



I  
E  
A



## Grading and Distributive Justice

Karl A. Smith  
Associate Professor  
Mineral Resources Research Center  
University of Minnesota Minneapolis,  
Minnesota

"For so it is, O Lord my God I measure it!  
But what it is measure, I do not know."  
| St. Augustine

Grading is frequently cited as the most difficult and least desirable task a teacher faces, yet, we continue to shell out grades at regular intervals. What do grades mean; what are the value bases for the grades we give? How do grades fit with instructional objectives? What do faculty, students, parents; and business recruiters think of grades?

This brief paper addresses the question of the value bases for grading, discusses the relationship between grading and the goals of instruction, and summarizes the results and recommendations of a recent national survey of grading in college.

### Value Bases for Grading

There are many possible bases for distributing grades to students. The teacher might, for example, give grades so as to produce equal outputs of student motivation, giving a high-ability student a lower grade than his or her performance warranted and vice versa for a low-ability student. Or the teacher might decide to give high grades to those who need them the most, for example, to those who need them to retain athletic eligibility. Or the teacher might decide to give them to those who work the hardest, those who accomplish the most, or those who help others the most.

These different values may conflict with one another: the most needy may not be the most able; those who work the hardest may not accomplish the most; giving everybody equal inputs may not result in their having equal outputs. Claims for inherent or natural priority among these values raise questions of distributive justice.

The concept of distributive justice centers on the fairness of the distribution of the conditions and goods that affect individual well-being<sup>1</sup>. Deutsch has listed nine characteristics and questions to identify the key features of any system of distributive justice (such as grading). Examples of the distribution of grades are also given.

1. The nature of the good or harm being distributed: its content, quality, and quantity. What is the meaning and value of a given grade? Is there a limited supply of high grades available so that only a few students can get high grades no matter how many excellent students there are?

2. Roles involved. To whom is the grade being distributed and by whom? Grades are usually given by teachers to students but there are other possibilities, such as, students grading one another or students and teachers grading each other.

3. Styling and timing of the distribution. How and when is the grade distributed? Secretly or publically? With or without explanation of its meaning and possible consequences? Some teachers provide grades to their students only at the end of a term; others provide their students with a more continuous evaluation of their work.

Sometimes students are given information only about their own mark; other times students have access to information about all students. In some courses, the factors determining a student's grade are clearly defined, whereas in others the whole procedure is unclear.

4. Values. What are the values underlying the distribution? A number of key values have been identified as the substantive values underlying the distribution of grades<sup>2,3</sup>. Justice has been viewed as the treatment of all people:

- (a) so that they can have equal "inputs" (for example, so that all students have equal educational resources available;
- (b) so that they have equal "outputs" (for example, so that all students have the resources necessary to enable them to achieve a given level of educational attainment even if some students require more inputs than others);
- (c) according to their needs;
- (d) according to their ability or potential;
- (e) according to their efforts and sacrifices;
- (f) according to their performance or improvement in performance;

- (g) according to the social value of their contribution;
- (h) according to the requirements of the common good;
- (i) so that none fall below a certain minimum;
- (j) according to what others choose to do for them; and
- (k) according to the principle of reciprocity.

5. Rules or criteria for defining the values. Suppose, for example, it is widely accepted that "merit" should be the basis for distributing "grades"; what are the criteria for defining "merit"? Shall it be defined in terms of the quality or quantity of work, in terms of the actual accomplishment, in terms of the improvement over prior accomplishment? Whether the distributive value is "merit," "equality," "effort," "ability," "need," or almost any other value, for this value to be implemented, there has to be specification of which of the relevant potential criteria shall be employed to represent the particular value.

6. Measurement procedures. How are the rules or criteria operationalized and implemented? The implementation of the criteria, the measurement, may be invalid (for example, instead of testing "ability," one measures "effort"); it may be unreliable, so that one can have little confidence that similar results would be obtained if the measurement procedures were repeated; or it may be insensitive, so that differences in the amounts of the value cannot be readily distinguished.

7. Decision-making procedures. What are the procedures for making decisions about any of the foregoing? Even if one feels that a distribution is fair, one may feel an injustice with regard to the procedures by which the values, the rules, or the measurement were determined.

8. Scope of the moral community. What is the scope of the distribution system? To whom does it apply? Who is included, who is excluded?

9. Effects. What are the effects of the distribution system on individuals within it, on different categories or groups of individuals, on the interrelations among different individuals and groups, and on the cohesion and productivity of the entire system?

Deutsch's theoretical analysis suggests that the distributive values operative within a group or society will and should depend on circumstances: Under some conditions, distributing goods according to individual need will be more just, and under other conditions, allocating goods in terms of individual productivity will be more so. Furthermore, Deutsch's analysis suggests that merit based on individual performance, i.e., equity, will be the dominant principle of distributive justice

in situations where an economic orientation predominates; equality will be the dominant principle in situations where a solidarity orientation predominates; need will be dominant in caring-oriented groups or institutions.

#### Grading and the Goals of Instruction

Wilbert McKeachie, Director of the Center for Teaching and Learning at the University of Michigan and author of Teaching Tips for the Beginning College Teacher, specifically addresses the problem of testing and grading in achieving the goals of instruction. He writes<sup>4</sup>:

"The standards stated as desirable of testing and grading--objectivity, reliability, coverage of content, and flexibility--are not nearly as important as the effect of the testing and grading procedures on student learning. For example, objective tests are likely to result in an objective and reliable test scoring, but their effect on student learning is to encourage inefficient learning strategies such as repetition and rote memorization, strategies that not only are likely to result in less retention of the subject matter but also result in less ability to use material for problem solving and effective learning in later courses or after leaving school. It seems to me that teachers' most important objective is to develop students' motivation and skills for continued learning, problem solving, and application of course learning after the course is over. Whatever grading or testing system one uses should, therefore, put these objectives as primary rather than the objectives of having a reliable test or an objective system of grading."

Alexander Astin<sup>5,6</sup>, creator of the talent development model of excellence, lists three key aspects for the development of excellence--student involvement, high expectations, and assessment and feedback. He claims that the main assessment problem is that most assessment is done for passive, evaluative purposes--to sort, review, and classify faculty and students--rather than to inform and enlighten them. His theory of involvement suggests that assessment should be used primarily as feedback, to increase the involvement of students and faculty members and to develop their talents as fully as possible. Such assessment is active rather than passive, since it is designed to facilitate and improve performance.

#### Effects of the Distribution

What are the effects of the distribution of grades on students' motivation, individual productivity, individual learning, and attitudes toward work or learning? The effect of grading is difficult to assess because the meaning of a grade is dependent on the context. Some factors that

moderate the meaning include: (1) the educational institution at which it is given, (2) the academic discipline in which it is given, (3) the instructor giving it, and (4) the student receiving it.

A few effects are clear, however. Educational Psychologist and expert on motivation, Edward Deci, says there are enormous differences between being "intrinsically motivated"--having the inner desire to be effective--and being "extrinsically motivated"--being pushed by external rewards and punishments to achieve. He comments on the undesirability of excessive external motivators such as grades as follows<sup>7</sup>:

"The frightening thing about the recent rhetoric advocating higher standards is its heavy emphasis on control. I, too, would like to see greater excellence in our educational system, but to get it, I think we need to support systems that encourage teachers to be innovative and self-determining and to promote innovation and self-determination in their pupils. By pushing harder with procedures such as standardized curriculum and competency tests, we are likely to end us with less excellence."

Stice<sup>8</sup> concluded that standardized tests and GPA's predict academic success but not much else. SAT and ACT scores predict freshman grades somewhat (correlation coefficient,  $r$ , varies between .30 and .39). The GRE does not predict graduate grades as well as the SAT predicts undergraduate grades, but the GRE is a better predictor of graduate GPA than is undergraduate GPA. These measures do not, however, correlate with significant adult accomplishments. Testing and grading do affect teaching and learning. Fredericksen<sup>9</sup> presents evidence that tests do affect teacher and student performance and that multiple-choice tests tend not to measure more complex cognitive abilities. Efficient tests (such as the more economical multiple-choice tests) tend to drive out less efficient tests, leaving many important abilities untested--and untaught. Finally, Deutsch's<sup>10</sup> research has shown that there is no consistent or reliable evidence to indicate that people work more productively as individuals or as group members where they are expecting to be rewarded in proportion to their performance than when they are expecting to be rewarded equally or on the basis of need. Deutsch concludes his essay on grading as follows:

"If the competitive grading system in our schools--a less corrupted version of a competitive merit system than the one that characterizes our larger society--does not foster a social environment that is conducive to individual well-being and effective social cooperation, why would one expect that such values would be fostered in a society that is dominated by a competitive, meritocratic ideology? If the competitive-hierarchical atmosphere is not good for our children, is it good for us?"

### National Grading Survey

A national survey was conducted of over 6000 students, faculty, parents, and business recruiters regarding their current thinking about grading in college<sup>11</sup>. Schools were selected to participate in the survey on the basis of two factors: type--community colleges, comprehensive universities, four-year colleges, highly selective four year colleges and highly selective universities--and location--Northeast, Southeast, Central, Southwest, and West Coast/Northwest. Abbreviated results from the first twelve questions asked of all respondents are as follows:

1. Over 90 percent in each group believed that grades are primarily intended for students.
2. Around three fourths of all groups preferred one of three grading systems that provide numerically quantifiable differentiation--A+, A, A-, 8+, so forth; percentages (0-100); A, B, C, D, F).
3. Approximately one half the participants in each group believed that there is a moderate relationship between grades and later success.
4. Responses varied among groups (students, parents, faculty, recruiters) on the question "Two students in the same class got different grades from their instructor--one received an A while the other received a C. How long do you think the difference in knowledge (achievement, learning, performance, etc.) represented by the grades A and C will last?" These responses show that grades, when considered as an evaluation of performance, have decidedly different meanings for different groups of people.
5. Between 78 and 85 percent endorsed the purpose of providing other educational institutions with information for making decisions about a student as being of major or moderate importance. The clear majority of participants in the survey forcefully suggest that grades ought to refer more to educational process and less to their evaluative use by society.
6. The fixed levels (100 - 90 = A) procedure was preferred by more than half the faculty and students, and two thirds of the parents and recruiters, over normal curve inspection of scores, student self-grading, personal change, amount of work completed, and students grading other students.
7. In describing personal feelings about each grade, A to F, the results suggest that, relative to other groups, parents overvalue all grades, students undervalue all grades (particularly the lower ones), faculty undervalue higher grades and find lower grades "not so bad," and business people tend to have



the most consistent valuation of grade worth across the complete range of scores.

8. Half the students, faculty and parents believed the professoriate should continue its emphasis on grades "as is." Over half the students, parents, and recruiters thought students should emphasize grades more; only one third of the faculty thought this. Almost half the faculty wanted students to emphasize grades less.

9. Over one third of all groups acknowledged having cheated to get a better grade and slightly over half of the current students acknowledged doing this.

10. Across all four groups, 90 percent mentioned dropping (or switching to audit) a course between one and five times because they were afraid of getting a poor grade.

11. Faculty were more concerned about the value of their symbols (grade inflation) than were any of the remaining groups.

12. Among the top four characteristics (out of thirteen listed) for all four groups indicating the usefulness of student grades or GPA were: academic achievement, ability to get grades, motivation, and self-discipline. Surprisingly, business people values grades as indicators of little other than the ability to get grades and tend to undervalue them as indicators of more broad-ranging personal abilities and skills.

Based on the results of this survey and an exhaustive literature review, the authors of this national grading study made five recommendations in response to the question "What changes do you recommend once we decide to employ grades in the service of learning?"

1. Clarify what we want grades to do. Are they to serve the purpose of promoting learning and teaching, or are they to serve the purpose of rank ordering students? If we select the former, tests and grades are in the service of teaching and learning. If we select the latter, tests and grades will continue as exercises in ranking, not teaching and learning.

2. Improve classroom tests. We must improve the quality of classroom tests so that whatever purpose tests and grades serve for us will be fairly and properly implemented. Test questions should be written more clearly than they are and should seek more than isolated factual information.

3. More and better information for students. We must supply considerable information to students (far more than the letter symbol) about their performance on course tests and other academic exercises.

4. Reduce the number of grade categories. Let us use fewer, rather than more, differentiated grading systems and let us not reify grades or any other metric used to describe academic performance. The perspective should be that grades are not more precise than the techniques used to create them; as it now stands, such techniques are relatively less precise than the metrics by which they are quantified.

5. Abolish the GPA. Let us abolish the GPA; it is a useless and misleading statistic for either teaching or research purposes. If administrators or researchers feel the need for an overall summary of students college learning, we should redesign transcripts so they show patterns in a student's academic career; let us never condone or condemn a student on the basis of a single number, artificially significant to two decimal places.

#### Improved Grading System

Characteristics of improved grading systems have been listed by Deutsch<sup>12</sup> and by Kirschenbaum, Simon and Napier<sup>13</sup>. A few essential features of Deutsch's grading system are (1) an ideal system would foster the view among students that they have a positive interest in the educational attainments of one another, (2) instead of emphasizing individual comparisons, such a system would provide individualized, particularistic feedback, (3) when prerequisites of specific skills and knowledge were necessary for students to engage in a course of study, criterion-referenced rather than norm-referenced tests would be developed and employed to assess the specific skills and knowledge, and similarly, criterion-referenced tests rather than norm-referenced tests would be used when it was necessary to certify the level of a student's educational attainment in a given area. The grading system of Kirschenbaum et al. would (1) eliminate the anxiety which usually goes with grading, (2) create a relaxed learning atmosphere in the class, (3) decrease competition for grades among students, (4) be meaningful (that is, a student's grade should mean something to him, personally), and (5) respect quality as well as quantity of work.

#### Conclusions

Forces for and against grading have been summarized in terms of equations, which seems appropriate and useful for an engineering audience.

The equations are:

History + Research + Experience = Arguments  
Against Traditional Grades

while

Teacher Ease + Administrative Convenience +  
Admission Procedures = Forces Which Maintain  
Traditional Grades

Since there are no clear cut paths through the grading maze, the intention of this paper has been to provide some heuristics for faculty and to pose questions to stimulate discussion. The current situation, especially for norm-referenced grading, is summarized in Dressel's definition of a grade: "An inadequate report of an inaccurate judgment by a biased and variable judge of the extent to which a student has attained an undefined level of mastery of an unknown proportion of an indefinite material<sup>14</sup>.

#### Acknowledgement

My thanks to Randal Barnes, Larry Bereuter, David Hurd, Wallace Whiting, and Malcolm Hepworth for reviewing this paper.

#### References

1. Deustch, M. (1985). Distributive justice. New Haven: Yale U Press.
2. Rescher, N.P. (1966). Distributive justice. New York: Bobbs-Merrill.
3. Deutsch, M. (1979); Education and distributive justice: Some reflections on grading systems. American Psychologist, 34, 391-401.
4. McKeachie, W. (1983). A dialogue on college teaching. Contemporary Education Review, 2, 165172.
5. Astin, A. (1985). Achieving educational excellence. San Francisco: Jossey Bass.
6. Astin, A. (1985). Involvement: The cornerstone of excellence. Change, July /August, 35-39.
7. Deci, E.L. (1985). The well-tempered classroom. Psychology Today, March, 52-53.
8. Stice, J. (1979). Grades and test scores: Do they predict adult achievement? Engineering Education, 390-393.
9. Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. American Psychologist, ~ 193-202.
10. Deutsch (see reference number 1).
11. Milton, O., Pollio, H.R., & Eison, J.A. (1986). Making Sense of college grades. San Francisco: Jossey-Bass.
12. Deutsch (see reference number 3).
13. Kirschenbaum, H., Simon, S., & Napier, R. (1971). Wad-ja-get? The grading game in American education. New York: Hart.
14. Dressel, P. (1983). Grades: One more tilt at the windmill. In A.W. Chickering, (Ed.), Bulletin, Center for the Study of Higher Education, Memphis: Memphis State University, Dec.



KARL A. SMITH

Karl A. Smith is an Associate Professor of Metallurgical Engineering at the University of Minnesota. He teaches courses in mineral and waste processing, extractive metallurgy, applications of systems (linear programming, network analysis, decision engineering and expert systems), and environmental issues.

Karl has served the ASEE as Director and Vice Chairman for programs of the ERM Division and Program Co-Chair of the 15th Annual FIE. He is a frequent writer and speaker on the use of active learning strategies including cooperative group learning and structured controversy discussion and on the use of personal computers as instructional tools in engineering education.