DOCUMENT RESUME

ED 114 025 HE 006 959

AUTHOR	Mayville, William V.
TITLE	The Trouble With Grading Is... ERIC Higher Education Research Currents.
INSTITUTION	George Washington Univ., Washington, D.C. ERIC Clearinghouse on Higher Education.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Oct 75
NOTE	5p.
AVAILABLE FROM	Publications Department, American Association for Higher Education, One Dupont Circle, Suite 780, Washington, D.C. 20036 ($0.40; orders under $15.00 must be prepaid)
EDRS PRICE	MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS	Academic Standards; Achievement Rating; Curriculum Development; *Evaluation Criteria; *Grading; *Higher Education; Learning Motivation; Pass Fail Grading; *Predictive Measurement; Student Development; *Student Evaluation; Undergraduate Students

ABSTRACT

That present grading practices are inadequate to their intended tasks, grading purposes as well as practices require examinations, and the social and educational effects of grading have been too long neglected are the basic axioms of educational reporters and academic professionals. This paper focuses on the problem of grade inflation, its causes and results. Included is a brief historical sketch of the evolution of grading systems in this country along with a discussion of current grading options, ways of measuring student achievement and the uses to which grades are put by the university and society. Whether the reinstatement of conventional (A-F) grading and the insistence on rigorous enforcement of standards by institution are possible or desirable are questioned. It is suggested that the growing belief that learning is independent of evaluation has profound consequences and could be the idea around which future grading systems will revolve. (Author)

ERIC Higher Education
Research Currents

American Association for Higher Education

October 1975

"THE TROUBLE WITH GRADING IS . . ."

William V. Mayville

In a 1971 study of college grading practices, Jonathan Warren warned the reader of certain "biases" on his part: namely, present grading practices are inadequate to their intended tasks; grading purposes as well as practices require examination; and the social and educational effects of grading should not be neglected any longer (Warren 1971, p.1). In 1975, these "biases" are the basic axioms of educational reporters and academic professionals, who readily acknowledge that grading issues have not been systematically dealt with.

By way of a sampling, one author in a recent opinion piece asseverated that "unless the corruption of the present grading system is changed, unless some kind of meaningful system of symbols for differentiating levels of ability can be restored, the colleges will simply forfeit this important function" (White 1975, p. 24). Another critic of current grading practices comments with equal earnestness that

> Dismissal rates [for academic insufficiencies] have gone down at the same time that initial quality of students has gone down. In this manner, the university or college passes on its product to the graduate or professional school—'Let the graduate school do the sorting and sifting.' It is time to ask how long this fraud and illusion shall be continued (Moulds 1974, p. 502).

Such commentaries testify to the sense of longing for a solution to problems that seem inherent to the process of evaluation. This study will consider how grading systems evolved in this country and how this evolution relates to the changing meaning of grades to the student, the teacher, professional and graduate schools, and society.

GRADE INFLATION

The most recent expression of concern over proper student evaluation is embodied in the phrase "grade inflation," meaning students nationwide are being awarded higher grades than ever before. This claim has gained public notoriety by being reported in widely circulated news magazines in a manner which gives to "grade inflation" an aura of public scandal. Some educational analysts believe this is an irresponsible and oversimplified version of the true state of things.

On the surface there does appear to be cause for alarm about current methods of student evaluation. The July 1, 1974, issue of *Newsweek* indicated that in 1961 one-half of the seniors at Harvard graduated with honors, while in July 1974 more than 82 percent graduated cum laude or better. Also the dean's list at the University of Virginia included 53 percent of the student body in 1973 compared with 21 percent in 1965. And the average college grade is now B. An article in the November 11, 1974, issue of *Time* magazine indicates that Stanford's undergraduate grade-point averages have climbed to 3.5+. *Time* also reported that at American University 75 percent of all grades in spring of 1974 were A's or B's, which led an undergraduate dean to ask for a faculty inquiry (p. 66).

Grade inflation seems to carry with it the accumulated frustration, both articulated and surpressed, of students and their teachers over what constitutes a valid statement of student achievement in any particular course. Some commentators have suggested it reflects a sense of cynicism on the part of many professors about the meaning of the grade they assign as well as denoting a fundamental shift in attitude among the professoriate and in society about what grades do or should measure. In the view of one writer faculty members have lost confidence in the value of what they are doing and are therefore unwilling to make rigorous judgments of whether students have mastered their subjects (Scully 1975, p. 1).

According to *Newsweek* (July 1, 1974, p. 50) most educators locate the roots of grade inflation in the disaffection with traditional marking that hit academe during the "tumultuous" 60's. The reason given is that professors became more lenient in their grading during the Vietnam era to keep undergraduates from being drafted and sent to war in Southeast Asia. Another reason put forward is the adoption of pass-fail grading systems, whereby students could choose traditional grades, neutral evaluations, or both. Also, many schools dropped the use of marks like "D" and "F" that signified substandard performance.

The perplexity over the nature of undergraduate evaluation is compounded because graduate schools place great reliance on the undergraduate grade-point average as a valid predictor of graduate or professional school success. Now graduate school admission personnel are confronted by a plethora of applicants with impressive cumulative grade-point averages. Because undergraduate grades are so high, graduate schools are "dismissing transcripts as plainly misleading and are concentrating instead on test scores." One analyst comments that turning the sorting and selecting function over to national testing agencies would afford some clear advantages, since "the clarity and reliability of E.T.S. scores contrast sharply with the confusion of course grades . . ." (White 1975, p. 24). However, this writer goes

Research Currents is prepared by the ERIC Clearinghouse on Higher Education, The George Washington University, Washington, D.C. The material in this publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to the American Association for Higher Education for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either AAHE or the National Institute of Education.

Copies of Research Currents may be ordered for 40¢ each from the Publications Department, American Association for Higher Education, One Dupont Circle, Suite 780, Washington, D.C. 20036. Payment must accompany all orders under $15.

William V. Mayville is a research associate at the ERIC Clearinghouse on Higher Education.

on to say that "college faculties remain the best means for sorting and selecting that we have. With all their flaws, the colleges still offer the best and most flexible route of advancement for merit" (White 1975, p. 24).

HISTORICAL VARIATION OF GRADING FORMATS

To gain a perspective on the issue of grade inflation it is necessary to review past approaches to student evaluation. The first grading systems in America were descriptive, with comments on the character of the student placed alongside comments on the student's intellectual worthiness. After 1800 a variety of grading scales came into existence. Examples of these scales are 1-4 at Yale in 1813, 1-20 at Harvard in 1830, and a scale of 100 at Harvard in 1877.

A radical departure from this approach took place at the University of Michigan in 1851. There a student either did or did not pass course work and in 1852 a plus mark was the only sign used to indicate a student had passed. If a student failed there were degrees of failure (conditional, incomplete). Also, there was strong sentiment that the emphasis on superior attainment and merit of one student over another was neither desirable nor democratic (Smallwood 1935, p. 83).

The first letter grade notation appeared in one reference to a student having a mark of "B," which occurred at Harvard in 1883. By 1895 Harvard had adopted a scale of merit that employed the terms "Failed," "Passed," or "Passed with Distinction." These expressions were qualified by a plus or minus sign (Smallwood 1935, p. 43). However, after 1900 there was much dissatisfaction with the plenitude of methods used to evaluate students. According to Smallwood, educators in this country, in an effort to determine normative judgments of student ability, "indulged in an orgy of new attempts to evaluate mental ability . . . which took the form of elaborate testing programs followed by still more complicated systems for recording the results of these tests" (Smallwood 1935, preface). This telescoping of some of the varieties of grading systems and the resultant move in the direction of external examinations would suggest the contemporary search for ways to accurately reflect student achievement has antecedents deep in American educational history.

GRADING OPTIONS AND THE USES OF GRADES

Many writers believe that changes in grading systems are related directly to changes in the curriculum. In Smallwood's view the reason for the great variety of grading scales in the past was not only due to "dissatisfaction with their efficiency," but also was "due to change in the curriculum; chief among these being the gradual growth of the elective system" (Smallwood 1935, p. 47). With a profusion of new courses being offered, ideas of acceptable achievement might vary widely from course to course, which would suggest disparity in grading methods.

Even though the elective system is now firmly in place in the traditional liberal arts curriculum, this stability has not engendered uniformity of grading. Pass/fail grading, when in vogue in the recent past, was intended to permit students to minimize risk when taking electives to encourage exploration outside their major field. However, because of the need for a high grade, according to Scully, students' exploration outside known areas of competency is sharply curtailed, and "when that happens, the broadening or liberating effects of liberal arts education are seriously jeopardized" (Scully 1975, p. 1). In a recent article on the pass/fail grading option, Philip Myerson, dean of New York University's College of Arts and Sciences is quoted as saying students use the system primarily to protect their averages; on campuses where pass/fail grades are optional (usually for one subject a semester), students often reserve them for difficult courses. And there is skepticism about the positive effect of pass/fail on motivation and learning. In the same article, a Stanford philosophy professor notes that "there is a great temptation for a bright student in a pass/fail course not to do any work and get away with it" (*Time*, Feb. 4, 1974, p. 66).

In a study done at Virginia Polytechnic Institute and State University it was determined that students who elected the pass/fail option performed better in quarterly grade average than other students. The total percentage of courses elected for pass/fail grading at V.P.I. is small at 3.3 percent. This same study cites extensive surveys at Princeton and the University of Southern Illinois that indicate students show some falloff in motivation in pass/fail courses and possibly learn somewhat less. Although students who have the pass/fail option do take a few additional courses they might not have taken otherwise, the willingness to explore new areas is not guaranteed by pass/fail grading (Delohery and McLaughlin 1971).

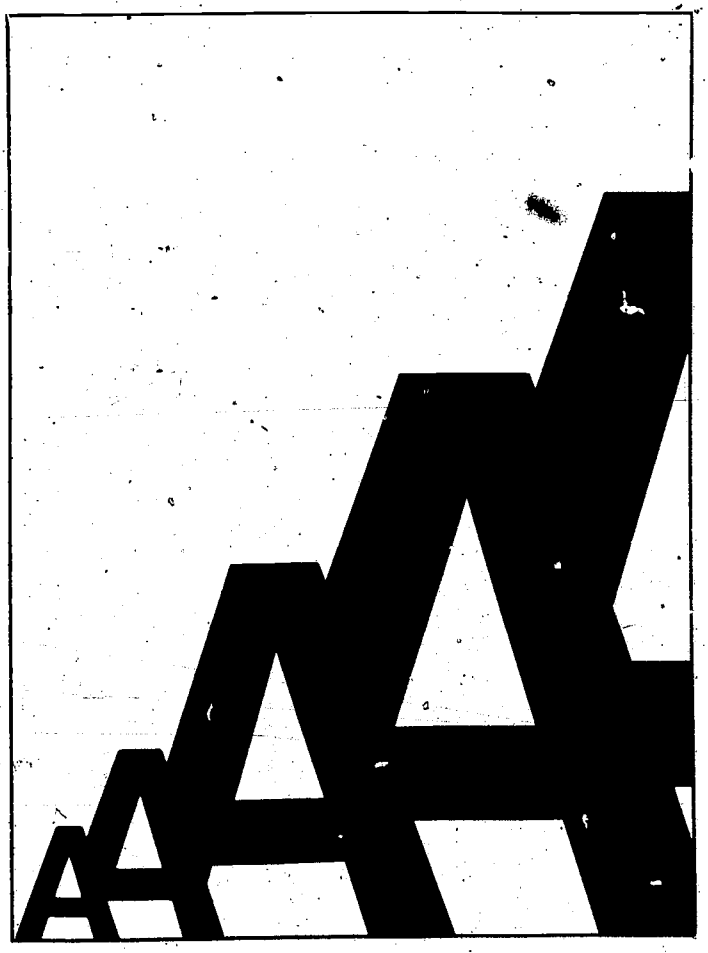
In a later study at the University of Wisconsin-Stevens Point on pass/no pass, the investigators concluded that this option produces lower quality scholarship, affects the grade-point average integrity, and causes students to take more courses per semester. Also, this option lowered rather than raised student motivation (*An Evaluation* 1973).

At the University of Alberta in Edmonton, a study was undertaken on their pass/fail system with much the same negative correlations of the other studies cited here. However, the author concluded that despite the evidence he would not recommend abandoning nontraditional grading because (1) it is too early in the innovational development to measure the real impact on the student, and (2) the nontraditional grading may do more to foster a lifetime of learning attitude (Otto 1972).

Two issues spring from the foregoing discussion of the pass/fail option. One is the negative effect of this option on motivation. The other has to do with the intrinsic, long-term value of a grading system on the student's interest in learning beyond the undergraduate or graduate/professional school degree.

On the matter of motivation there seems to be contradictory opinion about whether traditional grades (A to F) do motivate all students to learn. For example, one study indicates that students with competitive and manipulative personality styles significantly prefer interpersonal comparison with their peers. This suggests that students who are more passive may do less well in a highly competitive environment with grades as the reward for achievement (Levin 1973, pp. 67-72). A questionnaire survey at the University of Indiana indicated that students understand what they must do to get a grade and both students and faculty believe grades are motivators (males more than females) (Doerann, George et al. 1974, p. 4-5, 8-9). Yet an earlier study at the University of California at Berkeley suggested that the grading system be changed because, among other reasons, grading failed to stimulate students to learn (Miller 1965).

In relation to intrinsic, long-term value of "traditional" grades, the Berkeley study found that the way in which the grading system encourages a kind of competition is alien to the real purposes of university education and discourages the development of intrinsic and lasting intellectual interests of self-definition and evaluation (Miller 1975). This same theme is sounded by Murphy and Raushenbush, who comment that "in a university environment



where competition for high marks is taken for granted as a means of achieving intellectual distinction, of winning scholarship awards, and gaining social prestige, the full intellectual development of the student is likely to be inhibited" (Murphy and Raushenbush 1960, p. 14). The authors conclude that "a college which looks carefully into the effects of its own requirements on its students and is concerned for the development of an integrity of interest in ideas and intellectual development will seek means other than the conventional grading system" (Murphy and Raushenbush 1960, p. 29).

MEASURING STUDENT ACHIEVEMENT

Smallwood (1935) suggests that grades originated in an attempt to express to those who had been examined the professor's idea of their achievement (p. 107). In her view, during the last 300 years students at institutions of higher learning have "more or less placidly accepted the subjective character of the opinions expressed concerning their intellectual ability" (Smallwood 1935, preface). But during these same years faculty became greatly discontented with the methods used to evaluate their students. This discontent spawned a myriad of grading combinations that exist to the present day.

In the past it was the faculty who were aroused to change grading procedures. Now the students are concerned because their future careers are at stake. This has prompted a shift in emphasis to learner-centered rather than teacher-centered evaluation. Whether this shift will make it any easier for students to gain access to careers of their choice is debatable. In former times it seems to have been taken for granted that potential careers would not be jeopardized by a less than superlative grade-point average. In the 1970's this conclusion is corroborated by analytical surveys of careers of physicians, lawyers, and politicians who have succeeded in their professions despite undistinguished academic records.

Dawes comments that the standard variables considered in selecting students for graduate school do not correlate well with later measures of the success or attainments of selected students (Dawes 1971; Willingham 1974), and mentions one investigator (Morton 1971, 1972) who proposes the abandonment of the Graduate Record Examination. Dawes also develops the thesis that "the variables for admitting students to graduate school *must* have low correlations with future measures of the success of these students," because of the adjustment of the variables (grades versus test scores) by each individual school or department to suit whatever is considered the best formula to satisfy its educational purposes (Dawes 1975, pp. 721-723). The dilemma of who should be admitted still persists. With the tremendous number of students wanting to establish careers in prestigious and well paying fields, professional and graduate schools seemingly must rely on statistical indices of a student's intellectual promise. An important question to ask is what is being measured by these indices.

A 1974 study at Indiana University found that students believed good grades did not accurately reflect their true achievement in a course, while conversely faculty believed the grades they assigned were a true assessment of student achievement (Doerann, George et al., 1974, p. 4-5, 42). This difference in perception is crucial. One common view of student achievement is the response a student makes to the formal criteria established by the teacher at the beginning of a course. Good performance on an examination given by the teacher is enough to satisfy many who do the evaluating in higher education of the suitability of the grade to the level of achievement. Here the emphasis is on goal-setting by the teacher, with the students' intellectual needs being reflected in the professor's idea of what the subject requires. The rationale for this is summarized by one professor as follows: "The matter of how well students have mastered their own goals is of interest, but, since students register for a given course for a variety of reasons, and considering that the instructor, by virtue of his superior knowledge and experience, has the right to set course goals, the suggestion that the instructor orient his evaluation toward a certain student's goals is quite dubious" (Moulds 1974, p. 504).

However, even when the instructor sets the goals the consistency of evaluation procedure is anything but assured. In a study of freshmen English students at Hofstra University, those students who had demonstrated superior writing ability on achievement tests and in high school English courses were exempted from English 1 and 2. At course completion there was no statistical difference in grade distribution of A to F when the "better" students were exempted from the class. Since there was no significant difference in the percentage of A's and B's given before or after the exemption, it is suggested that instructors were using a relative rather than an absolute standard in their grading (Lichtenstein 1971, p. 3).

The growth of new approaches to learning already has created profound changes in the idea of evaluation that may or may not allow the student intellectual freedom. For example, two psychology courses, one an undergraduate course in experimental analysis of behavior, and the other an honors section and a night

class section of introductory psychology, were involved in a program designed to result in the same "terminal performance" for every student. With the exception of two students who received a "B," all other students received an "A" as their final grade. A course evaluation questionnaire indicated that students understood what was expected of them, felt their comprehension of the subject matter was above average, and were motivated to take other undergraduate psychology courses (Chapman 1971).

With this approach to learning, the question of evaluation takes on another dimension. Here learning has to do with instilling a desire to learn by use of a "mastery" approach. The future use of the earned credit seems of less inherent interest than the desire of the student to learn.

CONCLUSIONS

The shift from teacher-centered evaluation to student-centered evaluation has resulted in a multiplicity of grading formats that reflect the tremendous diversity of programs and constituencies in American higher education in the 1970's. A part of this process has been the devolution of various grades into one simple form, which represents mastery as an achievable goal, and results in the tendency to grant every student an "A" or "Pass." This devolution is viewed by many as a corruption of the differentiating or sorting-and-sifting function of educational institutions. In this regard, there is a concern on the part of some that the meaning of grades and degrees has been debased by attempts to respond to the learning patterns of individual students.

The question is whether the reinstatement of conventional (A to F) grades and an insistence on rigorous enforcement of "standards" by institutions nationwide is possible or desirable. The belief that learning is independent of evaluation has profound consequences and could be the idea around which future grading systems will revolve. Just how undergraduate, graduate, and professional school admissions personnel will adjust to this notion is a matter of great concern. Collins and Nickel (1974) comment that "as competency-based education becomes more widespread in higher education, it appears that additional modifications and changes in grading, recording, and averaging practices will come into being and that the traditional transcript/GPA approach will lose more of its historical meaning" (p. 10).

This seems to be happening. However, whether the emerging heterodox notions of how students should be evaluated will gain wide acceptance by those who teach in colleges and universities is uncertain. What is certain is the need for a systematic examination of the grading question from which clear guidelines can be drawn. Until this is accomplished the diverse student populations who are in the process of being educated are in danger of being poorly served in their desire for a higher education.

Some questions to consider are: Should there be a return to the "traditional" system of A-F, with guidelines across disciplines to ensure maintenance of standards? Can disciplines with scholarly traditions that go back centuries accommodate new learning strategies and new evaluation methods? Can any system of evaluation encourage individual students to learn a given subject and, at the same time, accurately measure what students have learned? And, if grades and test scores do not correlate well with adult achievement, should something be done to identify and build into evaluation mechanisms those variations that do predict such achievement?*

*See Munday and Davis (1974) for a discussion of this topic.

BIBLIOGRAPHY

Axelrod, Joseph. "The Creative Student and the Grading System." In *The Creative College Student: An Unmet Challenge*, edited by Paul Heist. San Francisco, Calif.: Jossey-Bass, Inc., Publishers, 1968.

Chapman, Richard F. *Production of Specified Terminal Performances in Every Student in Undergraduate Psychology Courses*. Pullman: Washington State University, 1971. ED 056 979. MF-$0.76; HC-$1.58.

Collins, Janet R. and Nickel, K. N. *A Study of Grading Practices in Institutions of Higher Education*. ED 097 846. MF-$0.76; HC-$1.58.

Dawes, Robyn M. "Graduate Admission Variables and Future Success." *Science* 187 (28 February 1975): 721-23.

DeLohery, Pat and McLaughlin, Gerald. *Pass-Fail Grading*. Blacksburg: Virginia Polytechnic Institute and State University, 27 September 1971. ED 056 651. MF-$0.76; HC-$1.58.

Doerann, George and Doerann, Judith. *Grading: Student and Faculty Opinions*. Indiana Studies in Prediction, No. 26. Bloomington, Ind.: Bureau of Educational Studies and Testing, 1974. ED 097 823. MF-$0.76; HC-$1.95.

An Evaluation of the Pass-Fail Grading Policy at UW-Stevens Point. Stevens Point, Wisc.: Office of Institutional Research, University of Wisconsin, February 1973. ED 076 098. MF-$0.76; HC-$1.95.

Lane, Michael. "Grade Inflation Not Really a Problem?" Letter to the editor, *Chronicle of Higher Education*, June 23, 1975, p. 11.

Levin, William C. "Social Comparison of Grades: The Influence of Mode of Comparison and Machiavellianism." *Journal of Social Psychology* 91 (October 1973): 67-72.

Lichtenstein, Pauline. *Distribution of English 1 (Day) Grades Before and After the Exemption of Students Based on Superior Writing Ability: A Comparison Between Fall 1969 and Fall 1970 Grade Distribution in English 1*. Hempstead, N.Y.: Center for the Study of Higher Education, Hofstra University, August 1971. ED 054 736. MF-$0.76; HC-$1.58.

Miller, Stuart. *Report on Methods of Evaluating Students at the University of California, Berkeley*. Berkeley: University of California, 1965. ED 076 631. MF-$0.76; HC-$1.58.

Moulds, Henry. "To Grade or Not to Grade: A Futile Question." *Intellect* 102 (Summer 1974): 501-504.

Munday, Leo A. and Davis, Jeanne C. *Varieties of Accomplishment After College: Perspectives on the Meaning of Academic Talent*. Iowa City, Iowa: American College Testing Program, 1974.

Murphy, Lois and Raushenbush, Esther. *Achievement in the College Years*. New York: Harper & Brothers, 1960.

The 1973 Boston Conference on Grading Alternatives. Amherst, Mass.: Value Associates, December 1973. Speeches given before the National Conference on Grading Alternatives. ED 087 131. MF-$0.76; HC-$4.43.

Otto, David J. *A Study of the Pass/Fail Grading System*. Edmonton: Office of Institutional Research and Planning, Alberta University, 1972. ED 077 472. MF-$0.76; HC-$1.58.

Scott, Robert A. "Grades: Inflated, Skewed, or Both?" *College Board Review* No. 96 (Summer 1975): 6-9.

Scully, Malcolm. "Inflated Grades Worrying More and More Colleges." *Chronicle of Higher Education*, May 19, 1975, p. 1.

Smallwood, Mary L. "An Historical Study of Examinations and Grading Systems in Early American Universities," *Harvard Studies in Education* 24, Cambridge, Mass.: Harvard University Press, 1935.

Trow, Martin. *Teachers and Students*. New York: McGraw-Hill Book Company, 1975.

Warren, Jonathan. *College Grading Practices: An Overview*. Report No. 9. Washington, D.C.: ERIC Clearinghouse on Higher Education, March 1971.

White, Edward M. "Sometimes An A is Really An F." *Chronicle of Higher Education*, February 3, 1975, p. 24.

To order documents in the bibliography identified by an ED number, write to ERIC Document Reproduction Service (EDRS), Computer Microfilm International Corporation, P.O. Box 190, Arlington, Va. 22210. Documents with HE numbers are presently being processed by EDRS and will be assigned ED numbers upon publication in *Research in Education* (RIE). In ordering, ED numbers must be specified. MF indicates microfiche and HC denotes hard copy; payment must accompany orders of less than $10.00; and all orders must be in writing.