

Some of the more obvious indications of this thrust are: (1) The Project to Improve College Teaching, sponsored by the American Association of University Professors and the Association of American Colleges —, with Kenneth Eble's report, **Professors as Teachers,** (2) the emergence of faculty development as a popular theme — as shown by Jerry Gaff's Toward Faculty Renewal, (3) one book published by the American Council on Education in 1967 entitled **Improving College Teaching** — still another in 1970 entitled **Effective College Teaching,** and (4) other serious efforts to deal directly with the subject such as Ohmer Milton's **Alteratives to the Traditional,** Brown and Thornton's **College Teaching: A Systematic Approach** that came out in a second edition in 1971, and Pat Cross's more recent **Accent on Learning.** Add to these the 1400 page **Second Handbook of Research on Teaching** and the 75th yearbook of NSSE, **The Psychology of Teaching Methods,** and you have a better indication of the concern and interest in teaching effectiveness.

This article is adapted from the *Institute of Higher Education Newsletter* Athens, GA, January, 1977. Dr. Cameron Fincher is professor of Higher Education and Psychology, and director of the Institute of Higher Education at the University of Georgia. We are indebted to Robert S. Wheeler, director of Resident Instruction at the University of Georgia College of Agriculture for encouragement of the author to make this manuscript available to NACTA Journal.