DOCUMENT RESUME

ED 109 209 TM 004 667

AUTHOR Otto, David
TITLE Pass-Fail Grading Systems: A Literature Review.
INSTITUTION Alberta Univ., Edmonton. Office of Institutional Research and Planning.
PUB DATE Sep 73
NOTE 68p.; Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.76 PLUS POSTAGE. HC Not Available from EDRS.
DESCRIPTORS *Academic Standards; Achievement Rating; College Students; Educational History; Evaluation Methods; Grade Point Average; *Grades (Scholastic); *Higher Education; *Literature Reviews; *Pass Fail Grading; Predictive Validity; Reliability; Student Evaluation; Student Motivation; Validity

ABSTRACT

What began as a review of Pass-Fail ended as a comment on the evaluation schema of student academic performance. "Standards," "Evaluation" and "Grades" are defined. A brief history of grading among early North American institutions of higher learning is compared to precursory practices in Europe. The validity and reliability of traditional grades are examined. Finally, the original review of the literature on Pass-Fail is presented. (Author)

OFFICE OF INSTITUTIONAL RESEARCH AND PLANNING

THE UNIVERSITY OF ALBERTA EDMONTON, ALBERTA T6G 2J9

PASS-FAIL GRADING SYSTEMS:

a Literature Review

Submitted to

THE G. F. C. COMMITTEE TO

INVESTIGATE TEACHING

by

David Otto, Ph. D.

Office of Institutional
Research and Planning

September, 1973

TABLE OF CONTENTS

LIST OF TABLES

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

PREFACE

This report began as a modest literature review of the advantages and disadvantages of the Pass -Fail Grading system so popular among today's college students. For one to understand Pass-Fail, however, he must also answer questions about the nature and purpose of grades. Hence, a new stratum of investigation appeared. But it seems that this new stratum contained questions which led to other strata, almost ad infinitum. How did grades come into use? What were they used for? What are they used for today? What is the relationship, if any, between grades received and material learned; between scholarship and performance? Did the grade received deal only with the intellectual ability of the student, or were other aspects (i.e., interpersonal relationships and human manipulation) involved? Were grades coveted as ends unto themselves, or employed for higher order ends?

Time and organizational considerations compelled me to draw lines and establish priorities. I have had to stop short on many ancillary questions simply because the literature is plentiful and time was scarce. Besides, the dominant feature of this report warranted all possible attention.

This report is no longer solely about Pass-Fail Grading. A more important question is that of the proper use of any set of symbols, called grades, in the avowed purposes of higher

learning.

I am indebted to many others for their assistance. Firstly, the members of the G. F. C. Committee to Investigate Teaching, for their support and counsel. Secondly, to the members of the Office of Institutional Research and Planning, for all their personal and technical aid. Finally, to the Librarians of the Education and Cameron Libraries, especially those in Interlibrary Loan, for their patience and composure in the presence of a madman.

March, 1973

Revised August, 1973

David Otto

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

I. DEFINITIONS AND DISTINCTIONS

A great many theses on grading practices, both orally and in the literature, lack cogency because authors fail to note the difference between terms used in evaluating various grading systems. I shall endeavor to adhere to the following definitions when speaking of grading systems in colleges and universities throughout the North American continent

STANDARDS

The term "Standards" refers to the calibre of the undergraduate and graduate education as it exists within a specific institution. The calibre of a student's education in college can be determined in a number of ways. It could be derived from the entrance requirements at both the undergraduate and graduate levels; it could be measured by the academic attrition rate, or conversely, the numbers of students winning national awards and scholarships; it could be measured by the type of programs (curricula) offered to students; it could be assessed by the requisite qualifications for graduation; etc. A host of separate indices are available (see Astin, 1968).

The single evaluator of the calibre of undergraduate education to be examined in this report is the system of grading students' performance while they are still undergraduates.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

GRADES

Basically I define a grade as a symbol (most commonly a percentum, number or letter) chosen from among a set of symbols, representing the evaluation of a learner's academic performance, as determined by his teacher, which at the end of a formalized segment of learning in an institution of higher learning is transcribed unto the learner's record for future reference. A Grade Point Average (GPA) is an average of these symbols within a particular institution of higher learning.

EVALUATIONS

The term GRADES must be distinguished from EVALUATIONS. A mark of excellence (a 9 at the University of Alberta; A's in some other institutions, 100% in still others) is the result of an evaluation, not the evaluation itself. Indeed an evaluation, as Elbow (1969) put it, occurs every time student and teacher meet. Few of these evaluations are summarized and formally recorded. Those that are recorded eventually find their way, as a composite, into the record of the student's vitae. Quite clearly the formal process of evaluation is comparing the student's progress to either the standards of the institution or the instructor's own standards to see where he stands, then assigning him his grade.

Now these three concepts (standards, evaluations, and

grades) may not be employed in the same manner from university to university, or even within a university. On the one hand, a college may set very, very high admission standards, but once the student is admitted will probably not grade students' work in any minute degree. On the other hand, undergraduate admissions standards may be intentionally set low, but a high attrition rate results when the faculty enforces a very rigid standard of evaluation during the student's first and second years. Or, as in the case of some Community Colleges, open admissions are complemented by a low standard of evaluation. In any case, the quality of the institution may or may not be affected.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

SECTION II: BRIEF HISTORY OF EVALUATIONS

The process of the tutor evaluating his student has probably existed for quite some time, but we have no evidence that either the results of the evaluation (i.e., marks and grades) or certification of learning (diplomas and degrees) existed in the days of the Greek philosophers. One presumes then that while evaluation took place the results of evaluations were either communicated orally or non-verbally.

Certification, as we have come to know it today, probably began with the creation of colleges and universities in medieval times. Students at the University of Paris, for instance, presented themselves to the Chancellor to be examined by him before receiving the licentia legendi (Hasking, 1965, p.46 et seq. describes both the University of Paris' and the University of Bologna's graduation exercises).

One can presume that this style of final examination for a degree continued even unto the discovery of the new world. Mary Smallwood (1935), in her study of five American colonial universities, tells us that public graduation examinations were common at Harvard and Yale in the mid-17th century. At Harvard each graduand was expected to be examined on the subjects of Latin, Greek, Hebrew, rhetoric, logic, physics and law--usually orally, and by members of the audience or external examiners. Graduation was not always the foregone conclusion it is in

today's convocation. The shift from an oral to a written examination was a late 18th and early 19th century event. Yale University instituted a 'biennial' examination in 1830.[1] As these five schools approached the close of the 19th century, not enough time was available to orally examine the candidates for graduation in depth. A number of forces led to the eventual decline of the practice of a single public examination for a degree. Firstly was the sheer numbers of students enrolled (paltry compared to the numbers in today's colleges and universities). As Ticknor wrote in 1825:

> "For it is idle to think of hurrying, in a single day, through the examination of sixty young men in the studies of a year; and it is unreasonable to expect the gentlemen [external examiners] will come from a distance and undergo considerable expense for many days, in order to do it more thoroughly. A good examination of this sort is one of the most laborious and one of the most valuable services that can be performed for the advancement of knowledge; and, if it is to be well done, it should not only be thorough, patient, and exact; but, should be public at the time, and public in its results."[2]

A second major force in the demise of the single final examination can be attributed to the proliferation of knowledge. Not only was it more difficult for students to have a universal grasp of knowledge but, perforce, the number of examiners had to

---

(1) It is interesting to learn that approximately 65 years earlier (1762), Yale students refused to be subjected to any examination "...except it be for a degree,...." (1935, p. 35).
(2) Smallwood, M., op. cit., p. 12

be increased to accomodate the increased specialization.

Thus, in the last 150 years, a shift has occur[illegible] from public examinations for a degree held by external examiners to written tests at the termination of a course, held in camera, and judged by the institution's own faculty.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

SECTION III. THE FUNCTIONS OF GRADES

Stanford Ericksen (1967) lists three possible functions of grades: 1) to satisfy social conditions, primarily employment opportunities, 2) to assist internal administrative procedure, and 3) to provide an evaluative feedback to the student.

SOCIAL CONDITIONS

Businessmen utilize grades as one of their criteria when selecting candidates. A recent survey by Bailey (1972) of governmental and commercial employers in Chicago and environs shows that undergraduate and graduate grades are second in importance only to "Previous work experience." Bailey concludes that:

> "The response of the survey group appears to be consistent with the generally held business view that an individual's previous track record is the best indicator of his future performance." p. 18

The second major group of users of the students' transcripts are the graduate and professional schools. Admissions committees closely review the candidates' transcripts as a routine pre-admissions procedure. The topics of the predictability of grades and of the graduate schools' admissions officials' views of non-traditional grades will be dealt with in later sections.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

INTERNAL PROCEDURE

Usually a single symbol, representing an extended period of learning on a specified topic, a grade encapsules all that is alleged to have transpired during that period. This single symbol indeed facilitates record keeping. Assume, for example, that each full-time day undergraduate student at the U. of A. elects five courses. One can easily see that the Registrar's office has 75,000 numbers to transcribe onto 15,000 individual records (source: Registrar's Statistics, 1971-1972). One stroke per evaluation is infinitely less time consuming than the briefest of essays.

Grades also serve other uses in the undergraduate student's career. The extremes of the grade point distribution are used by university officials to either bestow awards, as in the case of a high Grade Point Average, or take disciplinary action, as in the case of a low Grade Point Average.

Finally, grades are employed in periodic decisions about the status of the student. Firstly, they determine where a student ranks vis-a-vis his peers. Secondly, they indicate if the student should be promoted from one year to the next. Thirdly, they are used to ascertain the student's candidacy for graduation. Fourthly, they are a basis for the decision to permit a student to enrol in more advanced courses of the same discipline.

Fifthly, they may be used for certain research studies (e.g., the McLean report of 1971). Finally, they facilitate transfer procedures both inter- and intra-institutionally. Grades are very handy devices to the administrator.

FEEDBACK

But let us not overlook the student. How do the students benefit from grades? Much literature alleges that grades are useful feedback devices. I might add parenthetically that it's deplorable that this may be in some cases the only feedback device available to the student. As a feedback, however, Feldmesser (1972) feels that the student is in a position to assess his progress within the learning situation in two contexts: firstly, he can compare himself to the performance of his classmates, and secondly, he can realize his preferences, abilities and weaknesses, thereby identifying his career objective and areas of competence.

Haagen (1964) feels that grades guide the student as well as motivate him.

> "Without some form of appraisal that provides orientation and confirmation of achievement, attempts at learning are inefficient and the student becomes demoralized" p. 89

Warren (1971) points out, however, that one should maintain the distinction between grades and evaluation; between the end result and the process.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

> "The contention that feedback to students about their performance constitutes an important purpose of grades...confuses evaluation (the assessment of performance) with grading (reporting of the assessment results). A variety of procedures are available to inform students about the nature of their performance without the publication of a summarizing symbol to represent overall performance in a course." p. 15

SUMMARY

Grades have functioned as feeback to the student, and as such can be perceived as immediate and short lived stimulii. A student can see, during the term, how well (or poorly) he is progressing. Once the term ends, the final grade becomes the artifact of what was purported to have occurred between learner and tutor. Its stimulus effect upon the student is removed (except in the case of supplemental examinations), but the artifact remains, and affects the student's future. In the near future certain advantages of the system of higher education (more advanced courses, intra- and interinstitutional transfer, etc.) may be offered to or withdrawn from him -- depending on his academic record. In the more distant future, these artifacts affect the student's chances for graduate or professional education, or employment. To the student, grades influence not only the here and now, but also his chances at the career of his choice. Is it any wonder he is tempted to favor these symbols of high achievement?

SECTION IV. A LOOK AT TRADITIONAL GRADES

It is a common practice for the admissions committees of graduate and professional schools to examine candidates' transcripts. What do they look for? Hofeller (1972), in her survey of 668 deans of the U. S. Graduate and Professional Programs in liberal arts, sciences, education, law, medicine and nursing found among the 391 responses that:

> "Grade-point averages remain the single most important criteria for the evaluation of the graduate school applicants." p. 4

This one factor was followed very closely by two others: entrance examination scores and the prestige of the undergraduate institution.

Entrance examinations, such as the Scholastic Aptitude Test (SAT), the Graduate Record Examination (GRE) and the Miller Analogies Test (MAT), as commercial examinations, have undergone rigorous trials of validity and reliability. Can the same be said about the Grade Point Average (GPA)? It, too, purports to measure the concrete acquisition of knowledge, the ability to reason, to cipher, to expound, etc. It, too, is a result of an examination--or to be more precise, the mean of the weighted scores on a number of examinations given by a number of examiners. One wonders if the Grade Point Average has undergone similar tests of validity and reliability.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

RELIABILITY

(a) Test-Retest Reliability

We begin our examination of the GPA's reliability by assuming that, in the macro-educational sense, the nature of a high school education is equivalent to the nature of an undergraduate education; and that the nature of a college education is equivalent in nature to a graduate education. We now can assume that the whole of the student's high school education was evaluated and represented by his GPA, and the whole of his college experience is a retest of his high school education, which is in turn represented by his college GPA. If this is the case, then ample evidence exists to show that high school grades and Grade Point Averages produce significant correlations with first year college grades and Grade Point Averages. Similarly, undergraduate college grades and GPA's are reliable predictors of first year graduate school achievement (see Hills, 1964; Hills, et al., 1965; Klugh & Bierley, 1959; Lannholm, et al., 1968; Lewis, 1966; and Chamberlin, et al., 1942). It is interesting to observe that the efficacy of the correlations deteriorates as temporal distance between the two variables increases. Thus the high school GPA has a lower correlation coefficient to graduate school GPA'S than it does to first year undergraduate GPA's. Rogers (1937), has shown that each of three incoming classes of students had a high correlation coefficient between each pair of academic terms, but as the length of time between pairs of terms

increased, the correlation coefficients decreased.

Looking deeper into the Grade Point Average, one easily observes that there are two fundamental inputs: the student as the performer and the teacher as the evaluator. In the correlational studies so far, the student, as academic performer has been a fixed element. Clark (1964) studied the the second element when he reviewed the grades earned by students in eighteen separate freshman classes who entered a single university between the years 1931 and 1958. The composition of the group of teachers, while it undoubtedly did alter during those 27 years, can be assumed to have remained somewhat stable. Each freshman class was, on the other hand, entirely new. Clark concluded that

> "...the ratio of the weighted standard deviation of student averages to the standard deviation of all grades, using the credit hour as a unit, can be taken as an indication of the reliability of the individual Grade-Point Averages." p. 430

Thus, in one university at least, a certain consistency in grading practices by the faculty was present.

(b) Intramural Comparisons of Reliability

So far we have looked at the association between separate institutions: high school to college, college to graduate school. Clark has introduced a new dimension into our examination of college grades: intramural comparisons of reliability.

The equivalent forms test of reliability could be applied intra-institutionally, if one is to assume that the quality and the content of different courses offered in different departments is roughly equivalent. Thus, an introductory course in physics is thought to be equivalent to an introductory course in chemistry, and both are seen as on a par with a freshman English course -- at least this is the conclusion one draws from examining the weighting system of "credits" promulgated in most college catalogues (see Heffernan, 1973). What evidence exists of an interdepartmental reliability? Do the instructors in various subunits of an institution of higher learning assign the same grades to the student body within a given term?

"The proportions of "A's" given", says Chamberlin (1942), "...vary not only from college to college, department to department, and instructor to instructor, but also with the same instructor grading the same materials at different times." (p. 23)

A dozen years prior to Chamberlin's massive study, Crawford (1930) examined the grades given to six freshman classes by faculty members in twelve departments at Yale University, and observed that

> "...it is bad enough to have to tie together so many credits as if they were pieces of string which eventually must reach a certain definite length. It is worse to find some of these pieces are rubber bands of variable lengths and elasticity. Granted that "crediititis" is in

> itself a malignant disease, so long as it persists, its ravages may at least be lessened to some extent by establishing more uniform equivalents between the necessary credits and true performance." p. 234

Crawford went on to observe that

> "...those departments whose marks most vary from the class mean also exhibit the most noticeable fluctuations from year to year, as compared even with their own subject norms. These departments also tend to show the greatest divergence between grades assigned by the experienced and the inexperienced teachers." p. 240

Crawford's departments remained anonymous throughout his article. So, too, were the departments studied by Gamson (1969). Dr. Gamson did, however, assemble her departments under two disciplinary rubrics: social and natural sciences. The grading practices of the social science faculty at 'Hawthorn' college were more liberal than those of the natural science faculty in the same college (see Table 1 below).

I decided to replicate Gamson's findings with data at hand. I selected the grade distributions from each of five of the social and natural sciences departments at a large public university in the western part of the United States. I also selected the grade distributions given in ten introductory level (200 level) courses taught at the University of Alberta, five by social science departments and five introductory courses by four natural science departments (data provided by the McLean (1971) and Coull (1973) reports). The results are presented in Table 1.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

TABLE 1

Distributions of the Grades Assigned Students by Selected Social and Natural Science Departments at Three Institutions of Higher Learning: Gamson's 'Hawthorn' College, a large public University in the western United States, and the University of Alberta, Canada.

	GRADE						
Hawthorn College [1]	A	B	C	D	E	Sum	Number
Social Sciences	15	43	29	10	3	100%	251
Natural Sciences	13	25	37	14	10	100	214

$x^2 = 10.31$, df = 4, $P < 0.05$

Western U. S. University[2]							
Social Sciences	24	42	27	5	2	100%	2,691
Natural Sciences	21	28	38	8	5	100	4,341

$x^2 = 6.84$, df = 4, $P < 0.20$

University of Alberta[2]	9[1]	8	7	6	5	4	<4	Sum	Number
Social Sciences	4	11	21	22	22	12	8	100%	5,291
Natural Sciences	3	8	16	24	23	14	12	100	3,518

$x^2 = 2.36$, df = 6, $P < 0.90$

While the frequency distributions of the two scholarly divisions in Gamson's study were statistically significant, similar distributions in two large universities, one in the U. S., the other in Canada, were not. Based on these limited data

---

[1] Both U. S. institutions use a five-letter grading system while the University of Alberta is on a 9-point system (9 = excellent, < 4 = Fail).

[2] Departments in the U. S. Public University were: Anthropology, Economics, Political Science, Psychology, Sociology, Astronomy, Biology, Chemistry, Geography and Physics. For the University of Alberta the subjects were: Anthropology 202, Economics 200, Political Science 200, Psychology 200, Sociology 200, Biology 230, Chemistry 200 and 202, Geography 201, and Physics 200.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

While the frequency distributions of the two scholarly divisions in Gamson's study were statistically significant, similar distributions in two large universities, one in the U. S., the other in Canada, were not. Based on these limited data one could infer that membership in a particular field (social or natural sciences) does not, de facto, result in strict or lenient grading practices.

The difference in the results of 'Hawthorn' College and the two universities might be partially attributable to another variable: time. Crawford observed that those departments at Yale with the widest deviation from the mean also had the "...most noticeable fluctuations from year to year." Gamson's data were obtained in the fall of 1962, a few years after 'Hawthorn' was founded. The grade distributions from the two western universities were gathered some nine years later, in the academic year 1970-71, and fully half a century after the creation of these two institutions. Surely one aspect of reliability is that the device measured stands immutable to the forces of time. The Prudential Life Insurance Company selected the Rock of Gibraltar as its trademark because it connotes an everlasting reliability. Are grading practices similarly unchangeable? Crawford observed fluctuations. Table 1 hints at the possibility of change over time. Rasmussen (1969) provides more conclusive evidence. Rasmussen found the faculties in the schools of education at six midwestern universities during the 1950's differed significantly

in their grading practices from the practices of their colleagues in schools of business administration and liberal arts. He also discovered that the differences present in the 1950's had largely been erased in the 1960's Rasmussen concluded that the grading and standards of the Education faculty members rose to those of the faculties in Business Administration and Arts.

Again, using information provided in the McLean and Coull reports, I computed the grading practices of the Education, Business and Commerce, and Arts faculties at the University of Alberta.

COMPUTATIONAL PROCEDURE: McLean presents the numbers of 9-point grades given in 200,300,400,500 & 600 level courses in each of these faculties. These data appear in Table VIII of his report. I chose the 200 and 300 level courses for each of these faculties, primarily because many Arts students are still in a three-year general BA program, and to compare 400 level courses in Arts with the other faculties would be pitting honors students' capabilities against regular undergraduate students in Business & Commerce, and Education. The numerical distribution of each grade at each level was converted to a percentage for each of the three years in the report. The two levels were then combined. These percentages became the basis for chi-square applications, using the Department of Educational Research Services program NONP02. I realize a two-way analysis of variance would have been a better statistical technique. The data, however, were not in a form amenable to the department's analysis of variance programs.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

TABLE 2

Grade Distributions of the Teaching Staff in three University of Alberta Faculties, expressed as a Percentage of all the Marks given by each respective Faculty to Students Enrolled in 200 and 300 Level Courses, 1968-69 to 1971-72

PART I. INTRAFACULTY COMPARISON

FACULTY OF ARTS

	1	2	3	4	5	6	7	8	9	Number of Marks
67-68	0.9	2.3	3.6	11.9	21.5	27.7	20.0	9.5	2.6	21,822
68-69	0.7	1.7	3.2	10.8	19.6	27.0	22.1	11.4	3.5	25,891
69-70	0.6	1.7	3.5	10.2	19.3	26.8	22.4	11.9	3.6	30,311
70-71	0.5	1.3	2.8	9.2	18.3	27.0	24.2	13.0	3.7	29,551
71-72	0.3	1.3	2.5	9.0	18.6	28.3	24.0	12.4	3.6	29,339

$x^2 = 3.35$, df = 32, P = 1.00

BUSINESS & COMMERCE

67-68	0.9	3.5	6.4	17.3	26.2	24.4	13.7	5.4	2.2	2,793
68-69	1.3	3.7	5.1	15.9	20.6	24.9	17.0	7.9	3.6	3,747
69-70	3.7	4.7	7.2	14.5	22.6	23.7	14.2	6.6	2.8	5,252
70-71	2.0	4.1	4.4	13.0	20.1	22.5	19.0	9.9	5.0	5,985
71-72	0.9	4.3	3.1	14.9	21.6	24.1	16.9	10.4	3.8	6,419

$X^2 = 11.95$ DF = 32 P = 1.00

EDUCATION

67-68	0.1	0.9	1.1	7.4	17.8	32.3	25.2	11.5	3.7	8,878
68-69	0.1	0.5	0.8	4.7	12.8	30.1	32.1	15.0	3.9	10,133
69-70	0.3	0.7	1.0	5.7	14.2	29.9	29.3	14.1	4.8	9,332
70-71	0.3	0.6	0.8	5.7	14.9	29.4	28.0	15.3	5.0	9,054
71-72	0.3	0.5	1.1	6.7	13.8	28.4	28.6	16.4	6.2	8,773

$X^2 = 5.36$ DF = 32 P = 1.00

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

PART II. INTERFACULTY COMPARISON[1]

YEAR	x^2	P <
67-68	18.5	.30
68-69	22.4	.20
69-70	27.7	.05
70-71	15.3	.50
71-72	18.6	.30

[1] The grade distributions of each of these three faculties have been assembled by year. For example, the 1967-68 distribution of the Faculty of Arts was joined by the 1967-68 distributions of the Faculties of Business & Commerce and Education to make a 3 X 9 matrix. There were sixteen degrees of freedom

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

The teaching staff in the faculties of Arts, Business & Commerce, and Education remained consistent in their various grading practices over the five years reviewed (see Part I); but with each successive year the differences between the grade distributions given by each faculty varied, from a position of loose agreement in 1967-68 to one of statistically significant difference in 1969-70, and seemingly following a curvilinear trend back to homogeneity in 1970-71.

(c) Summary of Reliability

Standard reliability tests of college grades, when applied between levels of education, are extremely dependable. In the macro-educational sense, grades and Grade Point Averages are reliable, and can be used as predictors of near future grade performance. College grades appear to be consistently applied to alternating freshman classes at one university over a 27 year period of time -- so longevity adds credence to this examination of reliability.

Some writers (notably Crawford and Chamberlin) contend, however, that grading practices vary within institutions. As much evidence exists to suggest that there is a consistent grading pattern among departments and schools as there is evidence to imply that grading patterns by departments and schools are capricious. The presence of other, potential intervening conditions (calibre of each class of students, a change in

administration, the proscriptions and sanctions of university rules and regulations, etc.) inhibits more detailed analysis of these data. One can conclude that grades and grading patterns are fundamentally reliable, however. It should be stressed that even with a high correlation coefficient of .60, only one-third (.36) of the variance in grades received can be accounted for; and although one variable antedates the other, causality can not really be presumed.

VALIDITY

Reliability is only one half of the twin approaches to assessing the credibility of college grades. Both reliability and validity must pass their respective tests if the phenomenon under investigation is to be judged worthy.

Simply put 'validity' asks the question: How accurately is what we measure truly represented by our measuring devices? Obviously one way to test the validity of a grade is to examine what it symbolizes.

Warren (1971) suggests that a grade represents a composite of the evaluations covering many dimensions of the student's learning. Thus the letter "B" in philosophy may be equated with the student's ability to think discriminately and to communicate logically and cogently. The same "B" in a science course may represent his ability to handle mathematical formulae, manipulate laboratory apparati, and memorize facts. Gamson (1969), in her

study of 'Hawthorn' college, found that the social scientists tended to include the humanistic qualities of the student when arriving at a grade (e.g., the potential of the student, his enthusiasm, his willingness to work hard, his outside-of-class activities, etc.) while the faculty in the natural sciences limited themselves more to academic performance on quizzes, laboratory reports, and examinations. This may also account for the wide discrepancies in the grade distributions between the social and physical scientists at 'Hawthorn' college.

If evaluators make valid evaluations, then a group of raters should assess the same stimulus essentially the same way -- or statistically speaking, the variance about the mean should be extremely small and the distribution should be leptokurtic.

(a) Intrinsic Validity

Examine, if you will, please, the frequency distributions set out in Table 3.

TABLE 3

A Distribution of Percentage Marks on
Two High School English Themes,
a Geometry Problem and a History Paper.

%	English Theme A	English Theme B	Geometry Problem	History Paper
54	*			
55		*		
58			*	
59				
60			*	
61			*	*
62		*		
63			*	
64		*		
65	*	***		
66	*		*	*
67		**	*	*
68	****	*	**	*
69	***	**	******	
70	*******		******	**
71	**		*****	*
72	*****	*********	*******	**
73	****	****	*********	**
74	****	*********	******	**
75	*********	*********	********	****
76	******	*********	**	***
77	****	****	**********************	********
78	****	*******	*****	****
79	********		**********	***
80	********	*********	******	**
81	***	*	***	***
82	*******	******	****	*
83			***	
84		**	***	*
Total	81	80	113	42
Mean	75.11	74.95	75.09	75.55
s.d.	5.04	5.22	5.04	4.63

These histograms are not at all unlike the scores that a typical class of students have earned on an examination. The startling feature of these graphs is that they were not the scores of the students, but the evaluations by a number of experienced teachers who graded the same answer to the same question. Strach (1912, 1913a, 1913b), took a typical high school English theme, a representative History essay and a common Geometry problem and sent them to a group of high school teachers in the state of Ohio. He asked each teacher to grade the quiz, using the same standards as were currently employed in his high school. All of the high schools at this time were using percentages. Some of the high schools set the passing criterion at 65%, others at 70%, and still others at 75%. All of the graphs have been redrawn, using standard scores, where the criterion (passing) score was set at 75% and a standard deviation set at 5%.

Now, if grades are supposed to represent a degree of achievement and if the theme written by one English student reflects his command of the language, then all high school English teachers in this study should have observed the same level of mastery. As it was, one English teacher graded the paper fully four standard deviations below the average of all the English teachers.

The reader probably can think of numerous instances when he and his colleagues have discovered that they employ differing

criteria in assessing the scholastic ability of their pupils, and this phenomenon casts serious doubt on the credibility of each grade. If a single grade symbolizes a unique arrangement of evaluative devices drawn from a mosiac of all possible means to judge the student's performance, then it should follow that a collection of all these symbols, allegedly equal in meaning, should represent the level of scholarship of its owner.

Chansky (1964) argues that the grade, the essential ingredient to the GPA, has no inherently stable meaning (evidence of this has already been presented in Table 3).

Perry (1968) concurs. The hypothetical student can balance a sign of poor scholarship (a "D" in English, for example) against evidence of good scholarship (a "B" in mathematics). The average, "C", permits the student to continue his education, but does not, by any stretch of the imagination, assume that he is any better student of English.

(b) Predictive Validity

Grades are thought to reflect some facet of man's acumen. Other measures of man's intelligence and scholastic aptitude exist. These measures can be employed, by means of multiple regression, to predict the level of scholastic achievement in college, and a measure of the validity of grades, as symbols representing scholarship, can be interpreted from the results.

This is precisely what Aiken (1963) did. His work was replicated by Miller (1969) and Wilson (1970).

All three authors have separately demonstrated that the Grade Point Average does not move concommitantly with measured levels of intelligence. Aiken, for instance, examined three incoming freshman classes of 1959, 1960 and 1961 at Women's College of the University of North Carolina. Mean SAT Verbal scores increased in this three year period by some 6.8 points. Mean SAT Math scores increased 5.95 points. Moreover, high school GPA and class rank in high school at the time of graduation had increased noticeably during the same period for each successive Freshman class. Using these data, Aiken produced a multiple regression coefficient with which he could predict the year-end Freshman Grade Point Average for each cohort. This Grade Point Average, when contrasted to the actual freshman year-end GPA, showed the actual GPA to be much lower than than predicted.

Wilson poses a disturbing conundrum when he says:

> "...if the level of freshman-year grades does not increase as the quality of grade-related student input increases, then grading standards may be said to have changed--to have become more stringent during the period under consideration." pp. 47-48

It would appear as if, unlike the economy, faculties in institutions of higher learning have managed to enhance rather than devalue the worth of a grade. It seems quite evident from

these three studies that a "C" buys more student intellectual input in 1970 than it did in 1960 or 1950.

(c) Conclusions

College grades have been extant in higher learning since the middle of the nineteenth century. Numerous scales have been devised, ranging in calibration from one hundred (as in percentages) to one (as in the current Pass/No Record 'nontraditional' grade).

In this paper I have attempted to analyze the reliability and validity of this evaluative technique. The greatest number of studies so far have dwelt on the predictive association of the students' Grade Point Averages at a lower educational level to their academic performance in the next higher level. When the measured performance of large numbers of students are so studied, the criterion of reliability is easily met.

When one shifts his gaze to the grading practices of teachers, the criterion of consistency is not so easily met. Intramural grading practices are erratic and perhaps (although there is not enough evidence to sustain this statement) may even be influenced by some unknown cyclical conditions.

Very serious doubt arises when one examines the validity of grades. It would seem that a multitude of distinct characteristics are evaluated when instructors set out grades. Some professors rate recall highly, others novelty and creativity. Some professors stress verbal lucidity, still others problem solving abilities. Whatever the idiosyncratic properties

of the teacher, he resorts to placing his evaluation on a common scale. Thus myriad value judgments, each independently derived and based on ditterent value systems, are mixed and pureed into a mishmash dubbed the GPA. And, as we all know, Grade Point Averages are the pabulum which sustains higher learning.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

SECTION V. NEW FORMS OF GRADES

The American Association of College Registrars and Admissions Officers (Oliver, 1971) has recently observed an increase in the use of new systems for recording the evaluations of student progress. More common forms of recording formal evaluations (percentages, letter grades and stanines) have, at times, been replaced by a three part system (Honors-Pass-Fail); a two part system (Pass-Fail, Credit-No Credit); and a single class system (Written evaluations or Pass-No Record grades). All of these systems tend to come under the rubric "non-traditional grading practices" which seems rather ludicrous when the "traditional" system has been in existence little more than a century and a half.[1] This particular study is going to examine the Pass/Fail system and shall, temporarily, set aside other

---

(1) This depends on one's historical perspective. Harvard college, the first North American university (1636) did not practice "Grading" until the mid-eighteen hundred's, some 200 years after its foundation. Although Harvard antedates the University of New Brunswick by a century and a half (1875), it is, in turn, second to the University of Mexico (1551) in the academic procession of the western hemisphere. Mexico, first in the west, is two or more centuries junior to such European institutions of higher learning as: Universita Degli Studi, Bologna and Universite de Paris (both founded sometime in the 11th century), Universite De Montpelles (1180), Universita Degli Studi, Moderna (1175), Oxford and Cambridge Universities (both ca. 12th century), Universita Degli Studi, Naples (1224), Universita Degli Studi, Roma (1303), Ruprect Karl Universitat (Heidelberg) (1386), Aberdeen University (1494), and Albert Ludwigs Universitat (1454). It would seem reasonable that, if these western European universities had practiced term-end grading, that Harvard and the University of Mexico would have also engaged in the same tradition. Yet Harvard deferred such a procedure for 150 years. QED, the more traditional practice of a single public examination was in effect.

forms of non-traditional grading.

In this section we shall view the Pass-Fail form of grading from four perspectives. In the first instance we shall examine the extent of the popularity of the Pass-Fail grading system among students, institutions, and admissions officers. Because Pass-Fail is said to internalize the learner's motivation the next perspective will review what the literature has to say about the student working harder--or less hard--at his studies when he elects a Pass-Fail option. We will examine what has happened to the student's grade and GPA when he elects a course under the Pass/Fail option from the third perspective. Finally we will examine the issues of the student's desire to explore non-required subjects under Pass/Fail, the Pass/Fail election in relation to his academic workload, and the effect of two grading systems coexisting within the academic community.

ACADEMIA'S RESPONSE TO PASS-FAIL

(a) Popularity Among Students

The Pass/Fail grading system, so Smallwood (1935) reports, was tried briefly by the University of Michigan in 1851. The practice of Pass/Fail was again tried about a century later, when Princeton University (see Karlins, Kaplan & Stuart, 1969) offered this alternative to its upperclassmen. A review of reports from some of the earlier universities which tried Pass/Fail

(Washington, Illinois, Princeton, Brandeis, and the Berkeley campus of the University of California), seem to indicate the ratio of Pass/Fail symbols to all the other marks, expressed as a percentage, increases during the first year or two, then levels out at about 10% of the total population of undergraduate grades administered by a particular institution. Suslow (1973) found that in the six years Pass/Not Pass (P/NP) had been used at the University of California (Berkeley) the percentage of P/NP grades to all undergraduate grades stabilized between 10% and 11%.

Johansson, Rossmann and Zandell (1971) report that

> "...a loss of the popularity of the ungraded option after the first year was the outstanding finding from the data. Almost half of the 1969 registrants took a S-U [Successful-Unsuccessful] course; this percentage dropped to about one-third in 1970. Women and freshmen, in particular, tended to take the S-U option less frequently in 1970 even though freshmen were allowed an unrestricted number of S-U courses in 1970, only about 7% of the freshmen class opted for more than one ungraded course during the spring term, indicating that students might be concerned about not overloading their academic records with ungraded courses." p. 275

(b) Popularity Among North American Colleges and Universities

The Oliver report (1971) reports that 60% of the 1,301 North American colleges and universities who responded to the AACRAO questionnaire had indicated that they were now "...using some form of non-traditional grading" in their record keeping systems.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

Three of the twenty Canadian institutions of higher education in Oliver's survey reported using a combination of traditional and non-traditional grading. Non-traditional grading of undergraduate work is a very recent phenomenon, as most of the responding institutions in North America said the change had occurred within the last two years. It would be too early to estimate how much of a foothold non-traditional grading will have had on the more traditional systems.

(c) Popularity Among Admissions Officers and Admissions Committees

Lawrence Hanlon (1964) indicated that, as the admissions officer for his medical school, he and many of his counterparts are confronted with a number of problems which result from non-traditional grades. He said:

> "Knowledge of the comparative standing is a necessity in evaluating the academic attainment of students from the few schools that do not report in terms of any of the usual grading systems... In my experience with some of these unusual methods of porting academic standings, there is frequently so much uncertainty about a student's academic ability that he simply cannot be considered for admission." p. 95

Hofeller (1972), in her survey of three hundred and ninety-one deans of graduate and professional schools in arts, sciences, education, medicine, law and nursing, found that the clear majority of graduate nd professional schools report the practice of re-computing the Grade-Point Averages of applicants, not only

to gain a measure of the student's performance in his major or senior year, but also to justify the GPA for possible inclusion of Pass/Fail grades. Hofeller said that

> "...the outstanding student, whose credentials include exceptional test scores, an academically prestigious college, and outstanding recommendations, may suffer little jeopardy from a choice of non-traditional grades. However, a well qualified, but no(sic) noticeal unique applicant who opts for Pass/Fail grade may well be discriminated against in favor of his potentially less able but more traditional peer. Although this trend might alter drastically, its present direction is of immediate concern to contemporary students." p. 11

Rossmann (1970) obtained responses from thirty graduate schools and fifteen professional schools to his question concerning Satisfactory-Unsatisfactory (S-U) grades. He learned that both sets of deans do not like to see a proponderance of S-U or non-traditional grades on a transcript. Rossman contends that a large number of non-traditional grades weakens the candidate's chances of being admitted -- expecially if these grades were in his major.

This same concern extends even to Admissions Officers at the undergraduate level. Oberteuffer (1970), in her survey of 213 undergraduate institutions found that 57 of the 149 respondents (approximately 1/3) felt that non-traditional grading would be definitely detrimental to the admissions possibilities of a prospective student.

One undergraduate admissions officer responded to the

Oberteuffer questionnaire by saying:

> "Grades represent diligence, time-budgeting, self-discipline, effort, and perseverance. What are you afraid of?"

It becomes evident that Admissions Officers at the undergraduate and graduate levels are not overjoyed at the thought of reviewing transcripts containing a large number of non-traditional grades.

STUDENT MOTIVATION

A non-traditional grading system, such as the Pass/Fail system, is said to facilitate learning because it shifts an emphasis on learning away from the external motivation (getting a 'good' grade) towards internal motivation ('I want to learn something about this subject'). Under the non-traditional grading system, the student can look within himself for his own measure of cognition, that is, he can see for himself what it is he knows and does not know, and furthermore, precisely what it is that he want to know. The desire to know, rather than the desire to obtain a pleasing grade becomes the dominant drive.

Raimi (1967) puts this argument cogently when he said:

> "...the most telling criticism of all grading systems is this: that the incentive and discipline they foster are incentives to beat the grading system itself (rather than towards scholarship), and discipline in the direction of safe conformity (rather than in the habits of learning)." p. 311

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

Proponents maintain that the non-traditional form of grading also allows the teacher to put more emphasis on teaching. Marshall (1971) suggests that individual descriptions of the students (one form of non-traditional grading) serves a double function: it provides all the desiderata many graduate school admissions committees think they need (see Oliver, 1971, pages 29 & 30), and it permits the instructor to devote himself, in a small way, to the individual student. The reader may look askance at this last assertion, but a study by Page (1958) offers evidence that perscnal feedback, even in the inconspicuous medium of written comments on an objective examination, is a strong motivation to learning.

Opponents of the Pass/Fail system have suggested that such a system tends to create a feeling of indolence and slovenlyness among students. What evidence does one have from the literature to support this?

Karlings, Kaplin and Stuart (1969) in their study at Princeton Unversity, find that students reported 72% said they felt they had worked closer to their capacity "...in a numerically graded course rather than one graded Pass/Fail".

Further on in the report, these authors indicated that 79% of these 1,006 responding students said that "If I knew my course were being graded 1-7, I would work harder.", and 74% felt that taking a Pass/Fail course had allowed them more time to study for

other courses.

The student survey at Princeton also reported that their students did, on the average, 80% of the assigned readings in the graded courses buy only 61% of the assigned readings in the Pass/Fail course. Students also reported attending 85% of the graded courses' lectures and precepts, but only 74% of the Pass/Fail courses' lectures and precepts.

Stallings and Smock (1971), in a poll at the University of Illinois, learned that the most of the students did only a minimum of work required for a "D". Bain, et al. (1971) at Ohio University report that only 31% of the 344 responding students worked hard enough to pass the course, and an additional 62% worked hard enough to meet their Grade Point Average.

But contrary evidence also exists. In an earlier study by Stallings, Wolff and Maehr (1969) found that the students electing Pass/Fail carried a significantly higher course load ($F = 11.078$, $df = 1,150$, $P > .01$), which suggests that while the Pass/Fail student may have eased up on one particular course, he compensated for it by increasing his course load. Suslow (1973) also found that students with high levels of scholarship elected an extra course under Pass/No Pass (P/NP) in order to graduate sooner.

Merryman & Kirman (1971) reported that 80% of the respondents (N=63) said that the Credit/No Credit reduced the

pressure for grades and about 60% said they did as much work for the Credit/No Credit course as they did for courses graded in the conventional manner.

Jesse and Simon (1971) compared the 'diaries' of medical students at the University of California (San Diego) to similar ones kept by medical students at the University of Illinois. Average times reported studying and in related academic activities between the medical students at these two universities were essentially the same, even though the San Diago Medical School was completely under a non-traditional grading system.

A continuing debate has occurred in academic circles concerning the scholastic performance of students who elect courses under a Pass/Fail grading system. Proponents of the non-traditional grading system claim that students learn just as much and perform just as well as students laboring under the more traditional grading system. Opponents to the non-traditional system tactfully suggest that students do not learn as much when the motivation to achieve a grade has been removed. In reviewing the literature, I considered two issues: the grades these two groups of students received when the instructor was aware that both systems were used in his course, and the grades these two groups of students received when the instructor was blind to which students elected Pass/Fail. It is my assumption that if the instructor knows which of his students are under a Pass/Fail system, he will be either more punitive or more lax in his

grading, depending upon his personal attitudes towards the non-traditional grading systems.

Smith (1971) at the University of Illinois, Chicago Circle, examined the grades received by these students who elected a Pass/Fail on a the grades of 2,335 students who elected the Pass/Fail. The instructors in every case were blind to which students elected this particular option. The grades received by these students electing a Pass/Fail were better than the grades received by the non-Pass/Fail students in the same courses. Smith went on to point out, however, that the students who elected Pass/Fail received a Pass/Fail grade which would have been lower than their Grade-Point Average.

Barbara Von Wittich (1972) studied the students in beginning foreign language courses (French, German, Spanish and Russian). Three hundred and five students of the 895 students in these courses had elected Pass/Fail and had they received conventional grades, their grades would have been lower than the 570 who chose to remain under the traditional system. The instructors who awarded these grades did not know which of their students elected the Pass/Fail option.

Karlings et al. (1959), at Princeton University discovered in their survey of its students that Pass/Fail students would have earned a grade which was one full grade lower than their GPA.

PASS-FAIL GRADING SYSTEMS: LITERATURE REVIEW

Sgan (1970) found that students at every class level except the Senior level who elected a Pass/Fail course at Brandeis University had significantly lower Grade Point Averages than the average GPA of all the members of that class. Thus Freshmen who elected P/F courses had a GPA of 2.34 (on a four point scale where A=4.0), and the class GPA for all Freshmen was 2.90.

Perhaps the most extensive field experiment on the topic of Pass/Fail and academic performance occurred at the Cortland campus of the State University of New York. Gold, Reilly, Silberman and Lehr (1971) were able to divide a Freshman class into three groups and a Junior class into two groups. The Freshman class was divided by its choice of Pass/Fail election: where all the Freshmen in the first group chose to have all their courses under Pass/Fail; the Freshmen in the second group chose to have only one course under Pass/Fail, and the Freshmen in the third group chose to have none of their courses under Pass/Fail. The Junior standing students were given only one option: to elect one course under Pass/Fail, and they were then randomly divided, where some were allowed to have that option while the rest were denied it. The course instructors were not informed of the student's choice or his group membership. The instructors submitted letter grades (A,B,C,D,E) for all students. These grades were given the usual weighting by the experimenters (A=4.0, B=3.0, etc.) and a Grade Point Average (GPA) was computed for members of each group. A difference of means test (t-test)

was applied to these data, the results of which appear in Table 4.

Table 4

A Difference of Means Test of Freshmen who took all their Courses Under Pass/Fail to Freshmen who took all their Courses Under the Letter Grade, by Semester.

	EXPERIMENTAL GPA	Number	CONTROL GPA	Number	Signifcance
Freshman Year Fall of '67	1.67	29	2.26	27	$p < .01$
Sophomore Year Fall of '68	2.28	22	2.72	24	$p < .01$
Junior Year Fall of '69	2.68	18	2.85	20	n.s.

Significant differences in academic performance, as measured by traditional grades, appeared in the Freshman and Sophomore years of the Gold, et al. study. This difference was erased in the Junior Year. One factor the experimenters could not control was attrition, either for voluntary or institutionally enforced reasons. Thus the eleven subjects in the experimental group and seven subjects in the control group who withdrew may have actually been different in their academic perofrmance.

Gold and his colleagues then compared their second group of Freshmen, those who were allowed to elect only one course under Pass/Fail, to three other experimental groups: (1) Freshmen who were denied the one-course P/F option, (2) Junior level students who had a single course election under P/F, and (3) Junior level

students who were denied the one-course election under P/F. The semester and grades were collected and weighted in the same manner described above. Table 5 presents difference of means test results (t-test) for two categories: (1) the course which the experimental subjects took under P/F compared to the course which the control subjects wanted to take under P/F but were denied the option, and (2) all the other courses elected by the students in these groups. S

Table 5

A Difference of Means (t-test) of (1) Single course Grades of (a) Those who had the P/F option to (b) Those who were denied it, and (2) the Grades Earned in all the other Courses by (a) Those who had the Single Course P/F option, and (b) Those who were denied the P/F option, by Class Level

	Single Course Election					All Other Courses		
	Exp'l GPA	N	Cont. GPA	N	Sig.	Exp'l GPA	Cont. GPA	Sig.
Freshmen	1.67	71	1.83	69	n.s.	2.23	2.15	n.s.
Junior	2.07	89	2.40	80	$p < .05$	2.78	2.80	n.s.
Total	1.89	160	2.15	151	$p < .05$	2.54	2.52	n.s.

Junior level students obviously earned lower grades in the one course where they had the P/F option, but did as well as their control counterparts in all the other courses elected that semester. This suggests a laxity of motivation for the grade in the P/F course, but not necessarily a deterioration of effort in other courses.

There has been some reports in the literature where it was

not known if the instructor did indeed know of the students' elections. Quann (1972) examined grades over a three semester period of time and reported that more students electing P/F sustained failing grades (25.9%) than those who chose to remain in the more conventional system (7.7%). Conversely, fewer of the students who elected Pass/Fail earned A's (3.8%) when compared to the more traditional group (21.3%). A chi-square distribution of these two distributions is even more striking. ($x^2 = 29.88$, df = 4, $P < .001$).

Stallings and Smock (1971) report that the Pass/Fail students received a grade which was 0.7 lower (on a four-point scale) than the non-Pass/Fail students when these students took the same course. Egan (1970) found that students at every class level except the Senior level who elected a P/F course at Brandeis University had significantly lower GPA's than the average GPA of all members of that class. Thus Freshmen who elected P/F courses had a GPA of 2.34 (on a 4-point scale), and the class Grade Point Average for all Freshmen was 2.90, significantly higher.

THE EFFECT ON GRADE-POINT AVERAGES

Quann (1971) reports that the differences between Grade-Point Averages of those who elected Pass/Fail at Washington State University and those who did not were very minor. Stallings, Holtz and Mahr (1969) found, at the University of Illinois, that

the students who had elected a Pass/Fail course earned a significantly higher Grade-Point Average than those who had not elected Pass/Fail.

EXPLORATION, ACADEMIC WORKLOAD and GRESHAM'S LAW

(a) Exploration

In a survey of students electing Pass/Fail courses at the University of Illinois (Urbana), Stallings and Smock (1971) learned that the responding students expressed favor for the Pass/Fail system when exploring courses which might be academically demanding. In a sister university in the State of Illinois, Reiner & Jung (1972) discovered that the P/F was used more often in non-major courses, but those who used the P/F option (1) did not achieve as well as their previous academic record would indicate they should have and (2) did less well in the P/F courses than classmates in the same courses under the traditional grading system.

What may have been the opinion on one campus is not shared by the opinions of students on other campuses. Quann (1972) reports that the students at Washington State University did not use the Pass/Fail system to venture into untried ares. Bain, Hales and Rand (1972), in a survey of 344 students taking the Pass/Fail option at Ohio University, discovered than only 9%

(N=31) of the respondents said they took a course outside their required subjects because Pass/Fail permitted them to do so without experiencing severe competition. Smith (1971), discovered that few of the 2,335 students at the University of Illinois, Chicago Circle, who had elected Pass/Fail used the option in their elective courses and that only one student in a hundred went on to take a second elective course on a Pass/Fail basis. Suslow (1973), while he had a difficult time determining which course elections were outside the student's usual curriculum, did have the temerity to suggest that very few students used the Pass/Not Pass (P/NP) system in order to explore unknown subject areas.

(b) Academic Workload

Another motivational element considered here is that the non-traditional grading system reduces the academic pressure on the student. This presumes that a) students are carrying a full load and b) that pressure is exerted by every element of the load in a more or less uniform manner on the student. Granting these assumptions, common experience has indicated that the nature of the tasks to be performed by a student carrying a full load appearing to be increasing in quantity and quality, and that the tolerance threshold of human emotive stamina, among many of the students, appears to have been reached. The functional grading of the non-tradition system here is to permit a cushion in this

demand on the student. (see Johansson 1971; Morishima & Micek, 1970; Phillips, 1962; Trapesser, et al., 1968; Thayer, 1971; Stallings, et al. 1970).

(c) Gresham's Law and Academic Currency

Perry (1968), President of Whitman College, reported on his institution's decision to abandon the conventional grading system for a three-part system of high honors (HH), honors (H) and pass (P). The experience at Whitman indicated that many graduate schools were rejecting very well qualified young students because these schools equated a "P" with a barely passing grade (e.g., "C" or "D").

Noble (1971), of Brown University reported that when a five-part letter grading system co-existed with the Pass/Fail system and "P" was set at "D", many students chose to remaing with the five letter system.

This indicates to me that two currencies of the academic realm cannot co-exist as interchangeable currencies in a single college or university any more than two currencies can co-exist within a single nation. Gresham's Law holds for the academic community, and it appears that the non-traditional currency is the weaker coin.

VI. RECAPITULATION

Grading practices, something which began a little more than 200 years ago, are ir to be an extremely efficient means of measuring, weighing and sorting students. Such means are not entirely without dysbenefits, however, for some students have learne[illegible] shift their attention from learning per se to earning a grade.

Within the last decade students and educators have tried new, so called non-traditional, forms of grading. Fundamentally this new grading practice reduces the number of 'grades' or pigeon-hole catagories from as many as one hundred (as in the case of percentages) or five (as in the case of letters) to one (written evaluations and Pass/No Record), two (Pass/Fail, Satisfactory-Unsatisfactory, Credit/No Credit), three (Honors-Pass-Fail), or four (High Honors-Honors-Pass-Fail).

The literature on one of these newer forms of grades, Pass/Fail, shows that this grading practice has gained a wide acceptance among many students. The general practice in many colleges and universities is to limit the number of Pass/Fail courses a student may take to one per term and in some cases it may be limited to the class standing of the individual or his cumulative GPA.

The literature suggests that the source of the popularity of

the Pass/Fail system among students stems from their need or wish to re-distribute their time and effort among all their course elections in such a way as to concentrate on the courses elected under the conventional grading system while slighting the course (or courses) elected under the Pass/Fail system. Students do not seem to use this grading system as a vehicle for exploring subjects outside their major area of interest. Students do not do all of the assigned readings, nor attend as many classes. Students admit they do not work as hard in their Pass/Fail courses as they do in their more conventionally graded courses.

These conclusions soundly condemn the non-traditional grading system, but one must keep in mind that students are rational human beings working towards certain goals and they naturally will employ every ligitimate means of 1) improving their GPA, 2) graduating and 3) gaining either admission to a graduate or professional school, or a job. Modern society has increased the tempo of learning and the pressure to excel. This, in turn, raises a question not posed in this literature review, but one which I personally feel bears examination. Have we, as educators, increased the academic pressure and academic load on the student to such a point that the student no longer is capable of learning? Are our efforts to increase the academic standards of our own institutions becoming counter-productive?

BIBLIOGRAPHY

Aiken, Lewis R., Jr. "Interdepartmental Variability and Student Expectations of College Grades." Educational and Psychological Measurement. v. 24 (1964) pp. 823-829.

Aiken, Lewis R. Jr., "The Grading Behavior of College Faculty." Educational and Psychological Measurement, v. 23, #2 (1963) pp. 319-322.

American College Testing Program. College Student Profiles. Iowa City: American College Testing Program, 1966.

Anderson, Scarvia B. "Estimating Grade Reliability." Journal of Applied Psycholgoy v. 37 (1953) pp. 461-464.

Astin, Alexander W. Predicting Academic Performance in College: Selectivity Data for 2,300 American Colleges. New York: Free Press, 1971.

Bailey, Robert L. (Chairman) AACRAO Sub-Committee on Non-Traditional Grading Patterns. Park Forest South, Illinois. Office of Admissions & Records, Governors State University, 1972 (mimeograph).

Bain, P. T.;Hales, L. W. & Rand, L. P. "Does Pass-Fail Encourage Exploration?" College and University v. 47 (Fall, 1971) pp. 17-18.

Baird, L. L. & Fiester, W. J. "Grading Standards: the Relation of Changes in Average Student Ability to the Average Grades Awarded." Educational Testing Service (May, 1971)

Bardwick, Judith M. Psychology of Women: A Study of Bio-Cultural Conflicts New York: Harper & Row, 1971.

Bass, B. M. "Intrauniversity Variation in Grading Practices." Journal of Educational Psychology v. 42 (1951), pp. 366-368.

Battle, E. "Motivational Determinants of Academic Competence." Journal of Personality and Social Psychology. v. 4 (1966) pp. 634-643.

Becker, Howard; Greer, Blanche; and Hughes, Everett. Making the Grade: The Academic Side of College Life. New York: John Wiley and Son, 1968.

Berry, G. et al. "Use of Pass/Fail Grading System for a Graduate Course. Correspondence: April 23, 1971.

Bess, James L. & Hayes, Charlene. "Bene[illegible] Beneficiaries of Innovation: A Research Model." Insti[illegible]

BIBLIOGRAPHY

Research and Institutional Policy Formulation, 11 Annual Forum of the Association for for Institutional Research Claremont, California: Association Institutional Research 1971.

Birney, Robert G. "The Effects of Grades on Students" Journal of Higher Education v.35 (February, 1964) pp. 96 - 98

Bramer, George R. "Grading and Academic Justice." Improving College and University Teaching v. 18 #1 (Winter, 1970) pp. 63 - 65.

Brannigan, G.G. & Taylor, A. "Sex Differences in Adaptive Styles." The Journal of Genetic Psychology v. 119 (September, 1971) pp. 143 - 149.

Bundy, McGeorge. "An Atmosphere to Breathe: Woodrow Wilson and Life of the American University College." Woodrow Wilson Foundation, 1959.

Caldwell, E. "Sex Biases in College Grading." Journal of Educational Measurement v.4, #3, (1967) pp. 129-132.

Carter, F.S. "How Invalid Are Marks Assigned By Teachers?" Journal of Educational Psychology, v.43 (1952), pp. 218 - 228.

Chamberlin, Dean; Chamberlin, F. S.; Drought, N. E. & Scott, W. E. Adventures in American Education: Did They Succeed in College? New York: Harper & Brothers, 1942.

Chansky, Norman M. "A Note on the Grade Point Average in Research." Educational and Psychological Measurement, v. 24 (1964) #1, pp. 95 - 99.

Chansky, Norman M. "The X-Ray of the School Mark." The Educational Forum (March, 1962) pp. 347-352.

Clark, Edward L. "Reliability of Grade-Point Averages" Journal of Educational Research, v.57 #8 (April, 1964) pp. 428 - 430.

Clanch, S. J. "Effects of P-F Grading on Quality Point Averages." College and University v. 47 (Winter, 1972) pp. 93-105.

Clements, William H. "Pass-Fail: Were Your Hypotheses Valid?" College and University v. 46 (Summer 1971) pp. 535-542.

College Entrance Examination Board (Editorial Office) Effects of Coaching on Scholastic Aptitude Test Scores. New York: College Entrance Examination Board, 1965.

BIBLIOGRAPHY

Committee on Educational Policy and Courses of Instruction, Berkeley Division "A Report on Scholarship Grades--University of California, Berkeley: Per Cent Distribution for Fall Terms 1961, 1963, 1965, 1967, 1969." Berkeley, California: Office of Institutional Research, University of California, October, 1970.

Committee on Educational Policy and Courses of Instruction, Berkeley Division "A Report on Scholarship Grades--University of California, Berkeley: Per Cent Distribution for 1970-1971." Berkeley, California: Office of Institutional Research, December, 1971.

Committee on Educational Policy, Santa Cruz. "Report on Grading at the University of California, Santa Cruz." Santa Cruz, California: University of California, 1970.

Constantinople, A. "Perceived Instrumentality of the College as a Measure of Attitudes Toward College." Journal of Personality and Social Psychology. v. 5 (1967) pp. 196-202.

Coull, K. W. "Student Marks Analysis II: A Report on the Distribution of Students' Marks for the Years 1970-71 and 1971-72." Edmonton, Alberta: Office of Institutional Research and Planning, May 1, 1973 (Mimeograph)

Crawford, Albert B. "Rubber Micrometers." School and Society. v.32 #816, (1920) pp. 233 - 240.

Crooks, A. D. "Marks and Marking Systems, A Digest." Journal of Educational Research. v. 27, #4 (1933) pp. 259-272.

Cross, K. Patricia (ed). "On Creativity." The Research Reporter. v. II, (1967) pp 1-4.

Cross, K. Patricia. The Undergraduate Woman. Washington, D. C.: American Association for Higher Education, March 15, 1971 (Research Report Number 5).

Downing, Marjorie "Variations in College Grading Systems." Journal of Higher Education v. 35 (February 1964) pp. 98 - 100.

Dressel, Paul L. & Nelson, Clarence H. "Testing and Grading Policies." in Dressel, Paul L. & Associates. Evaluation in Higher Education. Boston, Mass.: Houghton Mifflin Co., 1961.

Edmiston, R.W. "Do Teachers Show Partiality Toward Boys or Girls?" Peabody Journal of Education. v.20 (1943) pp. 234-238.

Eells, W.C. "Reliability of Repeated Grading of Essay Type Examinations." Journal of Educational Psychology. v.21 (1930) pp. 48 - 52.

BIBLIOGRAPHY

Elbow, Peter H. "More Accurate Evaluation of Student Performance." Journal of Higher Education. v. 40, #3 (March 1969) pp. 219 - 230.

Elliott, Corinne; Warren, Richard J.; & King, Horace C. "Evaluating Evaluation -- Grading Practices." College and University, (Summer, 1970) pp. 349 - 359.

Engle, T. L., "Comparative Study of First and Final Marks." School Review, v. 40 (1932) pp. 61-66.

Fricksen, Stanford C. "Pass-Fail Grading." Memo To The Faculty #22 Ann Arbor, The Center for Research on Learning and Teaching, (April 1967)

Fala, Michael A., "Dunce Cages, Hickory Sticks, and Public Evaluations: The Structure of Academic Authoritarianism." Madison, Wisconsin: The Teaching Assistant Association, University of Wisconsin, 1968.

Feldmesser, Robert A., "The Positive Functions of Grades." American Educational Reserach Association , v. 53 (Winter, 1972) pp. 66-72. (Also ERIC report # ED 049 704).

Feldmesser, Robert A., The Option: An Analysis of an Educational Innovation Hanover, New Hampshire: Dartmouth College Press, 1969.

Gamson, Zelda F. "Performance and Personalism in Student Faculty Relations: Departments of Natural Sciences and Social Sciences." Sociology of Education. v. 40, #4 (Fall, 1969) pp. 279-301.

Gold, Richard M.; Reilly, Anne; Silberman, Robert & Lehr, Robert. "Academic Achievement Declines Under Pass-Fail Grading." Journal of Experimental Education v. 39, #3 (Spring, 1971) pp. 17-21.

Goldstein, K. M. & Tilker, H. A., "Attitudes Toward A-B-C-D-F and Honors Pass-Fail Grading Systems." Journal of Educational Research. v. 65, #3 (November, 1971) pp. 99-100.

Goodman Paul. "In What Ways Does the Present Marking and Credit System Inhibit or Promote Learning?" in Smith, G. Kerry (ed) Current Issues of Higher Education: Undergraduate Education. Washington, D. C.: American Association for Higher Education, 1964 (pp. 123-125).

Haagen, C. Joss. "The Origins of a Grade." Journal of Higher Education. v. 35 (February, 1964) pp. 89-91.

BIBLIOGRAPHY

Hanlon, Lawrence W., "College Grades and Admission to Medical School." Journal of Higher Education. v. 35 (February, 1964) pp. 93-96.

Haskins, Charles H., The Rise of Universities. Ithaca, New York: Cornell U Press, 1965.

Heffernan, James M. "The Credibility of the Credit Hour: The History, Use, and Shortcomings of the Credit System." Journal of Higher Education v. 44, #1 (January, 1973) pp. 61-72.

Heininger, Joseph F., "The Nongraded Transcript." Journal of General Education v.23 (January, 1972) pp. 301-308.

Hewitt, Raymond G., The Status of Pass Fail Options at Twenty-Two Colleges and Universities. Amherst, Mass.: University of Massachusetts, Office of Institutional Research, 1967.

Hills, J. R., "Predictions of College Grades for all Public Colleges of a State." Journal of Educational Measurement. v. 1, #2 (December, 1964) pp. 155-159

Hills, J. R.; Klock, J. A. & Bush, M., "The Use of Academic Prediction Equations with Subsequent Classes." American Educational Research Journal. v. 2, #4 (November, 1965) pp. 203-206.

Hodgkinson, Harold L., "Pass-Fail and the Protestant Ethic." The Chronical of Higher Education. v. VII, #12 (December 11, 1972) p. 8.

Hofeller, Margaret A., " A Survey of Graduate School Attitudes Toward Non-Traditional Grading Systems: Preliminary Report." Hempstead, Long Island: Hofstra University Center for the Study of Higher Education, 1972. (also ERIC # ED 056 677)

Holland, John L., "Prediction of College Grades from Personality and Aptitude Variables." Journal of Educational Psychology. v. 51 (1960) pp. 245-254.

Hoyt, P. Donald. "The Relationship Between College Grades and Adult Achievement." ACT Research Report Number 7, Iowa City: American College Testing Program, 1965.

Hoyt, P. Donald. "Rationality and the Grading Process." Educational Record v. 51 (Summer, 1970) pp. 305-309.

Hyman, Lawrence W., "Grades and Academic Standards." Washington, D. C.: U. S. Department of Health, Education and Welfare, ERIC report # ED 036 275.

BIBLIOGRAPHY

Jesse, William F. & Simon, Harold J., "Time Utilization by Medical Students on a Pass/Fail Evaluation System." Journal of Medical Education. v. 40 (April, 1971) pp. 275-280.

Johansson, C. B.; Rossmann, J. E. & Sandell, S. D. "The S-U Grading Option at Macalester." Educational Record. v. 52 (Summer, 1971) pp. 273-276.

Jordon, David Starr. The Trend of the American University. Stanford, California: Stanford University Press, 1929.

Kappel, F. R. From the World of College to the World of Work. New York: Telephone and Telegraph Company, 1962.

Karlins, Marvin; Kaplan, Martin & Stuart, William, "Academic Attitudes and Performance as a Function of Differential Grading System: an Evaluation of Princeton's Pass-Fail System." Journal of Experimental Education v. 37, #3 (Spring, 1969) pp. 38-50.

Kelly, Fldon G., "A Study of Consistent Discrepancies Between Instructor Grades and Test Results." Journal of Educational Psychology. v. 49, #6 (1958) pp. 328-334.

Kelly, Samuel P. & Thompson, Ralph. "Grading and the Nature of the Discipline." Journal of Higher Education. v. 39, #9 (December, 1968) pp. 517-518.

Kirby, B. C., "Three Error Sources in College Grades," Journal of Experimental Education. v. 31, #2 (December, 1962) pp. 212-218.

Kirschenbaum, Howard; et al. Wad-ja-get? The Grading Game in American Education New York: Hart Publishing Company, Inc., 1971.

Klugh, H. F. & Bierley, R., "The School and College Ability Test and High School Grades as Predictors of Achievement." Educational and Psychological Measurement. v. 19 (1959) pp. 625-626.

Kraft, B. & Lundquist, G. "Pass/Fail Grades Compared to Traditional Grading Procedures." SPATE: Journal of the Student Personnel Association for Teacher Education. v. 10 (September, 1971) pp. 1-13.

Lacey, Oliver, L., "How Fair Are Your Grades?" AAUP Bulletin. v. 46, #3 (September, 1960) pp. 281-283.

Lannholm, Gerald V.; Marco, Gary L. & Schrader, William B. "Cooperative Studies of Predicting Graduate School Success."

BIBLIOGRAPHY

Graduate Record Examination Special Report 68-3. Princeton, New Jersey: Educational Testing Service, 1968.

Lewis, W. A., "Early Prediction of College GPA Using Precollege Grades." Journal of Educational Measurement. v. 3, #1 (Spring, 1966) pp. 35-36.

Lovell, K., "Examinations and Marking." in Layton, David (ed) University Teaching in Transition. Edinburgh: Oliver & Boyd, 1968. pp. 135-146.

Maney, Charles A., "Sex Bias in College Marking." Journal of Higher Education v.4 (1933) pp. 29-31.

Marshall, Max S., "This Thing Called Evaluation." Educational Forum. v.23, #1 (November, 1958) pp. 41-53.

Marshall, Max S., "Why Grades Are Argued." School and Society. v. 99, #2335 (October, 1971) pp. 350-353.

Marston, Albert R., "It is Time to Reconsider the Graduate Record Examination" American Psychologist. v. 26 (July, 1971) pp. 653-655.

McClure, Robert F., "Birth Order, Income, Sex, and School Related Attitudes." Journal of Experimental Education. v. 39, #4 (Summer, 1971) pp. 73-74.

McLean, R. Alan. "Student Marks Analysis: A Report on the Distribution of Students' Marks for the Year 1969-70." Edmonton, Alberta: Office of Institutional Research and Planning, January 27, 1971.

Melville, George L. & Stamm, Eleanor. "The Pass-Fail System and the Change in the Accounting of Grades on Comprehensive Examinations at Knox College." Galesburg, Office of Institutional Research, Knox College. 1967. (also ERIC report # ED 014 788).

Merryman, Mary F. & Kerwin, Leo. " A Three Semester Study of the Credit-No Credit Option in the College of Arts and Sciences, University of Missouri, Kansas City." Washington, D. C.: U. S. Department of Health, Education and Welfare, ERIC report # ED 051 750.

Miller, Richard H., "Students Show a Preparation Increase But No Increase in Grades Was Shown." College and University. v. 45, #1 (Fall, 1969) pp. 28-30.

Miller, Stuart. Measure, Number, and Weight: a Polemical Statement of the College Grading Problem. Ann Arbor, Michigan:

BIBLIOGRAPHY

University of Michigan Center for Research on Learning and Teaching, March, 1967.

Milton, Ohmer. "Pass-Fail Grading." Teaching-Learning Issues. Knoxville Tennessee: University of Tennessee Learning Resources Center publication #974-2459, 1967.

Morishima, James K. & Micek, Sidney S. "Pass-Fail Evaluation: Phase I: Grade Analysis." Seattle, Washington: Office of Institutional and Educational Research, University of Washington, April 10, 1970

Morishima, James K. & Micek, Sidney S. "Pass-Fail Evaluation: Phase II: Questionnaire Analysis." Seattle, Washington: Office of Institutional and Educational Research, University of Washington, April 27, 1970.

Morishima, James K. & Micek, Sidney S. "Pass-Fail Evaluation: Summary." Seattle, Washington: Office of Institutional and Educational Research, University of Washington, April 28, 1970.

Noble, Milton E., "Advantages and Disadvantages Associated with the "No-Fail" Grading Systems." College and University. (Summer, 1971) pp. 717-726.

Nyberg, V. R. (Chairman), "The G. F. C. ad hoc Committee to Study the Nine-Point Grading System." The University of Alberta (mimeograph) ca. 1972.

Oberteuffer, Margaret., "College Reactions to a Pass/Fail Threat." National Association of College Admissions Counsellors Journal (National ACAC Journal) v. 15, #1 (May, 1970) pp. 8-9.

Odell, C. W., "High School Marking Systems." School Review. v. 33, #5 (1925) pp. 346-354.

Office of Institutional Reserach, University of California, Berkeley. "A Theoretical Examination of the Effect of Large Scale Enrollments in Pass/Not Passed Courses on the Grade Point Average of Undergraduate Students at Berkeley." Berkeley, California: Office of Institutional Research, December, 1968 (Mimeograph)

Oliver, E. E. (Chairman). The AACRAO Survey of Grading Policies in Member Institutions. Washington, D. C.: American Association of Collegiate Registrars and Admissions Officers, 1971.

Page, E. B., "Teacher Comments and Student Performance."

BIBLIOGRAPHY

Journal of Educational Psychology. v. 49 (1958) pp. 173-181.

Pederson, Eldor. "Undergraduate Pass/Not-Pass Grading at the University of California, Berkeley: Fall Quarter, 1966." Berkeley, California: Office of Institutional Research, University of California (Berkeley), May, 1967.

Perry, Louis B., "College Grading: A Case Study and its Aftermath." Educational Record. v. 49, #1 (Winter, 1968) pp. 78-84.

Priest, Robert F., "Why College Students Favor Grading Reforms." Journal of College Student Personnel. v. 12, #2 (March, 1971) pp. 120-125.

Quann, Charles James., "Pass-Fail-What are the Trends?" College and University (Summer, 1970) pp. 484-493.

Quann, Charles James., "Pass/Fail Grading: Panacea or Placebo?" National ACAC Journal. v. 16, #4 (February, 1972) pp. 11-13.

Quann, Charles James., "The Pass/Fail Option: Analysis of an Experiment in Grading." College and University. (Summer, 1971) pp. 542-549.

Raimi, Ralph A., "Examinations and Grades in College." AAUP Bulletin. v. 53, #3 (September, 1967) pp. 309-317.

Rasmussen, Glen R. "The Change in Grading Standards in Undergraduate Education Courses." Journal of Higher Education v. 40, #4 (April, 1969) pp. 311-317.

Reiner, J. R. & Jung, L. B., "Enrollment Patterns and Academic Performance as a Function of Registration Under a Pass-Fail Grading System." Interchange v.3, #1 (1972) pp. 53-62.

Rogers, Herbert W., "The Reliability of College Grades." School and Society v. 45, #1170 (May 29, 1937) pp. 758-760.

Rosenthal, R. and Jacobsen, L. Pygmalion in the Classroom: Self Fulfilling Prophecies and Teacher Expectations. New York: Holt, Rinehart and Winston, 1969.

Rossmann, Jack E., "Graduate School Attitudes to S-U Grades." Educational Research. v. 51 (Summer, 1970) pp. 310-313.

Schmuck, Richard & Van Egmond, Elmer., "Sex Differences in the Relationship of Interpersonal Perceptions to Academic Performance." Psychology in the Schools. v. 2 (1965) pp. 32-40.

BIBLIOGRAPHY

Sgan, Mathew R., "First Year of Pass/Fail at Brandeis University." Journal of Higher Education v.40 (Fall, 1969) pp. 135-144.

Sgan, Mathew R., "Letter Grade Achievement in Pass Fail Courses." Journal of Higher Education v.41 (November, 1970) pp. 638-644.

Shaw, M. C. & McCuen, J. T., "The Onset of Academic Under-Achievement in Bright Children." Journal of Educational Psychology. v. 51 (1960) pp. 103-108.

Shontz, Howard. "The Grading System at UCSC - A Critique." in "Pass-Fail--What are the Trends?" College and University. (Summer, 1970) p. 494.

Siegelman, Marvin, "SAT and High School Average Predictions of Four Year College Achievement." Educational and Psychological Measurement v.31 (Winter, 1971) pp. 947-950.

Singer, Jerome E. "The Use of Manipulative Strategies: Machiavellianism and Attractiveness." Sociometry. v. 27, #2 (June, 1964) pp. 128-150.

Smallwood, Mary L. An Historical Study of Examinations and Grading Systems in Early American Universities. Cambridge, Massachusetts: Harvard University Press, 1935.

Smith, Bardwell; Porter, L. E. & David, H., "Carleton and Curricular Changes." Liberal Education. v. 54, #4 (December, 1968) pp. 546-559.

Smith, Kirker. "Pass-Fail: Were Your Hypotheses Valid?" College and University v. 46 (Summer, 1971) pp. 549-555.

Stake, Robert E. "Objective, Priorities, and Other Judgment Data." Review of Educational Research. v. 40, #2 (April, 1970) pp. 181-212.

Stallings, William M. & Smock H. Richard. "The Pass-Fail Grading Option at a University: a Five Semester Evaluation." Journal of Educational Measurement. v. 8, #3 (Fall, 1971) pp. 153-160.

Stallings, William M.; Wolff, Joseph L. & Maehr, Martin L., "Test Anxiety, Grade Utilities, and the Pass-Fail Option." Research Report #289, Measurement & research Division, Office of Institutional Resources, University of Illinois, Champaign, Illinois, February, 1969 (Mimeograph).

Stallings, William M. & Yongkittikul, Chumporn. "Selected

BIBLIOGRAPHY

Data for the Evaluation of Pass-Fail Grading at the University of Illinois During the Fall Semester of 1972." Research Memorandum #125. Champaign-Urbana: Measurement & Research Division, Office of Instructional Resources, University of Illinois, April, 1972.

Starch, Daniel & Elliott, Edward C., "Reliability of the Grading of High School Work in English." School Review. v. 20, #7 (1912) pp. 442-457.

Starch, Daniel & Elliott, Edward C., "Reliability of Grading Work in Mathematics." School Review. v. 21, #4 (1913) pp. 254-259.

Starch, Daniel & Elliott, Edward C., "Reliability of Grading Work in History. School Review. v. 21, #10 (1913) pp. 676-681.

Strong, Foster. Letter to Ohmer Milton in , March 7, 1967, as reported by Milton the University of Tennessee's Teaching-Learning Issues.

Suslow, Sidney. "Pass-Fail Grading at Berkeley: Facts and Opinions." Berkeley, California: The Office of Institutional Research, The University of California (Berkeley), February, 1973 (Mimeograph).

Toaf, Howard M., Jr., "What Price Grades?" Journal of Higher Education. v. 35 (February, 1964) pp. 100-103.

Thayer, Robert E., "Do Low Grades Cause College Students to Give Up?" 1971 (Mimeograph)

Thompson, W. N., "A Study of the Grading Practices of Thirty-One Instructors in Freshman English." Journal of Educational Research. v. 49, #1 (1955) pp. 65-68.

Torbert, William R. & Hackman, J. Richard. "Taking the Fun Out of Outfoxing the System." in Runkel, Philip; Harrison, Roger; & Runkel, Margaret (eds) The Changing College Classroom. San Fransico, California: Jossey-Bass, 1969.

Tragesser, F. F. et al. "Are New Developments in Achievement Testing an Adequate Answer to P/F Systems?" College and University. v. 43 (Summer, 1968) pp. 565-567.

Trow, William C., "Grades and Objectives in Higher Education." Educational Record. v. 49, #1 (Winter, 1968) pp. 85-91.

Von Wittich, Barbara. "The Impact of the Pass-Fail System

BIBLIOGRAPHY

upon Achievement of College Students." Journal of Higher Education. v. 43, #6 (June, 1972) pp. 499-508.

Ward, Lewis, "Staple Exchange Rates for Academic Currencies." College Board Review #38 (1959) pp. 26-29.

Warren, Jonathan R. College Grading Practices: An Overview. Washington, D. C.: Department of Health, Education and Welfare, March, 1971. also ERIC report # ED 047 164.

Warren, Jonathan R. Current Grading Practices. Washington, D. C.: American Association for Higher Education, January 15, 1971. Research Report Number 3.

Webb, Sam C., "Measured Changes in College Grading Standards." College Board Review. #39 (1959) pp. 27-30.

Wilson, Kenneth M., "Increased Selctivity and Institutional Grading Standards." College and University. v. 46 (Fall, 1970) pp. 46-53.

Wrinkle, W., Improving Marking and Reporting Practices. New York: Rinehart & Co., 1963.

Wolfe, Dael, "Are Grades Necessary?" Science. v. 161, #3847 (September 1968) page 1203.