

AN ABSTRACT OF THE THESIS OF

JOHN MICHAELS PIKE for the DOCTOR OF EDUCATION
(Name) (Degree)

in EDUCATION presented on October 21, 1971
(Major) (Date)

Title: A COMPARISON OF THE EFFECTS BETWEEN LETTER-
GRADES AND PASS-FAIL GRADES ON THE
ATTITUDES AND ACHIEVEMENT OF ELEVENTH GRADE
UNITED STATES HISTORY STUDENTS

Abstract approved:

Redacted for privacy

Margaret Catherine Lumpkin

This study was designed to determine whether individual students at the eleventh grade level who were graded on a pass-fail basis in United States History would show any difference in academic achievement from those of a similar group marked by the traditional ABCD-F letter-grade method. In addition, the study was to investigate the effect, if any, of the pass-fail marking system on the study habits, study attitudes, and the over-all study orientation of the same students as compared with those being evaluated by the letter-grade system. Specific hypotheses to be examined were:

1. There is no difference between the experimental and control groups in the academic achievement of the United States History students as measured by the Crary American

History Test.

2. There is no difference between the experimental and control groups in academic behavior (study habits) of the United States History students as measured by the Survey of Study Habits and Attitudes.
3. There is no difference between the experimental and control groups in the study attitudes of the United States History students as measured by the Survey of Study Habits and Attitudes.
4. There is no difference between the experimental and control groups in over-all study orientation of the United States History students as measured by the Survey of Study Habits and Attitudes.

A review of the literature reveals considerable dissatisfaction with the traditional ABCD-F letter-grade marking system widely used in the secondary schools of the United States. The pass-fail concept appears to have merit as a more flexible system, and its use on the secondary level is increasing.

The sample consisted of six eleventh grade required United States History classes at Rex Putnam High School, Milwaukie, Oregon. One hundred and two individuals participated in the experimental groups. They received a pass-fail mark on all routine class work as well as grade card reports throughout the academic year. Data were

obtained for 96 experimental subjects. One hundred and five individuals participated in the control group. They were exposed to the traditional ABCD-F letter-grade marking system throughout the academic year. Complete data were obtained for 99 control subjects. Statistical treatment was applied to the data for 195 subjects.

The Crary American History Test and the Survey of Study Habits and Attitudes were administered to subjects prior to the beginning of the research and at the conclusion of the project. An analysis of the post-test score gains revealed no differences between the means of the experimental and control groups as tested by t at the .05 level of significance on all four measures; hence, all four hypotheses were supported in that subjects exposed to the pass-fail system showed no difference in academic achievement, study habits, study attitudes, and study orientation from those treated by the traditional ABCD-F letter-grade marking system.

Recommendations for further research in secondary marking systems include:

1. The merits and demerits of innovative grading practices as compared with traditional marking systems.
2. Investigation of the achievement and study attitudes of students taking all courses under the pass-fail system compared with students taking all courses under the letter-grade system.
3. Investigation of the pass-fail and letter-grade marking system

on attendance patterns.

4. Studies concerning the effect, if any, of the pass-fail grading system on high achievers as compared with students of low ability.
5. Experimentation in grading systems doing away with the use of the designation "F" denoting failure.
6. Development of criteria for the evaluation of student performance using the pass-fail system.
7. Investigation of the relationship between grading practices and cheating.
8. Relationships, if any, between grading practices and drop out patterns.
9. Study of problems relating to college requirements and how these impinge upon wider use of the pass-fail marking system.

© 1972

John Michaels Pike

ALL RIGHTS RESERVED

A Comparison of the Effects Between Letter-Grades and
Pass-Fail Grades on the Attitudes and Achievement of
Eleventh Grade United States History Students

by

John Michaels Pike

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Education

June 1972

APPROVED:

Redacted for privacy

Professor of Education

in charge of major

Redacted for privacy

Dean of the School of Education

Redacted for privacy

Dean of Graduate School

Date thesis is presented October 21, 1971

Typed by Clover Redfern for John Michaels Pike

ACKNOWLEDGMENT

I would like to express my gratitude and sincere appreciation to Dr. Margaret Catherine Lumpkin, Professor of Education, School of Education, Oregon State University for her guidance and direction in the development of this research study.

Acknowledgment is also made to Dean Dan William Poling, Associate Dean of Students and to Professors Carl Leonard Anderson, Andrew William Kennedy, James J. O'Connor, and Bud M. Tamler of Oregon State University.

I also want to express special appreciation for the support, understanding, and encouragement given to me by my wife, Emma, and by my children Wes, Dave, and Sharon, during the stress of study and thesis preparation.

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
I. INTRODUCTION	1
Statement of the Problem	5
Hypotheses	6
Limitations of the Study	6
II. REVIEW OF RELATED LITERATURE	10
Traditional Marking Systems	10
Need for Change in Grading Practices	16
Pass-Fail Concept	24
Summary	31
III. METHODOLOGY OF THE STUDY	33
Collection of the Data	33
Sampling Technique	33
Design of the Study	34
Procedure	34
Measuring Instruments	37
The Cray American History Test	37
The Survey of Study Habits and Attitudes	40
IV. RESULTS OF THE STUDY	43
Pre-test Analysis of Means and Variances	44
Post-test Score Gain Analysis	47
V. SUMMARY AND DISCUSSION	53
Recommendations for Further Research	57
BIBLIOGRAPHY	60
APPENDICES	63
Appendix I: Correspondence	63
Appendix II: Statistics	64

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Method of randomization.	33
2. Summary of means and variance pre-tests.	47
3. Raw score totals.	56

A COMPARISON OF THE EFFECTS BETWEEN LETTER-GRADES
AND PASS-FAIL GRADES ON THE ATTITUDES AND
ACHIEVEMENT OF ELEVENTH GRADE UNITED
STATES HISTORY STUDENTS

I. INTRODUCTION

Concerned people have been questioning the value of the traditional letter-grade marking system used by most secondary teachers in the evaluation of their students and the practices involved in the determination of the rating process. Answers to these questions have often been based on personal opinion rather than established fact. Most responses have been subjective in nature and do not give satisfaction to those involved. Adding to the controversy is the growing practice of greater utilization of the pass-fail concept in grade marking as well as other experimental procedures on the secondary level.

Dissatisfaction with traditional grading systems has become widespread. Many students and educators feel that worry and frustration over grades is becoming so critical that the cause of learning is being hindered. As a result, the traditional letter-grade marking system on the secondary level has become the target of criticism with ever-increasing intensity. The reasons for this are complex and varied. For one thing, many teachers are skeptical of the grading system, and feel at best a student's grade in a particular course is an imperfect measure of what he knows of the course content, to say

little of what he may retain (Wofle, 1968). Others feel that the existing system of teacher grading and report card marking of secondary students is not only harmful, but "inadequate, burdensome to teachers, and a barrier between teachers and students" (Sr. Scholastic, 1967). Seawall (1967) compares the letter grading system and its lack of precision to a machinist's use of the yardstick in place of a micrometer in gauging the thickness of a piece of sheet metal. All of this has contributed to the rather chaotic state of the grading system.

Often parents are reluctant to place their children in honors classes, enrichment programs, and other special situations, because they fear lower grades at the end of the grading period. Flinker (1967) says, "There is some validity to their complaints that marks are relative, not absolute. It is not uncommon to find half a class failed by a teacher who considers himself conscientious." Grades often result from such variables as the make-up of the class, the standards of the individual teacher, and the impressions made upon the teacher by the student, his personality as well as his effort and conduct. Melby (1966) declares the marking system to be damaging in its impact on the education of children and, "it should go the way of the hickory stick and dunce cages. It should be abandoned at all levels of education."

Lack of fairness in grading is a topic of much discussion both

among students and teachers. In the past as well as the present, however, little has been done on the secondary level to change, improve, or remedy the situation. Adding to the problem is the great variability of classroom teachers in their marking and reporting systems resulting in "discrepancies of graders in judgment and in consistency from one paper to another" (McKeel, 1951). Further, great differences are revealed between departments in grading courses of comparable levels. The differences are evident within courses in allied fields. In addition, grading standards vary greatly, with some instructors being very rigid and others being most lenient in their assessment. Berdie (1965) points out that college instructors and other teachers assign course grades on the basis of absolute achievement and relative standing. The first requires determination of what is superior, satisfactory, and poor achievement; the second is based largely on the characteristics of the student population. Because of this, significance differs within an institution from department to department, or even course to course, and from year to year.

A good example of inconsistency on a large scale is the well-known "immunity" to failure of seniors in high school. The reason is not given that seniors are more academically oriented or conscientious than the under-classmen, but because the twelfth-year status is somewhat of an automatic guarantee of teacher and/or administrative beneficence (Winkler, 1969). To be more specific, one need only

compare the greater incidence of failure in an eleventh grade United States history class with that of a twelfth grade American Problems class. Obviously, the reason is not because of subject matter, lack of effort, or other academic variables on the part of the student. One of the consequences from the above is that many secondary students, being aware of the imperfect nature of the grading system, tend to exploit these imperfections in diverse ways.

All of this is not to suggest that grading should be abandoned on the secondary level. Winkler (1969) emphasizes "standards must be maintained and compliance demanded of each student according to his ability to perform." The problem, however, often becomes more acute because for some educators a "chasm often exists between flexible and rigid standards, a chasm that is too contradictory to bridge" (Winkler, 1969). This may imply a basic conflict in the grading process that need not exist. The ultimate goal of the grading process should be an indication of the educational standing of the student. If the grade is to represent achieved "worth" it must reflect all circumstances leading to this value judgement, for the grading process, like so many other aspects of education, is a personal involvement between teacher and student (Winkler, 1969). Seawall (1967) concludes "the logic of the present letter grading system should be questioned." He further points out that examinations are the usual basis for a generalized letter grade, and the generalization from this

point into a letter grade is redundant and the logic questioned when one derives a specific grade-point average from a generalized letter grade.

Statement of the Problem

The problem is to determine whether individual students at the eleventh-grade level who are graded on a pass-fail basis in United States History will show greater, the same, or less academic achievement than those of a similar group marked by the traditional ABCD-F letter-grade method. Is academic achievement facilitated, restricted, or unaffected by the simplified pass-fail marking process? This study is designed to consider the merit of the pass-fail concept in grading versus the traditional letter-grade system.

In addition, the problem is to determine the effect, if any, of the pass-fail marking system on the study habits, study attitudes, and the over-all study orientation of the same students as compared with those being evaluated by the letter-grade system.

The dependent variables are the study orientation which is an over-all measure of the study habits and attitudes and, the academic achievement. Extraneous variables such as maturation, sex, and time are assumed to exert insignificantly different influences on both the experimental and the control groups, and are, therefore, not being considered in the design of this study.

Hypotheses

- H₁ There is no difference between the experimental and control groups in the academic achievement of the United States History students as measured by the Crary American History Test.
- H₂ There is no difference between the experimental and control groups in academic behavior (study habits) of the United States History students as measured by the Survey of Study Habits and Attitudes.
- H₃ There is no difference between the experimental and control groups in the study attitudes of the United States History students as measured by the Survey of Study Habits and Attitudes.
- H₄ There is no difference between the experimental and control groups in the overall study orientation of the United States History students as measured by the Survey of Study Habits and Attitudes.

Limitations of the Study

The city of Milwaukie is located within a suburban community six miles south of Portland, in Clackamas County, Oregon, on the east bank of the Willamette River. As a part of the greater Portland

Metropolitan area, the city has a population of 10,470 according to the 1970 census, and serves a larger community of approximately 60,000. From a socio-economic point of view, the community consists largely of residential, white middle class people. There are exceedingly few minority group students in Rex Putnam High School such as Blacks, Chicanos, Chinese, Japanese, Indian American, or people of Jewish ancestry. The same is true for the other schools of Union High School District Number Five. In addition, there is little representation of these groups on the teaching staff at Rex Putnam High School, or, for that matter, in the other high schools of the district. For the most part, the community may be thought of as white, Anglo-Saxon, and Protestant, with the exception of numerous Italian Catholic families in the Milwaukie area stemming from an original colony of these people engaged in truck farming during the early beginning of the current century.

At the present time, there are three high schools in Union High School District Number Five, and sites have been selected for a fourth and a fifth high school. Eighteen elementary schools and five junior high schools provide educational facilities for children who eventually enter one of these high schools. On March 1, 1971, voters approved the unification of Union High School District Number Five with the four elementary school districts, thus forming Oregon's fifth largest district. Each high school was designed for a maximum enrollment of

1200 students. At the oldest school, Milwaukie High School, the enrollment is currently 1374. The second oldest building, Clackamas High School, has 1655 students, and the newest of these, Rex Putnam High School, built in 1963, numbers 1382 as of December 31, 1970.

The curriculum of Union High School District Number Five is a general comprehensive school program which offers college preparatory courses and pre-vocational education. A regular summer session is conducted at Rex Putnam during June and July for the high school students of the district who may take make-up courses in required subjects, enrichment classes, regular courses in an accelerated program for early graduation, and driver education. The summer enrollment in 1970 was 889 students.

In addition to the three comprehensive high schools, Union High School District Number Five has gained recognition for its emphasis on vocational education and its establishment of the Owen O. Sabin occupational skills center. The approach is to provide a center for part-time attendance for occupational training purposes only. The skills center philosophy is based on "strong feeling that the student should retain his or her identity with the home high school" (Crain and Martin, 1968). Thus, all activities for students are carried on in the home school and provision is made so that interested students may participate in these objectives.

The skills center program concentrates on job clusters and

specific general education is retained in the home high school. Only junior and senior students from the three high schools in the district attend the skills center program. The juniors (Level One) attend classes for a two-hour period and the seniors (Level Two) for a three-hour period. Students from Rex Putnam High School are transported morning and afternoon to participate in the program of the skills center.

Because this study was conducted with the students from one high school, it is possible that generalizations will not apply to other districts that differ in size, structure and composition. However, assumptions might be made that the findings could have significance for the other two high schools in Union High School District Number Five in that the curriculum, teaching methods, size, structure and composition of the schools are similar to those in Rex Putnam where the research was undertaken.

Nevertheless, it should be kept in mind that even though the findings of this study can be applied with reasonable certainty only to Rex Putnam High School, the results of the study certainly suffice as an excellent reason to further explore the merit of innovative grading systems as compared with traditional practices. It should also emphasize the necessity for teachers to develop a sharp awareness of the need to create an effective and fair appraisal system for evaluating student achievement in the classroom.

II. REVIEW OF RELATED LITERATURE

Traditional Marking Systems

One of the most difficult and perplexing problems facing teachers and administrators throughout the history of secondary education has been that of rating and evaluating the educational progress and achievement of children. Flinker (1967) states that in evaluation, as in other fields, "a little learning is a dangerous thing." If such be the case, then the history of educational marking systems on the secondary level has been fraught with error and danger. DePue (1967) labels this the "great fault" in school marks. He relates that a sharp displacement occurred in the early years of the present century in the most common and effective system of notation for educational measurement. Subsequently, or concurrently, and unfortunately, an additional disturbance further compounded that displacement. More recent attempts at a satisfactory system of notation have merely glossed over and hidden the original fault, so that consequent errors in measuring academic achievement continue to plague with fallacious results those who, for various reasons, would attempt to evaluate and predict academic performance (DePue, 1967). DePue's conclusion is that by a simple strategy of four practical elements, classroom teachers everywhere right now can achieve more unified marking, approximating a common goal and justice in marking. These are: (1) using an accurate scale

of zero to 100; (2) equalizing for guessing; (3) aiming for an arithmetic mean of 50; (4) counting as failures only the stragglers. In lieu of the conventional letter grading system of assessment, Seawall (1967) also suggests that each student be ranked in an array from highest to lowest based on performance. After the students are ranked in the array, the assignment of a percentile score to each student becomes simply an administrative function. The class rank can be converted to a base of 100 with the arithmetic mean for each class to be 50. This ranking would develop a spirit of competition within the class and, consequently, enable the students to develop more critical judgment of their work with the resulting improvement in quality.

In response to what is to take the place of the present "unjust and outmoded" grading system, Johnson (1967) suggests the use of a project plan conducted under the teacher's supervision, the successful completion of which will be adequate proof that he has learned the fundamentals of that particular course and is capable of further study in it. Johnson (1967) says that the project plan of motivated effort by a student is superior to the use of arbitrary grades in determining his grasp of each area of the subject matter for the following reasons: (1) it encourages creative thinking and individual production; (2) it indicates subjective interest in the material studied; (3) it generates a desire to help himself and encourages self pride in his work; (4) it creates motivation within the student which is the vital need.

According to Wrinkle (1947) there are six fallacies in grading. They are: (1) the conventional letter grade is an effective conveyor of information; (2) a student can achieve any mark he wishes--if he is willing to make the effort; (3) the student's success in his after-school life compares favorably with his grade mark success in school; (4) the student's mark is comparable to the worker's pay check; (5) the competitive system provides a worth-while and justifiable introduction to competitive adult life; (6) the mark can be used as a means without its being recognized as an end in itself. Aikens (1968) suggests three false assumptions upon which most of our grading systems are based: (1) there is a consistent value or level of achievement implied in the letters, percentages, or numerals used in grades; (2) grades are objective; (3) grades motivate pupils to better achievement. Mouley (1968) says grading is not without danger, and when misused, can easily negate all the school is trying to accomplish through sound teaching and evaluation. From a psychological point of view, he lists the following objections raised against most of the grading done in the schools: (1) grades are inadequate indicators of the child's total development and a single letter grade or even two cannot possibly cover all aspects of child growth in any one subject-matter area; (2) grades exert a strong influence in raising the child's extrinsic motivational level which may be desirable for some students, but for others, have detrimental effects ranging from psychosomatic

disorders to feelings of resentment, hostility, frustration, and discouragement; (3) grades often become ends in themselves so that learning is cast into a secondary role of vehicle toward a grade; (4) teachers often find that giving the child a low grade destroys the pupil-teacher relationship they have been trying to cultivate; (5) grades are of questionable dependability from the standpoint of validity and reliability.

One of the problems in the traditional grading system is the need for a wide sampling of oral and written responses necessary in order to make a fair assessment of the student's standing. However, some teachers make report card ratings based solely on a single test or on a mid-quarter or final test of their own devising. Standard practice in making assignments is a frequency distribution following the pattern of a bell-shaped curve. Teachers use this pattern to justify the wide distribution of marks in their own classes. "Forgotten are the two basic conditions for achieving the bell-shaped curve, viz., the very large number of children tested (over 1000) and the random heterogeneity of these children" (Flinker, 1967). The above practice does not take into consideration, for example, such special situations as honor classes of superior students or special high schools with unusually high standards of admission.

The practice of using the bell curve in determining classroom grade distributions is examined in greater detail by Fensch (1966).

He says:

Grades are often the results of a mysterious and weird process. Each teacher has his own philosophy and system, so that the basis for arriving at grades differs widely from teacher to teacher and from school to school.

The practice of applying a curve to each class in determining grade distributions is almost universal in its use. One determines by the grade curve how many students will receive A's, B's, C's, D's and F's (Fensch, 1966). In addition, some teachers use the bell curve to show that approximately five to seven percent ought to get A's and F's and 20 to 24 percent should receive B's and D's, which leaves approximately 38 to 50 percent who are to receive C's.

The fact that bell-shaped curves are statistical devices based on a large sampling and were never intended to be used with such small numbers of cases as 20, 25, or even 30 or 35 individuals does not seem to be known by the average teacher (Fensch, 1966).

The same probably is true for school administrators who often set the pace in policy determining the grading practices of the members of the staff.

The obvious conclusion drawn from the above is that without a representative sampling, the teacher's attempt to establish on the basis of the grade curve how many students should receive A's, B's, D's or F's, might be questionable. It does happen that teachers are assigned classes where the students have ability meriting considerably more than average grades. In these cases most of the students

should receive A's, B's, and possibly some C's. This is particularly true of an honors class. Yet, in many cases on the secondary level, teachers still assign marks on the basis of the curve so that there are top grades, bottom ones, and the majority in between. It appears that a great number of teachers use the curve system because they feel it gives support to back up their grade distributions. Fensch (1966) concludes:

It is possible that these teachers would be seriously frightened to find themselves awarding no grade below a B in one class. In fact, it is a safe guess that many of them could not bring themselves to do this because it would mean lowering their standards.

In dealing with the question as to the significance of the traditional letter-grade system, Sparks (1970) says,

It would be unpleasant to discover in plain fact, as many already suspect in a subjective sort of way, that our present system of academic bookkeeping has very little to do with the total capacities of an individual and often fails to represent what did actually happen to him in a given learning situation.

Some imply that higher grades are rewards for conformity to the established order and, hence, do not always justify the reliance society places on them as a basis for selecting future leaders, innovators, and problem-solvers (Sparks, 1970). In other words, the transcript and the grade point average have become "negotiable instruments in their own right, serving as a medium of exchange in the academic community. However, one can never be

quite certain just which values they actually represent" (Sparks, 1970). The conclusion reached by Sparks (1970) about the traditional letter-grade marking system is best stated in the form of a question: "Have we come to the place and a time in which the negotiability of academic symbols is more important than knowing (or caring) what the symbols mean?"

Need for Change in Grading Practices

As most educators are aware, there is considerable unrest in the high schools of America. Many educators are apprehensive that it will reach the proportion of the revolt on college and university campuses. Racial antagonism is blamed for much of the unrest, but other issues including the grading system are contributing causes.

In view of this, many questions are being asked about education and much criticism is being directed at the public schools. Oregon's Superintendent of Public Instruction, Dale Parnell (1969) gives a critical analysis of the situation as it exists today:

The American public is being deluged with fuzzy thinking on what the schools should and should not be doing, but few critics are coming up with workable solutions that will push the universal education experiment up the road. Although Americans have called for universal education, schools have given them the bell-shaped curve which says that unless you receive an "A" or a "B" in algebra, geometry, physics, chemistry, etc., you're a terminal student, not advanced, and you really don't amount to

much. Grades are the hard currency, critical thinking a major goal, yet grades don't generally seem to reflect even remote attainment of this goal. The result of this is that school life has become separated from real life, and students are restless and disenchanting. If the secondary schools are to be truly comprehensive necessary changes must be made including the move away from the caste system of college prep with its traditional emphasis upon grades and grade point averages.

The 1970 Second Annual Survey (Gallup, 1970) of the public's attitude toward the public schools shows that people continue to have a high regard for the schools of their community and that they believe firmly that education is the road to success in America. Yet, the poll shows that there is undeniably a new mood in the nation with which education must reckon. Student protests, both at the high school and college level, have, in the case of the curricula, opened the whole issue of whether education in America is reality-related. This new mood also involved the question of grading practices. Goodlad (1964) identifies the current curriculum reform movement with the launching of the first Russian satellite in the fall of 1957. Little reform, however, in the marking systems has been undertaken on the secondary level.

A paramount reason for change in traditional grading practices is to be found in the realm of learning theory, particularly in view of the important relationship between intrinsic motivation and the learning process. Bruner (1967, p. 113) says, "The most characteristic thing about human beings is that they learn." Interest in the material

to be learned should be stressed, rather than such external goals as grades or later competitive advantage. In fact, "where grades are used as a substitute for the reward of understanding, it may well be that learning will cease as soon as grades are no longer given--at graduation" (Bruner 1963, p. 51). This is illustrated by the remark of Kirkendall (1939) when he says, "The avidity with which pupils work for marks is equalled only by the assiduity with which teachers teach for marks." Bruner (1963, p. 66) further questions the present system of rewards and punishments as seen by pupils in school as tending to inhibit the use of intuitive thinking:

The assignment of grades in school typically emphasizes the acquisition of factual knowledge, primarily because that is what is most easily evaluated; moreover, it tends to emphasize the correct answer, since it is the correct answer on the straightforward examination that can be graded as "correct". It appears to us important that some research be undertaken to learn what would happen to the development of intuitive thinking if different bases for grading were employed.

Another variable contributing to the questionable value of the traditional secondary marking system, and hence the need for change, is the extent of personality factors and the "halo" effect contributing to the higher grades of the "good" students. In an analysis of the reasons for grade failures in an economics course, Rosenthal (1967) made observations with reference to certain characteristics predominating in a progressive manner from the low students to the higher students. These may be summarized as follows: (1) the more

academically inclined students were collectively dressed better, revealing a higher socio-economic standing; (2) the higher the achievement level, the more open were the students demonstrating a greater freedom in asking for assistance from the teacher; (3) the higher the grade standing of the student, the more pronounced the attitude of concern about parental displeasure if the student did not achieve.

More is said about grades and halos by Musselman (1967). What is the relationship, if any, which exists in the minds of teachers between the academic success and personality traits in high school was sought by studying samplings from 187 high school graduates in four graduating classes. From these graduates of a large city high school were selected the highest ten percent, the middle ten percent, and the lowest ten percent according to grades received in four years of high school (Musselman, 1967). It was found from the study that those who get the highest grades were, in the opinion of their teachers, better appearing, more socially adept, more courteous, more honest, more dependable, more industrious, and more cooperative than their fellow graduates who made lower grades. Numerous questions were raised by the study as to the effect of halo and grades. Do pupils early and unconsciously begin to function in personality and character spheres at a level set by their academic success in school? Do teachers, without knowing exactly when or how they do it, create a

"class society" based upon early competence in school work? Is the public school responsible in part for a lower level of personality development for many because the school has "typed" children? Has the school caused the child to evaluate his personal worth and his pride on an academic normal curve (Musselman, 1967)? Obviously these questions may not readily be answered, but they are of great concern and cannot easily be evaded or dismissed by the educational establishment.

In view of this, Melby (1966) concludes that the present marking system is no longer relevant to the needs and educational programs of our society. Stated reasons are that it says nothing about a pupil that is meaningful. Further, it glosses over exceptional effort on the part of some pupils and lack of effort on the part of others. It says nothing about the most important outcomes of education. Finally, it tells us little about what the pupil has accomplished in the subject he studies and nothing about the effect his study of the subject has had upon him.

Melby (1966) therefore indicts the marking system for not only being irrelevant and mischievous, but as being destructive to the self-concept of millions of children every year. Particularly is this true of the deprived child:

He often enters school at six with few of the pre-school experiences that the middle-class children bring to school. We ask him to learn to read. He is not ready to read. We give him a low mark--we repeat the low mark for each marking period--often for as long as the child

remains in school. At the end of perhaps the ninth grade, the child drops out of school. What has he learned? He has learned he cannot learn. We have told him so several dozen times. Why should he think otherwise?

We have lied to him. He can learn. If we were worth our salt as teachers and as a school, we should have taught him he can learn. We should have asked him to do things he can do, not what we know he can't do. Every day we should have sent him home with more confidence in himself, liking himself better than when he came in.

Commenting on the need to do away with F's in the conventional grading system, Nelson (1968) describes the deleterious effect of grade failure:

A failing grade probably does more harm than good. Success breeds success, while failure leads to failure. Each time that a child is told that he is a failure, he loses a little more self-confidence. The failing grade is not, as we sometimes try to use it, a prod for increased effort. Instead it is generally one more blow to an ego that has already been eroded by multiple failure. Rather than resulting in greater effort, an F will most likely lead the child to reject the academic area as a means of obtaining satisfaction and turn to other less desirable ways of achieving notice. The boy or girl with the failing marks is also quite often the discipline problem and eventually the dropout.

Melby (1966) points out that in addition to the effect of the marking system on the deprived child, all children suffer to a certain extent from the marking system. Some of the reasons for this are that they are injured because they are induced to seek the wrong goals. They become satisfied when their performance reached a given level, rather than when they have done their best. They are injured because they develop a lifelong dislike for subjects in which they get low

grades, which is an obstacle to self-actualization. Johnson (1967) points out that some of the slowest poor-grade students make the most success in their careers, while too often the top high graders make little use of their talents, often being dangerous and destructive in the use of the abilities they have.

The injustice of grades is outlined by Johnson (1967):

(1) Students are not encouraged to think for themselves. (2) Students find various means of cheating to make the grade. (3) Fear of grade failure dominates their attitudes because too many teachers use grades as incentive for a student's work. (4) A love of learning for personal satisfaction is seldom enjoyed. (5) When students work for grades, knowledge becomes a secondary goal. (6) Grades become a student's chief goal in his school career and it isn't his fault.

Aikens (1968) says, "From a psychological standpoint, the emphasis given to grades by parents, teachers, and honorary groups is frequently damaging to the mental health of the pupil." Tensions are created and feelings induced by low grades often reduce further achievement rather than stimulate the pupil to do better work. We condemn the child of low potential to a life of failure by penalizing him for his lack of ability with failing grades. Furthermore, according to Aikens (1968) the problem of cheating in the schools is directly proportional to the emphasis placed on grades. The more emphasis on grades, the greater the cheating problem.

After teaching English in the high school to what she refers to as the "bottom" group, Sparks (1967) came to the conclusion that the traditional grading system is not just "weak," but positively disastrous and has a "shattering effect on the unmotivated, the socially unadjusted, the slow learner, the mentally undeveloped, the confused." She proposes an alternative to the traditional grading system dominated by teachers, a more flexible student-oriented one and has tried to de-emphasize the importance of grades.

Frost and Rowland (1969, p. 65) discuss the plight of the under-achiever in the classroom:

The usual American classroom tends to be completely oriented toward achievement. Successful achievement is rewarded, while a lack of achievement is punished in one form or another. The achiever in the classroom makes the highest grades, is supported by the teacher in independent study efforts, enjoys extraordinary classroom privileges, and is often recognized publicly for his ability to achieve by memberships in honor societies or special groups, or receives rewards of high value such as scholarships, plaques or scrolls.

The non-achiever is a problem in any classroom, for he is most easily conceived as a threat to the teacher's competence. If a child cannot be taught, or will not be taught, he is challenging the teacher's ability to teach. The teacher will scheme, organize, plan, and devise a panorama of activities, each designed to "motivate" or "guide" the underachiever or non-achiever to become an achiever. There is no acceptable place for the non-achiever in the usual classroom. He is a thorn and threat who must be dealt with

One of the most questionable practices is the teacher's use of the marking system for disciplinary purposes or for classroom

control, especially when all other methods fail. Students are forced to conform by the threat of poor marks. The point here is illustrated by the young mother in the doctor's office warning her restless child, "If you don't behave, I'll have the doctor stick that long needle in you!" By the same token, the teacher often uses the threat of an "F" with possible long-lasting and disastrous results.

Pass-Fail Concept

It has been pointed out in this study that grades and grading systems have become the targets of criticism for a number of reasons. On the college level grades have symbolized, "objective examinations, machine records, and other efficient devices and practices that critics have sometimes indicated as evidence of the dehumanization and regimentation of higher education" (Wolfe, 1968). College students further object to having grades released to draft boards, being graded in courses outside the major field of interest in which the student lacks background, and receiving grades for required courses considered irrelevant by the students.

These objections are also being expressed by students on the high school level. With the changes in draft regulations and lowering of the draftable age, high school students feel even more threatened today. Many high school students would like to take courses outside their major field of interest, but because of the threat of lower grade

point averages due to the uncertainty of how well they would do in the course, high school students often evade the opportunity for greater educational growth. In addition, high school students feel that many courses are irrelevant and lacking intrinsic motivation their grade point will be lowered. Many feel they would like a greater choice in electives and a reduction of courses required for graduation.

The provision for students to take some courses without credit or without grades is spreading. Wolfle (1968) states,

The practice is not one to be made universal and mandatory, for grades provide some students with standards for self-appraisal, and for motivation, and they provide the institution with comparative information concerning students.

However, educators are finding that educational effectiveness and the ability to make effective decisions in the pursuit of educational excellence do not necessarily require that grades be given to all students in all courses. Even on the secondary level experimentation should be encouraged to give students a wider latitude in determining such things as choice of courses, length of time for graduation, and as to whether or not in some cases, the student should receive a grade in a specific course (Wolfle, 1968).

Ebel, (1967) pursuing a similar line of thought, pinpoints the problem by posing several significant questions. Should colleges (or secondary schools) institute pass-fail grading for certain courses in order to encourage the students to broaden the range of their

educational experience? It has often been said that even good students are reluctant to take courses outside of their major field, courses in which they would have to compete for grades with students who were majoring in that field. Now, Ebel (1967) asks, is it in fact true that a significant proportion of good students who want to take a course outside their major, and who feel that they are prepared to study it profitably, are deterred by the prospect of a lower grade? Do they choose instead to take a less valuable course that may give them a higher grade?

Stallings, Smock, and Elwood (1968) refers to the pass-fail grading policy as a "palliative for some of the alleged evils of the traditional marking systems and as an inducement to students wishing to explore the terrae incognitae of academic disciplines not directly related to major fields of study." Although there are many variants, a pass-fail option on the college level usually implies that a student may have a choice in the selection of a small number of courses for which his grade will be either pass or fail; and that, upon receiving a pass, he obtains credit toward meeting graduation requirements (Stallings, Smock and Elwood, 1968).

Basic questions raised in the implementation of a pass-fail grading system on the high school level are similar in nature to those on the college level, but with some exceptions. These questions are discussed by Stallings, Smock, and Elwood (1968) in some detail.

Should the instructors be notified as to which students are enrolled on a pass-fail basis? Should the pass-fail option be restricted to certain groups? Should only "good" students be allowed to take courses on a pass-fail basis? Should a "fail" be considered in the computation of the grade point average? How is the pass-fail system to be articulated with grade point average? What effects will pass-fail grades have upon the continuation of a scholarship or athletic eligibility? Should the pass-fail option be extended to required courses? Should the student be able to take some work in his minor field for pass-fail? Suppose a student discovers a new interest by taking a pass-fail course and suppose he changes his major to conform to that new interest. Should the credit previously earned on a pass-fail count toward the requirements of the new major? Should the student be allowed to change from pass-fail to regular grading in a given course? If so, what time interval for making this change should be allowed (Stallings, Smock, and Elwood, 1968)? One can readily see that the impact of the pass-fail option upon a particular student is dependent, in large part, upon the answers to many of the above questions.

Haskell (1967) reports on an experimental pass-fail course instituted for a semester at Longmeadow High School in Massachusetts. Some of the questions resulting from that study are: (1) What system of controls should be used? (2) What are the criteria for grading student performance? (3) How will the pass-fail concept be

received by college admissions officials? In this course Haskell (1967) makes an attempt to help the student free himself from the standard of an "arbitrarily fixed grade, and, to substitute a kind of self-motivation that enables an attainment level that he might not reach under the pressure imposed by a letter grade!"

Other questions about the merits and demerits of the pass-fail system posed by Stallings, Smock, and Elwood (1968) are: Do students in the pass-fail system achieve at the same level as students who are on the regular grading system? Do students in a pass-fail situation show more favorable attitudes toward the course than on the regular grading system? Do students working in a pass-fail situation attend scheduled class periods as often as those on the regular grading system? Do students working in a pass-fail situation spend as much time in preparation as in the regular situation? Do students in a pass-fail structure audit more courses than those in the regular grading system? What courses are commonly elected on a pass-fail option? Does the preparation of pass-fail to regular students vary over time? Do students who elect the pass-fail option differ from others in grade point average, scholastic aptitude scores, or non-intellective characteristics? Does the supposed broadening of students' interest under a pass-fail system lead to change in academic majors (Stallings, Smock, and Elwood, 1968)?

The question might be asked, if grades are so time consuming

and unsatisfactory, why even bother with them? Emmens, Gage, and Rummel (1965, p. 288) cite eleven specific purposes of marking systems, four of which are pertinent to this study: (1) grades provide a basis for information and recommendations for parents on student status and progress; (2) grades may be used as indices for promotion and guidance; (3) grades give information to a school which the student may later attend; (4) grades may reinforce learning through extrinsic motivation. Mouley (1960, p. 395) says that higher grades can be thought of as a general reinforcer with great potential for influencing a student's behavior; but, also, "fear of failure can be thought of as a stimulus which contributes to a higher level of anxiety over testing and grades, resulting in--for some students--a situation too stressful for optimum learning." Mouley (1960, p. 105) describes the responsibility of the schools to strive in helping the student maintain his self-image intact:

The school plays a significant role in determining the quality of the child's self-concept, especially with respect to academic materials and other contents with which he has had only limited previous experience. . . . It must provide each and every child with the opportunity for productive achievement and maximum utilization of his potentialities. The school can also do harm: indiscriminate grading and reporting to parents followed by condemnation at home may be a significant factor in the development of a negative self-concept.

Mouley (1960, p. 105) also discussed other ramifications of the grading process, much of which is detrimental to the learning

environment:

Grades are so important to some students that accumulating a good scholastic record or surpassing others becomes the measure of academic and personal worth. Others restrict their learning to the minimum necessary to obtain a passing grade. The precaution teachers take to protect the security of examinations, for example, is a reflection of the objectionably extrinsic nature of grades in the overall learning process. Under these conditions, grades actually impede learning by forcing the student to study not what is important but what is likely to be covered in the test. Under these circumstances, grades simply serve to teach the child shortcuts such as cramming, cheating, apple-polishing, and other means of circumventing true learning. Furthermore, although over-emphasis on incentives can promote some learning, this learning tends to cease the minute the incentive is removed. The emphasis must remain on the real goal, for that is what will continue to be significant after the reward is removed.

Cronbach (1963) states that fear of low grades stimulates anxious students to a level of arousal such that learning effectiveness is lessened. However, non-anxious students are stimulated to a less complacent arousal level with a resulting increase in learning effectiveness. Furthermore, Cronbach (1963, p. 595) says, "anxiety tends to facilitate performance in a task where carefulness and reliance on authority pays off. . . . it tends to interfere with original thought, which always implies a departure from the safe, well-marked path."

Some fear that the possibility of not being motivated by grades, through the adoption of a pass-fail grading system, would lower the learning effectiveness of non-anxious students. On the other hand, non-anxious students may be motivated to a high degree of learning

without the external threat stimulus of fear of low grades. One might predict that students' behavior would more apt to be changed in some desirable way if the students were motivated intrinsically to learn course content rather than extrinsically motivated to learn course content because of a desire for earning high grades (Stallings, Smock, and Elwood, 1968).

One of the big, but not insurmountable difficulties, in wider use of the pass-fail system on the secondary level is the attitude of the colleges and universities which for the most part prove to be somewhat less than enthusiastic. In the consideration of using teacher evaluations rather than grades for its students, Wilson High School in Portland, Oregon, polled colleges and universities to see what reception its graduates might expect (Scholastic Teacher, 1970). There were 149 replies. Sixty percent of the public institutions responding said they would be apprehensive about a high school student from a pass-fail system. Private colleges split nearly 50-50. The only institutions to show little concern over whether or not students had grades to show were community colleges; 16 out of 18 said Wilson's students would be welcome and the other two were uncertain.

Summary

A review of the literature reveals widespread and growing dissatisfaction with conventional marking practices and grading

systems. Many feel that the marking system currently used is damaging in its impact on the education of children and should be changed. Grades often result from a great many variables and differ from student to student, class to class, course to course, teacher to teacher, and school to school. In particular, the marking system may be especially harmful to the deprived child, and hence, destructive to the self-concept of many children. Furthermore, the fear of failure induced by low grades may have a negative effect upon the mental health of the pupil. The emphasis upon grades is said to encourage cheating and stifle independent thinking, causing knowledge to become a secondary goal. Many feel that the present marking system is no longer relevant to the needs and educational programs of our society. A review of the literature reveals little positive support for the conventional letter-grade marking systems as currently in use. Yet, few would abandon the grading system altogether, but rather recognize the need for innovation and change.

The pass-fail concept as a method of grading is becoming more widespread on the secondary level. There is a need for experimentation and carefully controlled research in order to evaluate the merits and demerits of this system as it reflects course achievement and student learning.

III. METHODOLOGY OF THE STUDY

Collection of the Data

Sampling Technique

The sample consisted of six eleventh grade required United States History classes at Rex Putnam High School, Milwaukie, Oregon. These six classes were selected from a larger population of approximately 425 students, all required to take classes in United States History. Each student in the total population of 425 was assigned a number from a table of random digits and a sample of 210 students was drawn from the original population. The sample was likewise then divided into two groups of 105 students each. A coin was then tossed to determine which group was to be the control group and which was to be the experimental group (Figure 1).

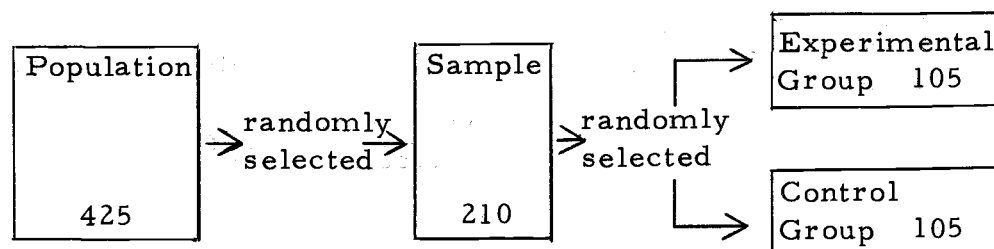


Figure 1. Method of randomization.

Design of the Study

The sample contained members of both sexes, approximately equally distributed and homogeneous as to academic ability. Both experimental and control groups were divided into three classes numbering about 35 individuals and were similar then as to sex, age, academic ability and size.

The design of the research included the assignment of the sample classes to a team of two United States history teachers working and teaching together so that all classes, both experimental and control, were exposed to the same instruction as well as curriculum materials. The teaching team consisted of Mr. Glen Jackson, chairman of the history department, and Mr. Larry Hermo, instructor in United States history. Lectures, films, television programs, and other aspects of instruction were the same for both control and experimental classes. Textbooks, handout materials, unit tests and other curriculum materials were also the same for all classes.

Procedure

Before the assignment to the experimental classes, a letter was mailed directly to the home of each student explaining the nature of the program to the parents (Appendix I). It was explained that the credit received from the pass-fail class would in no way affect college

entrance requirements, honor role, or the student's grade point average. In fact, the pass-fail credit was to be excluded from the student's grade point average. Parents were encouraged to call the administration for additional information about the project. Enclosed in the letter was a form to be returned by the parents signifying permission for their child to participate in the pass-fail classes. In the case of those parents not returning the form, it was assumed there was no objection to the program. One set of parents objected strongly to the project and their child was transferred to a regular class. A number of other parents called for further information expressing interest in the proposed plan. Several parents expressed wholehearted approval of the pass-fail grading system and expressed various ideas as to why they favored changes in grading procedures.

At the beginning of the school year a brief explanation was made by the teaching team to the sample United States history classes as to the pass-fail method of grading and standard procedures to be followed in the traditional letter-grade system. Two weeks later the entire sample was given the pre-test of both the Crary American History Test Revised Form E and the Survey of Study Habits and Attitudes Form H administered by the guidance department. The Crary American History Test took approximately 50 minutes to administer. The time used for the Survey of Study Habits and Attitudes was approximately 40 minutes. Of the 210 students enrolled in

the six United States History classes, three, because of prolonged absences at the outset, did not complete the pre-testing and consequently were not included in the pre-test analysis. These three were in the experimental group, and later proved to be drop outs.

Throughout the year both experimental and control groups were exposed to similar instruction and curriculum materials. Class sessions were often held in the large team-teaching room so that at times large groups of both control and experimental classes were given the same lectures covering the same material. The same was true during the showing of films and various television programs utilizing the closed circuit television facilities of the district. At times both control and experimental classes were broken down by the teaching team into small groups for individual research projects, discussion and other instructional techniques. Procedures for instructional methods were uniform at all times for both the control and the experimental groups.

In all grading including routine assignments, tests, as well as the nine week and semester report card marking, the treatment of the experimental classes consisted of a pass or fail mark. The control classes received the traditional letter-grade mark.

Instructional sessions as described were continued throughout the academic year. The roll was taken each day and recorded in the grade book as well as the attendance office in accordance with

procedures at Rex Putnam High School, with an average absence of 4.1 per period. No consistent individual absence pattern among students or classes was noted.

After the treatment both experimental and control groups were post-tested with the Crary American History Test Revised Form E and the Survey of Study Habits and Attitudes Form H. Pre-testing was completed for 102 individuals in the experimental group and 105 individuals in the control group. Six dropped out from the experimental group and six from the control group throughout the year. Post-testing was completed for 99 in the control group and for a total of 96 in the experimental group. The control group consisted of 48 males and 51 females and the experimental group consisted of 46 males and 50 females.

Measuring Instruments

The Crary American History Test

The Crary American History Test (Crary, 1965) was first published in 1950-52 and later revised in 1964-65. The test measures student achievement of important educational objectives of the American history course typically taught in grade eleven. It samples not only command of factual knowledge but skills, interpretation, understanding, and inference as well (Crary, 1965). A class period of 45

to 50 minutes is normally required for administering the test.

Each form, of 76 items, gives balanced coverage to the politico-national development of the United States, social and economic problems, cultural and scientific aspects of American history, foreign relations, and study skills that are important in the social studies (Crary, 1965). The two forms are approximately equivalent to each other in terms of content and statistical balance.

The item-tryout and standardization of the Crary American History Test were carried out in a carefully supervised dual research program conducted in 1964. Experimental forms were administered to 9,523 American history students in 24 public high schools. Community population ranged from 10,000 to 100,000 and were proportionately distributed through 16 states in representative sections of the country. In order to obtain comparative items analysis and standardization data for the two forms of the test, the experimental forms were administered in alternating order. Also, the experimental forms were subjected to careful review by many classroom teachers and their criticisms were taken into consideration in the final forms (Crary, 1965). From the 130 items in each experimental form, 76 were selected for each of the revised Forms E and F. The correlations of the long experimental forms with the 76-item final forms were found to approximate .98 for each of the standardization groups. The magnitude of this correlation is evidence that in shortening and

refining the test there was no loss of efficiency between the 76-item forms and the experimental forms (Crary, 1965). Comparison of the raw score distribution for the two Forms E and F revealed that the scores on one form were equivalent to the scores on the other form at all points along the distribution curves. Consequently, performance data from one form of the test are directly comparable with performance data from the other form (Crary, 1965).

The Crary American History Test has been constructed to provide a representative coverage of the important outcomes of American history instruction as attested by authoritative judgments and by a consensus of current practices. Thus, in relation to commonly accepted instructional emphasis, the test may be considered as having a high degree of validity (Crary, 1965).

A unique feature of the revised test is the addition of a guessing distracter to each item. Students are instructed to mark (DK) for "Don't Know" whenever they are unable to decide on the correct answer. The conclusion from this is that the contamination of students' scores by wild guessing is less than it would be if the DK options were not used (Crary, 1965).

Reliability coefficients for the Crary American History Test have been determined by the split-half method and are further complemented by the standard errors of measurement. Thus, the Crary American History Test should provide a valid and reliable instrument for

evaluating achievement, guidance in determining educational goals, comparing ability and achievement, and as a research tool for investigations of the educational or counseling process.

The Survey of Study Habits and Attitudes

Construction of the Survey of Study Habits (Brown and Holtzman, 1967) began with an exhaustive review of the literature and a series of group discussions with college freshmen concerning the motivational differences between good and poor students and culminated in the development of Form C for college students in 1959. Form H was developed in 1959 to meet the demands for a version of the SSHA that would be suitable for use in junior and senior high schools (Brown and Holtzman, 1967).

The SSHA has proven to be useful as a screening instrument, a diagnostic help, a teaching aid, and as a research tool. Because of its appreciable relationship to academic success, it is suitable for inclusions with other scales in research investigations of the educational or counseling process. The total score may represent the Study Orientation variable in the research design, or specific subscale scores may serve as checks or controls on the study habits and/or attitudes of the groups being studied (Brown and Holtzman, 1967). The great majority of students complete the SSHA within 20-35 minutes, depending on the grade level. There is, however, no

time limit imposed and students are instructed to complete the inventory.

There is a total of 100 items on the questionnaire which has four basic scales, two subscales, and a total score. The four basic scores--Delay Avoidance (DA), Work Methods (WM), Teacher Approval (TA), and Education Acceptance (EA)--are obtained directly. The Study Habits (SH) score is found by adding the DA and the WM scores; the Study Attitudes (SA) is found by adding the TA and EA scores; the Study Orientation (SO) is obtained either by adding the SH and SA scores, or by summing the DA, WM, TA, and EA scores. For the purpose of this study only the scores from the Study Habits (SH), the Study Attitudes (SA), and total score representing the Study Orientation (SO) will be used.

Form H (the high school edition) of the SSHA has been validated in a large number of junior and senior high schools throughout the United States. In 1965, the SSHA--Form H was administered to representative samples of junior and senior high school students in six states. Scores on appropriate scholastic aptitude tests, as well as course grades, were obtained from cooperating counselors or administrators in each school system. The correlations were statistically significant and positive for all schools and grade levels. The mean correlations range from .46 for Grade 12 to .55 for Grade seven (Brown and Holtzman, 1967).

Two hundred thirty-seven ninth-graders in San Marcos High School were given the SSHA--Form H twice, with an interval of four weeks between sessions. The test-retest reliability coefficients were .95, .93, .93, and .94, respectively, for the Delay Avoidance, Work Methods, Teacher Approval, and Education Acceptance scales, and .95 for SSHA total score (SO). These studies indicate that the four subscale scores are sufficiently stable through time to justify their use in predicting future behavior or in assessing the degree of change in study habits and attitudes after counseling (Brown and Holtzman, 1967).

IV. RESULTS OF THE STUDY

The study was designed to investigate whether or not students at the eleventh-grade level who are graded on a pass-fail basis in United States History, would show greater, the same, or less academic achievement than those of a similar group marked by the traditional ABCD-F letter-grade method. In addition, the study was to determine the effect, if any, of the pass-fail marking system on the study habits, study attitudes, and the over-all study orientation of the same students as compared with those being evaluated by the letter-grade system.

From the original sample of 207 students taking the pre-test, complete results were obtained from 195 subjects, 94 males and 101 females. The control group was composed of 48 males and 50 females. Only those students who completed all the designated testing for both the Crary American History Test and the Survey of Study Habits and Attitudes were included in the statistical treatment.

Specific hypotheses tested were:

1. There is no difference between the experimental and control groups in academic United States History achievement.
2. There is no difference between the experimental and control groups in academic study habits.

3. There is no difference between the experimental and control groups in academic study attitudes.
4. There is no difference between the experimental and control groups in overall academic study orientation.

The statistical treatment used was (1) a pre-test to determine that the control and experimental groups were not significantly different; and, (2) tests of difference between the means and variances for the two groups on the four measures.

Pre-test Analysis of Means and Variances

In the pre-test analysis it was found that there was no difference between the means of the experimental and control groups as tested by t at the .05 level of significance on all four measures using the formula

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right)}}$$

where

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

for the means and the formula

$$F = \frac{S_1^2}{S_2^2}$$

for the variances (Appendix II).

There was no difference between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance on the Crary American History Test.

$$t = \frac{(24.05 - 24.49)}{\sqrt{72.9198(1/105 + 1/102)}} \qquad F = \frac{75.303}{70.466}$$

$$= -0.3706 \qquad F = 1.06864$$

There was no difference between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance on the Survey of Study Habits.

$$t = \frac{(40.59 - 38.64)}{\sqrt{218.913(1/105 + 1/102)}} \qquad F = \frac{217.175}{220.702}$$

$$= .9480 \qquad F = 1.01624$$

There was no difference between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance on the Survey of Study Attitudes.

$$t = \frac{(53.06 - 49.97)}{\sqrt{179.532(1/105 + 1/102)}} \qquad F = \frac{200.53}{157.911}$$

$$= 1.6588 \qquad F = 1.26989$$

There was no difference between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance on the Survey of Study Orientation.

$$t = \frac{(93.06 - 88.61)}{\sqrt{660.886(1/105 + 1/102)}} \qquad F = \frac{698.114}{622.553}$$

$$= 1.4101 \qquad F = 1.12137$$

Hence conclude from above that:

- (1) There is no difference between the means of the experimental and control groups as tested by t at the .05 level of significance on all four measures.
- (2) There is no difference between the variances of the experimental and control groups as tested by F at the .05 level

of significance on all four measures.

A summary of the means and variance pre-tests of the four measures may be seen in the following schedule (Figure 2).

	<u>t</u>	<u>F</u>
Crary American History Test	-0.3706	1.06864
Survey of Study Habits	.9480	1.01624
Survey of Study Attitudes	1.6588	1.26989
Survey of Study Orientation	1.4101	1.12137

Figure 2. Summary of means and variance pre-tests.

Thus, the randomness of the sample was established in the pre-test since no difference was found between the means of the experimental and control groups as tested by t at the .05 level of significance on all four measures, and no difference was found between the variances of the experimental and control groups as tested by F at the .05 level of significance on all four measures.

Post-test Score Gain Analysis

An analysis of the score gain between experimental and control groups was made for the Crary American History Test, in order to test hypothesis number one (Appendix II).

H_1 : There is no difference between the experimental and control groups in academic United States history achievement.

H_0 : There is no difference between the means of the experimental and control groups on the Crary American History Test,

as tested by t at the .05 level of significance.

The control scores when compared with the experimental scores yielded $t = -0.0959$ not significant at the .05 level. The control scores when compared with the experimental scores yielded $F = 1.05$ not significant at the .05 level.

$$t = \frac{(32.24 - 32.38)}{\sqrt{103.871(1/99 + 1/96)}} \qquad F = \frac{106.487}{101.172}$$

$$= -0.0959 \qquad F = 1.05$$

Thus, no difference was found between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance, and the first hypothesis was supported. Subjects exposed to the pass-fail marking system showed no difference in academic achievement than those treated by the traditional ABCD-F letter-grade marking system.

An analysis of the score gain between experimental and control groups was made for the survey of study habits, in order to test hypothesis number two (Appendix II).

H_2 : There is no difference between the experimental and control groups in academic study habits.

H_0 : There is no difference between the means of the experimental and control groups on the Survey of Study Habits, as tested by t at the .05 level of significance.

The control scores when compared with the experimental scores yielded $t = 1.2737$ not significant at the .05 level. The control scores when compared with the experimental scores yielded $F = 1.12$ not significant at the .05 level.

$$t = \frac{(40.44 - 37.81)}{\sqrt{207.802(1/99 + 1/96)}} \qquad F = \frac{219.398}{195.840}$$

$$= 1.2737 \qquad F = 1.12$$

Thus, no difference was found between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance and the second hypothesis was supported. Subjects exposed to the pass-fail marking system showed no difference in study habits than those treated by the traditional ABCD-F letter-grade marking system.

An analysis of the score gain between experimental and control groups was made for the survey of study attitudes, in order to test hypothesis number three (Appendix II).

H_3 : There is no difference between the experimental and control groups in academic study attitudes.

H_0 : There is no difference between the means of the experimental and control groups on the Survey of Study Attitudes, as tested by t at the .05 level of significance.

The control scores when compared with the experimental scores yielded $t = 0.7620$ not significant at the .05 level. The control scores when compared with the experimental scores yielded $F = 1.16$ not significant at the .05 level.

$$t = \frac{(48.43 - 46.84)}{\sqrt{212.223(1/99 + 1/96)}} \qquad F = \frac{227.902}{196.049}$$

$$= .7620 \qquad F = 1.16$$

Thus, no difference was found between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance, and the third hypothesis was supported. Subjects exposed to the pass-fail marking system showed no difference in study attitudes than those treated by the traditional ABCD-F letter-grade marking system.

An analysis of the score gain between experimental and control

groups was made for the survey of study orientation, in order to test hypothesis number four (Appendix II).

H_4 : There is no difference between the experimental and control groups in overall academic study orientation.

H_0 : There is no difference between the means of the experimental and control groups on the Survey of Study Orientation, as tested by t at the .05 level of significance.

The control scores when compared with the experimental scores yielded $t = 0.8176$ not significant at the .05 level. The control scores when compared with the experimental scores yielded $F = 1.031$ not significant at the .05 level.

$$t = \frac{(88.88 - 85.66)}{\sqrt{755.892(1/99 + 1/96)}} \qquad F = \frac{743.178}{769.007}$$

$$= .8176 \qquad F = 1.031$$

Thus, no difference was found between the means and variances of the experimental and control groups as tested by t and F at the .05 level of significance and the fourth hypothesis was supported. Subjects exposed to the pass-fail marking system showed no difference in over-all study orientation than those treated by the traditional

ABCD-F letter-grade marking system.

Hence conclude from above that: (1) There is no difference between the means of the experimental and control groups as tested by t at the .05 level of significance on all four measures. (2) There is no difference between the variances of the experimental and control groups as tested by F at the .05 level of significance on all four measures. To summarize: all four hypotheses were supported in that subjects exposed to the pass-fail marking system showed no difference in academic achievement, study habits, study attitudes, and study orientation from those treated by the traditional ABCD-F letter-grade marking system.

V. SUMMARY AND DISCUSSION

The key question in this research study is what effect, if any, does a pass-fail grading system have upon academic achievement and how does it influence classroom learning? Or, do students in a pass-fail system achieve at the same level as students who are on the regular grading system? In relation to the above question, the findings of this study would seem to indicate in a positive way that eleventh grade United States history student at Rex Putnam High School are capable of equal academic achievement in a pass-fail grading system as those students in United States history receiving the traditional letter-grade mark. In coming to this conclusion one must consider the limitations of this particular study. However, the ramifications of these findings may take on larger proportions. In the case of this study, the pass-fail standards were related directly to the traditional letter-grade system in that all students involved, both in the experimental and control classes were placed on the same curve in arriving at the decision to pass or fail or to receive the regular letter grade. In all the tests and examinations students were to achieve at least 50 percent of the correct answers on a scale of 100 points. Those achieving less than 50 percent in the experimental pass-fail class on tests and examinations received a "fail" mark and those achieving above the 50 percent level received a "pass" mark. In the case of the control classes, those

achieving less than 50 percent on the scale received the "F" grade and those achieving above the 50 percent level received a traditional letter grade following the general practice of a grade curve allowing for a distribution of approximately five to seven percent of A's, 20 to 24 percent receiving B's and D's, leaving approximately 38 to 50 percent receiving C's.

It would seem that a greater use of the pass-fail grading system would allow more flexibility in the entire grading process, which in turn, would enable to secondary student and teacher to be more concerned with achievement and less concerned with problems related to receiving graduated letter grades.

In addition to this, the more simplified pass-fail system on the secondary level might facilitate teaching methods that would enhance a greater stress on intrinsic motivation in the learning environment rather than extensive use of a letter grade as an extrinsic device used as a motivational factor. Children could be better induced to seek higher goals in learning. Rather than being satisfied with performance at a given level, students with academic ability might be challenged to performance of a higher level when that performance is not tied to a system of grade reward or punishment. Students of low potential perform better in an atmosphere of encouragement and support, rather than one dominated by fear of failure. Reduction of threat among these students might also reduce problems of a disciplinary nature.

Greater use of the pass-fail grading system might tend to alleviate much of the unfairness inherent in the conventional system, one consequence of which is the alienation that often takes place between student and teacher as a result of disappointment in receiving a lower letter grade than was anticipated. Sometimes the lower grade is the result of a point or two on an inflexible grade curve. Greater use of the pass-fail system would also de-emphasize the halo effect as an important variable in the conventional grading system that has long been a troublesome element to both teachers and students. In all of this, it would appear that greater use of the pass-fail grading system would help to reduce some of the conflict and tension that exists in the process of education and thus help to foster a greater threat-free learning environment.

As to the effect of the pass-fail grading system on attendance, the study at Rex Putnam High School revealed little. Unlike on the college level, as well as at some secondary schools, class attendance at Rex Putnam High School is required, roll being taken at regular intervals. The study showed no discrepancies in attendance patterns between the experimental pass-fail group and the letter-grade control group when the number of absences from each class were compared.

In addition to the findings concerning academic achievement, this research was also concerned about study habits, attitudes, and the study orientation of the eleventh grade United States history

students at Rex Putnam High School. Again, the findings of the study in these areas was in agreement with the hypotheses postulated prior to the actual classroom learning process. The study habits, study attitudes, and overall study orientation of the United States history students at Rex Putnam High School did not appear to be affected by the pass-fail grading system as compared to the letter-grade method of assessing student progress. In contradistinction to the increase in academic achievement throughout the time period of the research; however, it was noticed that there was a decline in total score gain for the same period in study habits, study attitudes, and study orientation for both experimental and control groups involved in the study. This may be seen by the following description: (Figure 3).

	<u>Beginning of Year</u> (Total Score)	<u>End of Year</u> (Total Score)
Experimental Group:		
Crary American History	2498	3108 (up)
Study Habits	3941	3630 (down)
Study Attitudes	5097	4497 (down)
Study Orientation	9038	8223 (down)
Control Group:		
Crary American History	2525	3192 (up)
Study Habits	4262	4004 (down)
Study Attitudes	5571	4795 (down)
Study Orientation	9833	8799 (down)

Figure 3. Raw score totals.

In each of the three measures, for both control and experimental groups, it may be seen from the above schedule that there was a decline in total score from the beginning to the end of the year. In American history achievement there was an increase in total score from the beginning to the end of the year for both control and experimental groups. This may suggest a gradual breakdown in study habits, study attitudes, and overall study orientation as the student progresses throughout the academic year. From these observations one can only conjecture as to reasons for this negative trend in study habits, study attitudes, and study orientation as reflected by the decrease in total score as measured by the Survey of Study Habits and Attitudes. The decline in scores may indicate physical and mental "tiredness" as the year progresses or increased negativism towards teachers and the learning process in general. These observations may have added implications for districts considering a 12 month secondary school year.

Recommendations for Further Research

Even though the findings of this study can be applied with reasonable certainty only to Rex Putnam High School, and with possible significance for the other high schools in Union High School District Number Five, the results of the study suffice as a justifiable reason to further explore the credibility of the merit of the pass-fail grading

system as compared with traditional practices on the secondary level. In consequence, the following recommendations for further research and investigation are made:

1. More controlled research is needed on the merits and demerits of innovative grading practices as compared with traditional systems.
2. Investigation of the achievement and study attitudes of students taking all courses under the pass-fail system compared with students taking all courses under the letter-grade system should be made.
3. Studies concerning the effect, if any, of the pass-fail grading system on high achievers as compared with students of low potential **is** needed.
4. More extensive investigation of the pass-fail and letter-grade marking practices on the attendance patterns of secondary students in a controlled study should be made.
5. Greater experimentation in grading systems moving away from the concept of "F" as a term denoting failure needs to be undertaken.
6. Further study into the problems of college requirements for secondary students concerning grade point averages, honor roll, scholarships and entrance requirements, and how these impinge upon wider use of the pass-fail system on the

secondary level is imperative.

7. Research into the development of controls and criteria for the evaluation of student performance under the pass-fail method of grading on the secondary level should be made.
8. Study is needed to explore the relationships between grading practices and cheating.
9. Research is needed concerning grading practices and problems of discipline and the connection between these variables and drop out patterns.
10. Research concerning articulation between the public and school personnel involving pass-fail, as well as other innovative grading practices, would be helpful in fostering needed change in secondary marking practices.

BIBLIOGRAPHY

- Aikens, Harold H. It's time to change our grading system. *Ohio Schools* 46:104. 1968.
- Berdie, Ralph H. A solution to the problem of distributing course grades. *School and Society* 93:373-75. 1965.
- Brown, William F. and Wayne H. Hotlzman. Survey of study habits and attitudes, manual. New York, The Psychological Corporation, 1967. 30 p.
- Bruner, Jerome S. Toward a theory of instruction. Cambridge, Harvard University Press, 1967. 176 p.
- _____ The process of education. New York, Random House, 1963. 92 p.
- Cronbach, Lee J. Educational psychology. New York, Harcourt, Brace, and World, 1963. 658 p.
- Crain, Robert and Clyde Martin. An occupational skills center--why--where--what--how. Scope and Sequence, Union High School District Number Five, Milwaukie, Oregon 8:1-5. 1968.
- Crary, Ryland W. Crary American history test manual. New York, Harcourt, Brace and World, 1965. 19 p.
- DePue, Palmer. The great fault in school marks. *Journal of Secondary Education* 42:217-22. 1967.
- Ebel, Robert L. Some limitations of basic research in education. *Phi Delta Kappan* 49:81-3. 1967.
- Emmers, H.H., N.L. Gage, and J. Francis Rummel. A practical introduction to measurement and evaluation. New York, Harper, 1965. 390 p.
- Fensch, Edwin A. Tone down the bell curve. *Clearing House* 40:469-71. 1966.
- Flinker, Irving. How do we rate? *Clearing House* 42:245-7. 1967.

- Frost, Joe L. and G. Thomas Rowland. Curricula for the seventies. Boston, Houghton Mifflin, 1969. 454 p.
- Gallup, George. Second annual survey of the public's attitude toward the public schools. Phi Delta Kappan 52:98-112. 1970.
- Goodlad, John I. School curriculum reform. New York, The Fund for the Advancement of Education, 1964. 95 p.
- Haskell, John M. Pass-fail? A system worth trying. Clearing House 42:172-3. 1967.
- Johnson, Icie F. The injustice of grades. School and Community 54:24-5. 1967.
- Kirkendall, Lester A. Teaching for marks. School and Society 49:642-44. 1939.
- McKeel, Willis Newton. Analysis of variability of grades and students. Master's thesis. Corvallis, Oregon State University, 1951. 51 numb. leaves. (Microfilm)
- Melby, Ernest O. It's time for schools to abolish the marking system. Nations Schools 77:104. 1966.
- Mouley, George J. Psychology for effective teaching. New York, Holt, Rinehart, and Winston, 1968. 639 p.
- Musselemann, Dayton L. Grades and halos. Clearing House 42:26-8. 1967.
- Nelson, Gerald E. Let's do away for F's. Minnesota Journal of Education 48:25-6. 1968.
- Parnell, Dale. The Oregon way. American Vocational Journal 44:14-17. 1969.
- Rosenthal, Geraldine H. Where did the system fail? Clearing House 41:558-61. 1967.
- Scholastic Teacher, November, 1967, p. 4.
- _____, October, 1970, p. 7.

Kirschbaum's book

Sparks, Merla. An alternative to the traditional grading system.
English Journal 56:1032-4. 1967.

Sparks, Richard K. How much is a beesworth? Clearing House
44:544-46. 1970.

Stallings, William M., H. Richard Smock, and Elwood K. Leslie.
The pass-fail grading option. School and Society 96:179-80.
1968.

Winkler, James E. Grading. Peabody Journal of Education 46:350-1.
1969.

Wofle, Dael. Are grades necessary? Science 161:1203. 1968.

Wrinkle, William L. Improving marking practices in elementary and
secondary schools. Rhinehart and Company, New York, 1947.
120 p.

APPENDICES

APPENDIX I

Correspondence

REX PUTNAM HIGH SCHOOL

Milwaukie, Oregon

Dear Mr. and Mrs.

In many of our high schools there has been intensive interest in a more simplified and effective method of grading and report card marking. One of these methods is the increasing use of the pass-fail concept in evaluating the student's progress. This method is being used very extensively in our colleges and its use is increasing on the secondary level.

This year, we would like to organize and grade several of our United States History classes on this basis. Your son or daughter has been recommended to participate in one of these classes. From these special classes, we hope to gather information that will enable us to evaluate our current grading practices and thus to improve the effectiveness of teachers and students involved together in the learning process. We are enlisting your help in making this special program possible. Above all, we want to assure you that the credit received from this class on a pass-fail basis will in no way affect college entrance requirements, honor roll, or the student's grade point average. The grade point average of students participating in the program will be determined from courses taken by the student exclusive of United States History.

We are thus requesting your permission to allow us to include your child in the program by returning the permission form that is enclosed. If you have questions or are in need of further information about this program, I will be most happy to discuss it with you. Please feel free to call me at this phone number, 659-3330, ext. 66, between the hours of 7:30 a.m. and 5:30 p.m. With your continued help and cooperation, we can seek better ways of improving the quality of our educational program.

Sincerely yours,

Leonard Suchsland, Principal
Rex Putnam High School

LS/eh

APPENDIX II

Statistics

Pre-test statistics to determine means, variances, and standard deviations of the two groups to establish the randomness of the sample on all four measures.

Crary American History Test

Experimental:

$$N = 102, \quad \sum_{i=1}^k f_i x_i = 2498, \quad \sum_{i=1}^k f_i x_i^2 = 68362, \quad \bar{X} = 24.49,$$

$$S'^2 = 70.466, \quad S.D. = 8.393$$

Control:

$$N = 105, \quad \sum_{i=1}^k f_i x_i = 2525, \quad \sum_{i=1}^k f_i x_i^2 = 68627, \quad \bar{X} = 24.05,$$

$$S'^2 = 75.303, \quad S.D. = 8.677$$

Survey of Study Habits

Experimental:

$$N = 102, \quad \sum_{i=1}^k f_i x_i = 3941, \quad \sum_{i=1}^k f_i x_i^2 = 174781, \quad \bar{X} = 38.64,$$

$$S'^2 = 220.702, \quad S.D. = 14.856$$

Control:

$$N = 105, \quad \sum_{i=1}^k f_i x_i = 4262, \quad \sum_{i=1}^k f_i x_i^2 = 195800, \quad \bar{X} = 40.59,$$

$$S'^2 = 217.175, \quad S.D. = 14.74$$

Survey of Study Attitudes

Experimental:

$$N = 102, \quad \sum_{i=1}^k f_i x_i = 5097, \quad \sum_{i=1}^k f_i x_i^2 = 270807, \quad \bar{X} = 49.97,$$

$$S'^2 = 157.911, \quad S.D. = 12.566$$

Control:

$$N = 105, \quad \sum_{i=1}^k f_i x_i = 5571, \quad \sum_{i=1}^k f_i x_i^2 = 316637, \quad \bar{X} = 53.06,$$

$$s'^2 = 200.53, \quad S.D. = 14.161$$

Survey of Study Orientation

Experimental:

$$N = 102, \quad \sum_{i=1}^k f_i x_i = 9038, \quad \sum_{i=1}^k f_i x_i^2 = 864338, \quad \bar{X} = 88.61$$

$$S'^2 = 622.553, \quad S.D. = 24.951$$

Control:

$$N = 105, \quad \sum_{i=1}^k f_i x_i = 9833, \quad \sum_{i=1}^k f_i x_i^2 = 994138, \quad \bar{X} = 93.65,$$

$$S'^2 = 698.114, \quad S.D. = 26.422$$

Post-test score gain statistics to find means, variances, and standard deviations used in determining the difference between the means and variances for the two groups on the four measures.

Crary American History Test

Experimental:

$$N = 96, \quad \sum_{i=1}^k f_i x_i = 3108, \quad \sum_{i=1}^k f_i x_i^2 = 110334, \quad \bar{X} = 32.38,$$

$$S'^2 = 101.172, \quad S.D. = 10.058$$

Control:

$$N = 99, \quad \sum_{i=1}^k f_i x_i = 3192, \quad \sum_{i=1}^k f_i x_i^2 = 113460, \quad \bar{X} = 32.24,$$

$$S'^2 = 106.487, \quad S.D. = 10.319$$

Survey of Study Habits

Experimental:

$$N = 96, \quad \sum_{i=1}^k f_i x_i = 3630, \quad \sum_{i=1}^k f_i x_i^2 = 156060, \quad \bar{X} = 37.81,$$

$$S'^2 = 195.840, \quad S.D. = 13.994$$

Control:

$$N = 99, \quad \sum_{i=1}^k f_i x_i = 4004, \quad \sum_{i=1}^k f_i x_i^2 = 183660, \quad \bar{X} = 40.44,$$

$$S'^2 = 219.398, \quad S.D. = 14.812$$

Survey of Study Attitudes

Experimental:

$$N = 96, \quad \sum_{i=1}^k f_i x_i = 4497, \quad \sum_{i=1}^k f_i x_i^2 = 229477, \quad \bar{X} = 46.84,$$

$$S'^2 = 196.049, \quad S.D. = 14.002$$

Control:

$$N = 99, \quad \sum_{i=1}^k f_i x_i = 4795, \quad \sum_{i=1}^k f_i x_i^2 = 254805, \quad \bar{X} = 48.43,$$

$$S'^2 = 227.902, \quad \text{S.D.} = 15.096$$

Survey of Study Orientation

Experimental:

$$N = 96, \quad \sum_{i=1}^k f_i x_i = 8223, \quad \sum_{i=1}^k f_i x_i^2 = 778176, \quad \bar{X} = 85.66,$$

$$S'^2 = 769.007, \quad \text{S.D.} = 27.731$$

Control:

$$N = 99, \quad \sum_{i=1}^k f_i x_i = 8799, \quad \sum_{i=1}^k f_i x_i^2 = 855619, \quad \bar{X} = 88.88,$$

$$S'^2 = 743.178, \quad \text{S.D.} = 27.261$$