KINDSVATTER, Richard Hughes, 1930—
THE DYNAMICS OF CHANGE IN MARKING SYSTEMS
IN SELECTED INNOVATIVE AND NON-INNOVATIVE
HIGH SCHOOLS OF OHIO.

The Ohio State University, Ph.D., 1966
Education, general

University Microfilms, Inc., Ann Arbor, Michigan
THE DYNAMICS OF CHANGE IN MARKING SYSTEMS IN SELECTED INNOVATIVE AND NON-INNOVATIVE HIGH SCHOOLS OF OHIO

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Richard Hughes Kindsvatter, B. S., M. A.

The Ohio State University
1966

Approved by

Hugh Laughrin

Adviser
Department of Education
ACKNOWLEDGMENTS

It is abundantly clear to me that this study, modest though it is, could not have been completed without the help of several gracious, willing, and capable people. These include Professor Hugh D. Laughlin, major adviser; Professors Frank J. Zidonis and Elsie J. Albery, reading committee; Professor W. Frederick Staub, temporary major adviser; Jack Howell, Ernest Spaights, and Ralph Pruitt, graduate students; and Lynn Lemming, typist.

Each of these people will readily find the evidence of his influence reflected in the pages that follow. To each of them, I am very grateful.
VITA

November 29, 1930  Born - Steubenville, Ohio

1956 . . . . . . . B.S., The Ohio State University, Columbus, Ohio

1956-1957 . . . . Mathematics Teacher, North Royalton High School, North Royalton, Ohio


1958 . . . . . . . M.A., The Ohio State University, Columbus, Ohio

1963-1964 . . . . Research Assistant, The Ohio State University, Columbus, Ohio

1963 . . . . . . . Assistant Instructor in Secondary Education, The Ohio State University, Columbus, Ohio

1964-1965 . . . . Mathematics Teacher, Finland Junior High School, Grove City, Ohio

1965-1966 . . . . Assistant Principal, Pleasant View High School, Grove City, Ohio

FIELDS OF STUDY

Major Field: Secondary Education

Educational Administration - Professor Roy Larmee

Higher Education - Professor Everett Kircher

History of Education - Professor Robert Sutton
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background of the Problem</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>5</td>
</tr>
<tr>
<td>Review of the Related Literature</td>
<td>8</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>25</td>
</tr>
<tr>
<td>Limitations</td>
<td>27</td>
</tr>
<tr>
<td>Plan of the Study</td>
<td>28</td>
</tr>
<tr>
<td>II. REVIEW OF PAST MARKING PRACTICES</td>
<td>30</td>
</tr>
<tr>
<td>1770-1826</td>
<td>30</td>
</tr>
<tr>
<td>1826-1900</td>
<td>35</td>
</tr>
<tr>
<td>1900-1930</td>
<td>41</td>
</tr>
<tr>
<td>1930-1965</td>
<td>49</td>
</tr>
<tr>
<td>III. DESIGN OF THE STUDY</td>
<td>53</td>
</tr>
<tr>
<td>Key to the Design</td>
<td>53</td>
</tr>
<tr>
<td>Selection of Schools Included in the Study Group</td>
<td>54</td>
</tr>
<tr>
<td>The Design and Transmittal of the Questionnaire and Cover Letter</td>
<td>56</td>
</tr>
<tr>
<td>Analyses of Responses to the Questionnaire</td>
<td>58</td>
</tr>
<tr>
<td>Data from Ohio State Department of Education Publications</td>
<td>61</td>
</tr>
<tr>
<td>Interviews with Principals</td>
<td>62</td>
</tr>
<tr>
<td>Summary</td>
<td>63</td>
</tr>
<tr>
<td>IV. PRESENTATION AND DISCUSSION OF RESPONSES TO THE QUESTIONNAIRE</td>
<td>65</td>
</tr>
<tr>
<td>Procedure for Analyzing Responses to the Questionnaire</td>
<td>65</td>
</tr>
<tr>
<td>Analysis of Data from the Questionnaire</td>
<td>66</td>
</tr>
<tr>
<td>Summary</td>
<td>100</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>V.</td>
<td>102</td>
</tr>
<tr>
<td>PRESENTATION AND DISCUSSION OF SUPPLEMENTARY DATA</td>
<td>102</td>
</tr>
<tr>
<td>Relationship of School Size and Certain Financial Measures to Change in Grading Practices</td>
<td>102</td>
</tr>
<tr>
<td>Impressions from Interviews with Principals</td>
<td>105</td>
</tr>
<tr>
<td>Information on Marking Practices Obtained from Grading Materials Submitted by the Respondents</td>
<td>108</td>
</tr>
<tr>
<td>Summary</td>
<td>116</td>
</tr>
<tr>
<td>VI.</td>
<td>117</td>
</tr>
<tr>
<td>DISCUSSION OF FINDINGS AND RECOMMENDATIONS</td>
<td>117</td>
</tr>
<tr>
<td>Statement of the Position of the Investigator</td>
<td>117</td>
</tr>
<tr>
<td>Conclusions from Review of Past Marking Practices</td>
<td>118</td>
</tr>
<tr>
<td>Conclusions from Responses to the Questionnaire</td>
<td>119</td>
</tr>
<tr>
<td>Conclusions from Discussion of Supplementary Data</td>
<td>131</td>
</tr>
<tr>
<td>General Findings</td>
<td>133</td>
</tr>
<tr>
<td>Recommendations</td>
<td>137</td>
</tr>
<tr>
<td>Concluding Statement</td>
<td>141</td>
</tr>
</tbody>
</table>
TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Explanation of the Meaning of the Symbols in the K Formula</td>
<td>59</td>
</tr>
<tr>
<td>2.</td>
<td>Number of Schools Reporting the Use of Suggested Marking Practices</td>
<td>67</td>
</tr>
<tr>
<td>3.</td>
<td>Number of Schools from Each Sub-Group Reporting the Occurrence or Non-Occurrence of Special Courses for Very Able Students</td>
<td>72</td>
</tr>
<tr>
<td>4.</td>
<td>Number of Schools from Each Sub-Group Reporting the Use of Differentiation of Grades in Special Courses for Very Able Students</td>
<td>73</td>
</tr>
<tr>
<td>5.</td>
<td>Number of Schools in Each Sub-Group Reporting the Occurrence or Non-Occurrence of Special Courses for Slow-Learners</td>
<td>74</td>
</tr>
<tr>
<td>6.</td>
<td>Number of Schools in Each Sub-Group Reporting the Use of Differentiation of Grades in Special Courses for Slow-Learners</td>
<td>75</td>
</tr>
<tr>
<td>7.</td>
<td>Number of Schools in Each Sub-Group Reporting the Occurrence or Non-Occurrence of a Suggested Philosophy of Grading</td>
<td>76</td>
</tr>
<tr>
<td>8.</td>
<td>Number of Schools from Each Sub-Group Reporting the Occurrence or Non-Occurrence of a Written Policy on Grading</td>
<td>77</td>
</tr>
<tr>
<td>9.</td>
<td>The Bases Used to Determine Students' Grades as Indicated by the Schools in Each Sub-Group</td>
<td>78</td>
</tr>
<tr>
<td>10.</td>
<td>Rank Assigned to Each of Five Possible Functions of Grades by the Schools in Each Sub-Group</td>
<td>79</td>
</tr>
<tr>
<td>11.</td>
<td>Number of Schools in Each Sub-Group Reporting the Occurrence or Non-Occurrence of a Recent Change in Marking Practices</td>
<td>81</td>
</tr>
<tr>
<td>12.</td>
<td>Persons or Groups Initiating Change in Each Sub-Group of Schools</td>
<td>83</td>
</tr>
</tbody>
</table>
13. Reasons Indicated by Schools in Each Sub-Group for Changing Grading Practices ................................................................. 85
14. Sources of Suggestions for Change in Grading Practices as Indicated by Each Sub-Group of Schools .......................... 88
15. Means of Implementation of Change in Grading Practices as Indicated by Each Sub-Group of Schools .......................... 90
16. Persons or Groups Supportive of Change in Marking Practices as Indicated by Each Sub-Group of Schools .................. 92
17. Persons or Groups Resisting Change as Indicated by Each Sub-Group of Schools .............................................................. 94
18. Changes That Principals Would Like to Effect in Their Present Marking Practices ......................................................... 95
19. Means with Which Principals Propose to Change Marking Practices ................................................................................ 97
20. Comparison of School Size and Certain Financial Measures in Specified School Groups .................................................. 103
21. Resume of Decisions Regarding Acceptance or Rejection of the Hypotheses .............................................................. 120
CHAPTER I

INTRODUCTION

The research described in this study is concerned with marking practices and grading systems. Perhaps no other facet of the educational enterprise is the source of more frustration and less progress. A description of the setting of the problem will clarify the basis for this statement and the rationale for the study.

Background of the Problem

The assigning of marks\(^1\) to students by teachers is a traditional practice in American education, performed routinely, and accepted \(a\) \(p\)\(riori\) as an apparently essential function. It rarely occurs to some educators and most laymen to consider thoughtfully what the purposes of grades are, or whether student evaluation as it is commonly carried out is justified in existing as an element of the total educative process. Grades, to the person who accepts marking practices unquestioningly, are the certain measure of student achievement—and achievement from what and to what is unspecified.

The teacher, whether concerned or not with the purposes, effectiveness, or justification of grades, is in any case the

---

\(^1\)The technical terms "mark" and "marking" and the more casual terms "grade" and "grading" are used interchangeably in this study.
dispenser of grades and generally finds the marking function a source of consternation. Assigning students' marks is usually considered a necessary evil in the educative process—necessary because there is as yet no adequate substitute for marks, and evil in the sense that the very nature of a grading system produces effects counter to the best interests of both the educators and those being educated.

Grading is essentially separate from the teaching-learning process. It occurs after the fact, and ideally is a measure of some specific academic achievement or behavioral change that has occurred as the result of a planned learning experience. But because of the desirability and prestige value associated with grades, grades themselves too frequently become the object of the students' efforts rather than learning for its own sake. Educational growth and grades, in a perverted and insidious manner, thus exchange places as ends and means from that which educators intend for them.

Grades also involve an emotional aspect, for they are neither nonchalantly given nor indifferently received. For the teacher, the necessity for making a judgment that in many cases will be unpopular—damning, even—is perhaps the most unpleasant of all his tasks. For the student, a grade may represent a number of things besides achievement in the classroom. The school mark may be associated with value- and emotion-oriented experiences involving pleasure, success, and social adjustment. Examples of these are participation in various school activities, acceptance in desirable peer groups, parental
approval and parent-granted privileges and responsibilities, and the opportunity to attend college.

Teachers, charged as they are with the function of assigning grades to the pupil, have by that token the power possibly to effect a change in the very course of the student's life. It is no small responsibility. And yet the informed teacher is aware of the variability in and unreliability of teachers' grades as demonstrated by educational researchers as long ago as the first decade of this century. Grades are shown by research to be somewhat more than the product of whimsy, but generally far removed from any scientific or objective basis.

Another criticism of grading systems is their tendency to cause unfair competition of unequals. Wrinkle contends that

The competition of unequals does not provide a fair basis for determining penalties or the granting of honors. ...The encouragement of competition by individuals of inequality...is in violation of the principle of individual differences, is unfair, does not conform to mental health practices, and is negative in many of its results. (19.,48)

Jarrett states that "The acceptance of a democratic philosophy of secondary education and the continued use of aristocratic and selective devices and practices in marking presents one of the interesting inconsistencies in our educational thinking and practice" (25.,36).

The above discussion provides support for the contention that marking practices and grading systems include several inescapable disadvantages: they are ineffective in serving many of the objectives commonly assigned them, they are misunderstood by many educators and
most laymen, they exist outside the teaching-learning process but
distort the objectives of that process, they have undesirable emotional
aspects involving students, parents, teachers and administrators,
and they are often undemocratic and unfair.

One might ask himself in light of the above discussion why
responsible educators continue to condone a practice that involves
such grave and insidious aspects. McKean partially answers this
when he states that

Although a variety of reporting practices may be
observed in the schools, most secondary schools
use the traditional report card which lists a
single letter or numerical mark for each course in
which the student is enrolled. Unjustified parental
confidence in this reporting procedure and the
weight of tradition are probably the two strongest
factors in this persistence. (9.,216)

Another source of support for the continued use of conventional
grading systems is the fact that there are advantages to their use as
well as disadvantages, for they perform with varying degrees of
success many functions for which no adequate substitute has yet
been devised. Wrinkle lists these functions as follows:

1. Administrative functions: Marks indicate whether
a student has passed or failed, whether he should
be promoted or required to repeat the grade or
course, and whether he should be graduated. They
are used in transferring a student from one school
to another and in judging candidates for admis-
sion to college. They may be used by employers
in evaluating prospective employees.

2. Guidance functions: Marks are used in guidance
and counseling in identifying areas of special
ability and inability, upon deciding upon the
advisability of enrolling the student in certain
courses and keeping him out of others, and in
determining the number of courses in which he
may be enrolled.
3. Information functions: Marks are the chief means employed by the school in giving information to students and their parents regarding the students' achievement, progress, and success or failure in his schoolwork.

4. Motivation and discipline functions: Marks are used to stimulate students to make greater effort in their learning activities. They are used for the same purpose in determining eligibility for honors of many different kinds such as participation in school activities, eligibility to play on the team, membership in selected groups, the winning of scholarships, etc. (19.,31-2)

Finally, even though high school marks are in varying degrees invalid, unreliable, and variable, Bloom and Peters report that "Most researchers have come to the conclusion that high school average grades are the best single measure from which to predict college success" (2.,9).

Statement of the Problem

There is a relatively large amount of literature relating to the general topic of marking practices and grading systems. It may be subdivided into three categories: articles and studies on the variability of grades; descriptions of a grading system or suggestions for marking practices; and surveys of marking practices to discern trends. The study proposed by the writer does not fall into any of these categories although elements of the latter two categories will be incorporated as the writer pursues the main topic. Neither will the proposed study have as its final objective the setting forth of some panacea that will alleviate the numerous difficulties of grading systems, desirable as such a study might be.
The assumption of this investigator, along with that of Wrinkle, is that "there is no one perfect marking system for all schools," and that a school involved in changing marking practices "is brought face to face with the fact that what it...(may have thought) was a problem in marking is fundamentally a curriculum problem" (19.,51). So as curriculums differ from school to school, so must marking systems differ as well.

The purpose of the study is to investigate the dynamics of change in marking practices and grading systems. Change is considered to have two dimensions: positive and negative. The positive dimension is simply evidence of change; the negative, absence of any change.

Concerning the positive dimension, the study attempts to provide some answers to the several-faceted question, "What changes in marking practices and grading systems have been made, by what means, and by whom?" The negative dimension poses the question, "What are the differences in the educational climate in those schools that provide a setting conducive to change vis-a-vis those schools in which change is inhibited?" Answers to these questions were obtained through submitting questionnaires to principals of selected innovative and non-innovative schools, and by conducting personal interviews with several principals from each type of school in order to obtain from them pertinent information, experiences, and insights.

In order to examine the dynamics of change in marking
practices and grading systems, the following null hypotheses will be tested:

1. There will be no significant difference between innovative and non-innovative schools in the use of unusual marking practices.

2. There will be no significant difference between innovative and non-innovative schools in the use of differentiation of grades in enriched, advanced placement, accelerated, or honors courses or classes.

3. There will be no significant difference between innovative and non-innovative schools in the use of differentiation of grades in slow-learner or remedial classes.

4. There will be no significant difference between innovative and non-innovative schools in the number of schools with a suggested philosophy of grading.

5. There will be no significant difference between innovative and non-innovative schools in the number of schools having a written policy on grading which stipulates that teachers shall subscribe to certain standards and/or employ certain practices.

6. There will be no significant difference between innovative and non-innovative schools in bases employed for determining students' grades.

7. There will be no significant difference between innovative and non-innovative schools in what is considered to be the primary function of course marks.
8. There will be no significant difference between innovative and non-innovative schools in the position (i.e., area of responsibility) of the person who provided the impetus for change in the marking practices or the grading system.

9. There will be no significant difference between innovative and non-innovative schools in the stated reasons for making a change in marking practices.

10. There will be no significant difference between innovative and non-innovative schools in the source of the idea for the suggested changes in marking practices.

11. There will be no significant difference between innovative and non-innovative schools in the procedures for implementing changes in marking practices.

12. There will be no significant difference between innovative and non-innovative schools in the persons who actively supported changes in marking practices or the grading system.

13. There will be no significant difference between innovative and non-innovative schools in the persons who actively opposed changes in marking practices or the grading system.

Review of the Related Literature

Marking practices—descriptions, principles, criticisms, trends

In the opinion of the investigator, the single most complete and most valuable reference bearing upon the topic of marking practices
is Wrinkle's small book, *Improving Marking and Reporting Practices in Elementary and Secondary Schools* (18.). The book resulted from ten years of experimental work with the grading system in the experimental school, College High School, Colorado State College of Education, of which Wrinkle was director. In the book, Wrinkle describes the several different plans effected in the school in an attempt to achieve a successful and promising grading system. Many failures ensued in this endeavor, but out of it Wrinkle and his faculty at the school arrived at many valuable insights which he reports. A concise summary of the most significant information included in the book would include: the grading system must be inextricably interrelated with the curriculum; the grading system must be based upon the stated and practiced objectives of the school; any single mark can measure but one single aspect of the student's educational or social progress; competition for grades which matches unequals is undemocratic and unjustified.

Also of interest in Wrinkle's book is his discussion of the six fallacies that a great many persons, professional educators and otherwise, are guilty of attributing to marking practices. The following six statements are Wrinkle's own discussion of the fallacies. The first sentence of each paragraph describes the fallacy, and the subsequent statements highlight Wrinkle's argument.

1. Anyone looking at a mark can discern with a high degree of accuracy what it means. ... Basically, the inability of marks to serve the various functions which they are supposed to serve lies in the fact that they do not represent fixed
values in terms of which they may be interpreted. ... This fallacy alone should be sufficient to eliminate the use of the single ABCDF marking system from general use.

2. A student can achieve any mark he wishes if he is willing to make the effort. ... The competitive marking system, whether we recognize it or not, assumes the equality of students. ... The conventional system as it operates in most schools is not concerned with the student's working up to the level of his ability. The unfairness of this practice is obvious to anyone who possesses even a meager understanding of individual differences.

3. The student's success in his after-school life compares favorably with his success in school. ... In the assumption that success in school and success in after-school life are closely related the fallacy is that in life the individual is compared with all other people. He is not. He is compared with those in the group with which he is associated.

4. The student's mark is comparable to the worker's pay check. ... To face activity apart from interest and a recognition of the values to be derived is very likely to be unproductive of real or lasting value. If the student cannot be stimulated to apply himself without the inducement of pay-check compensation, it is probable that the activity is either inappropriate to his needs or has been so presented that he does not recognize its appropriateness.

5. The competitive marking system provides a worthwhile and justifiable introduction to competitive adult life. ... the school by its marking practices is doing much to promote the development of antisocial attitudes and practices. A desire to win even at the expense of others cannot be countenanced as a desirable educational attitude. The competition of unequals does not provide a fair basis for determining penalties or the granting of honors.
6. School marks can be used as a means to an end without their becoming thought of as ends in themselves. ... The emphasis given to marks by most teachers in most schools tends ultimately to convince the student that the mark, rather than what it is supposed to represent, is the most important outcome of learning. (19.,35-48)

Nearly all the popular educational journals occasionally include articles concerned with the various aspects of marking practices and grading systems. Also, the textbooks on secondary education, and the textbooks dealing with student evaluation, generally include a chapter on marking practices and grading systems.

As this investigator perused the literature dealing with the broad field of student evaluation, it became increasingly apparent that there was little that dealt directly with his immediate concern, i.e., the dynamics involved in implementing improved methods of marking student achievement. However, the literature does provide information and insights into the status of grading systems generally and the problems confronting educators who are interested in improving marking practices. Without this background information, the immediate concern of this study is less well understood. Therefore, representative excerpts from recent articles and textbooks are included in this review of the literature.

For the sake of convenience in analyzing the recent literature, arbitrary categories have been established by this investigator, and the excerpts are assigned as was deemed appropriate by him. The categories are (1) descriptions of grading systems, (2) principles of grading, (3) criticisms of grading, and (4) status studies and trends.
in marking practices. Many articles, however, include elements of more than one of the categories.

An example of the first category is an article by Keller, in which he described a dual system of grading using ABCDE as a rating of competitive achievement and S-U as a rating of effort. Keller understood, however, that manipulation of symbols was not in itself a sufficient remedy, and he offers a suggestion.

Changing a method of reporting or the discarding of school marks will not solve the fundamental problem of grading or help to alleviate it very much. Hence, unless we find sufficient reason and proof for eliminating most of the purpose now inherent in the giving of grades, it seems more realistic and logical that the best approach to the problem would be to educate and train teachers better how to use a sound and improved basis of grading. (29.,73)

Another article reports an innovation in the grading system in Yakima, Washington. This system proposed defining achievement symbols as follows: A--top 5 percent; B--75-95 percent; C--25-75 percent; D--5-25 percent; F--lowest 5 percent. It also includes these work habit symbols: E--excellent; S--satisfactory; and U--unsatisfactory. The system also employs the symbol Sp--special education student. This innovation is described as a "vast improvement in Yakima" (42.,14).

A discussion of reported innovations in grading practices could be extended considerably, for a plethora of literature on the topic is available. Other articles discuss dual grading systems, use of unusual symbols or an unusual arrangement of symbols, means of
weighting grades, new report card designs, and many more. In Chapter IV the investigator indicates the extent to which innovations such as these are actually being used in Ohio insofar as his sample is concerned. It becomes apparent at that point that such innovations exist more commonly in the literature than in practice.

The second category, principles of grading, includes a variety of ideas and suggestions from recent writers. Grace, for one, directs his comment to the necessity for assigning a single meaning to each marking symbol. He maintains that

...confusion exists because there is no single meaning to a grade. Once we understand what grades mean, then we will probably use them correctly and be neither confused nor taking issue with the importance of grades. However, as long as we continue to leave grades undefined or misuse them, we may expect to be hung on the horns of a dilemma. (23.93)

Johnson emphasizes the necessity for defining the meaning of any symbol used to mark student achievement, whatever the symbol may be.

It makes little difference whether marks are in the form of numbers, letters, or geometric figures, so long as their meaning is understood. But some of the confusion surrounding marks might be cleared up in formulating a policy which indicates their purposes, the attributes to be marked, the sources of evidence used, the bases of comparison and curricular reference of the marks. At least there might be less of a mess in marking. (28.30)

Anderson, Grim, and Gruhn stress the importance of the principle of individual rather than competitive marking.

Marks, in order to show progress, should reflect individual growth. Competitive marking and relative marking deny known psychological facts regarding individual differences in ability,
interests, and home background by assuming that all types of pupils learn the same things in the same way at the same time. Each child should be evaluated in terms of his own abilities and interests rather than compared with others who may differ markedly. Equal achievement does not necessarily indicate equal abilities or effort. Undue emphasis on competitive marks also frequently results in bad morale and conditions of poor mental health in the classroom. (1.,266-7)

Muessig is concerned in that we irrationally attempt to do that which by definition is impossible, i.e., measure the immeasurable.

You cannot use quantitative means (letters, numbers, rank order, percentiles, and so forth) to assess qualitative aspects of life, of people, of high-school English without losing something. The something is an amorphous, indescribable essence which makes a man man; makes a poet instead of a bricklayer; makes a Shelley rather than an Einstein...(35.,416)

Criticism of grading practices, as the third category of articles, probably occupies more space in the periodicals than any other aspect of this subject. Jarrett provides an example as he condemns overemphasis on marks.

The major problem in the classroom use of marks and grades is not so much the question of how marks and grades should be computed or even of which marking or grading system should be used. Much more important is the discovery of means of keeping teachers, students, and parents from attaching too much importance to marks... The overemphasis on marks interferes seriously with the quality of the learning activities in the school. (26.,37)

Hosmanek and Hueittner point out an inequity that is allowed to exist in the grading system of many schools.

Lost in the maze of educational problems is one which virtually affects numerous of our 'gifted'
students who find themselves in the increasing enriched, accelerated, and 'honors' classes. Simply stated, it is that these students are very frequently penalized for being in such classes when evaluations in the form of grades are made for their records. (26.,115)

Melby is critical of grades from a mental health perspective.

We are more interested in what children study than in what study does to children. Our expertise lies in material we teach, not in understanding children. If we were primarily concerned about children we could not stubbornly hang on to our destructive marking systems and continue to disregard maturation of children in our teaching nor could we as teachers continue to pressure children and youth to get them into prestige colleges regardless of what we do to children's personalities. (34.,449)

Briggs also finds the concern with grades a dangerous preoccupation of both students and teachers.

One of the most alarming situations in many high schools today is the extreme concern of most high school students over grades. This preoccupation with 'grade-getting' is like a fever—it's contagious; and by the upper high school years, it has become a fashionable disease among the high school population. In a sense, this 'grade-mania' sweeping through the secondary schools is much like the common cold. It is easy to diagnose, but virtually impossible to find an immediate cure. (21.,280)

The tendency (of the teacher) is to stress the importance of facts per se rather than to teach for ideas, understandings, and concepts... The student becomes a passive agent relying on the teacher and the textbook as authorities. This kind of teaching leads to conformity and conformity leads to mediocrity. (21.,282)

An unfortunate characteristic of marks which are not founded in
clear policy and definition (and they usually are not) is their
meaninglessness. Leese describes this condition.

Real parent—and pupil—confusion comes when the
wondering and the wounded stir up the temerity to
seek an explanation. It is then one really
discovers the world of intrigue, mesmerism, and
double talk known as the school marking system.
What is so simple, so pristine, so final on the
report card becomes involved, inelegant and
speculative in a conference or at a PTA meeting...
After so many years of fumbling on this subject
we cannot expect to resolve our problems over­
night. But in the name of common sense every
school should have clear policy on the purposes
of marks and the ingredients used to make marks.
We can eliminate the contamination of letter
grades with numerical translations, remove
intangible and unmeasured character traits from
mathematical scores, prevent the teacher from
distorting the academic records with petulant
ratings of personal discipline. (31.,9-10)

The fourth and final category of articles is that of
surveys and/or trends in marking and reporting practices. Wrinkle
describes the trends during the 25 years between 1922 and 1945 as
follows:

1. Many schools report by use of a multiple point
scale commonly involving the use of the letters
ABCD.
2. The greatest single innovation in marking
practices has been the substitution of letter
grades for percent grades.
3. Most schools in addition to reporting a letter
grade also report on a variety of character
traits by checking one or more items in a
printed list of undefined terms or statements.
4. Most of the departures from conventional practice
have been made by elementary schools and very
few by secondary schools.
5. Most schools send out reports each six weeks.
6. Most of the schools that have made departures from conventional practices (represented by the use of informal letters, parent-teacher conferences, the substitution of S and U or H, S, and U for ABCDF, etc.) are satisfied with their new practice and are working to improve it still further.

7. The marking and reporting problem ranks close to the top among those about which most teachers and schools are seriously concerned, but since they do not know how to improve on what they are doing, they do not change. (19.,30)

The Journal of Education included in 1954 a list of grading practices which it claimed represented the trends at that time. The article did not include the criterion used by the author to decide what constituted a trend.

Decreasing

1. Unexplained single marks
2. Competitive marking
3. Use of percent marking

Increasing

1. Defining the meaning of marks
2. Grouping of subjects in broad fields
3. Listing objectives and skills to be accomplished in individual subjects
4. Including explanations of grades, reasons for lack of accomplishment, suggestions for improvement
5. Reporting special abilities, interests, activities
6. Including profiles of student performance, including objective test data
7. Reporting social and emotional development and behavior
8. Reporting on physical and health development
9. Describing philosophy, purpose and organization of the school
10. Individualized personalized letter form reports
11. Parent-teacher conferences as a substitute or supplement to the report card
(38.,191)
Terwilliger conducted a study in 1963 of 129 high schools selected to be representative of all high schools to ascertain what marking practices they employed. Certain findings from his study have been referred to for comparative purposes in Chapter VI of this study. His conclusion from the findings was as follows:

At the risk of sounding grandiose the author would like to suggest that what is truly needed is a revolution in marking. We should be concerned not only with the standardization but also with the justification of grading policies and practices. This means calling into question such cherished notions as the 'absolute standard,' penalizing or rewarding students for classroom behavior or absences by revising grades, etc. Our data suggest that such notions underlie much of the evaluation presently being done in secondary schools. However, they reflect blind adherence to tradition much more than what is presently known about sound educational measurement. (42,36)

Members of the Ad Hoc Committee on Marking Practices of the National Association of Secondary School Principals reacted to the above study with suggested action guides for principals.

1. Decide that this problem is really serious and that your school will try to contribute its share toward solution.
2. Read (Terwilliger's report) and study the implications of the data.
3. Survey your school and find out what teachers and department heads think their marking policies are, what primary bases for marking are used, what the actual marking practices seem to be, how class rank is established.
4. Reconstruct your marking, reporting, and ranking practices so that they are consistent with agreed-upon policies and consistent among departments.
5. Provide periodic checks to satisfy yourselves that your agreed-upon policies and practices are being followed consistently. (42,4)
For that person who is interested in looking further into the recent literature on the topic of grading, this investigator recommends a selected bibliography published by the United States Department of Health, Education, and Welfare, Pupil Marks and School Marking Systems (49.). This is an annotated bibliography that includes the best references from books, pamphlets, and periodicals through January, 1963, with most references being from the latter half of the fifties and the early sixties. Also, Chapter II of this study includes many more references to the literature, particularly those from earlier years, many of which would be appropriate in the review of the literature discussed in this section except for the duplication that would occur.

As one looks more extensively into the literature than could be presented in this brief review and few representative excerpts, he may expect to find many dozens of articles in each of the categories suggested by this investigator. Insofar as the specific marking practices that are suggested, there is a great deal of incompatibility from author to author, and often downright contradiction. One may find documentation and support for practically any marking practice or point of view he may wish to adopt, whether it be increasing the number of grading scales or using just one, increasing the number of symbols for more specific descriptions or using simply S-U, increasing emphasis on making grades more meaningful or eliminating grades entirely. They wide variety of opinion, and
the incompatibility and contradiction are testimony to the pernicious nature of the problems of marking and reporting pupil achievement. Nevertheless, one finds general agreement of the authors on certain of the needs for improving marking practices such as more clearly defining grading policy, greater acceptance of responsibility by administrators for improving marking practices, and increased understanding of the mechanics and philosophy of marking by the teachers.

There is a particularly discouraging feature in the whole area of grading. The literature attests to the strides made in nearly every other field of education, particularly since 1957. Yet in the past ten years, no evidence of substantial progress can be detected in marking practices. One might even consider that marking practices have regressed because so many schools have adopted computerized reporting procedures that limit and depersonalize school marks. However, as will be pointed out in Chapter II, it is the nature of marking practices to lag behind developments in other areas of the educative process.

Innovation in education

Innovation in education is a topic rarely considered in the literature prior to the late fifties. The general speed-up of change in education following the White House Conference in 1955, and even more so following Sputnik in 1957, has led investigators to inquire into the nature of change in education.

Innovation in education does not exist as an entity separate
and distinct from all others, but rather is one aspect of a more
general study of change. The study of innovation in such fields as
agriculture, medicine, and industry preceded the same sort of investi-
gation in education. Such concepts as change agents, change
mechanisms, and target systems, and such topics as the nature of
innovation and diffusion of innovations were developed prior to their
application to the study of change in education.

Several attempts have been made to classify the character
of innovation. Chin (46.) suggested a set of descriptions of innova-
tion arranged in escalating order indicating the increasing extent
of change from the original (i.e., prior to change) structure, or
target system: (1) substitution of one element for another, (2)
alteration of basic elements, (3) perturbations and variations in
which temporary oscillations in a system occur, (4) restructuring or
basic structural modifications, and (5) value-orientation change re-
quiring basic attitudinal or cultural changes. Rogers (13., 146)
attempted to define the factors which affect the rate of adoption of
any innovation: (1) relative advantage— the degree to which in-
novation is superior to the idea it supersedes, (2) compatibility—
the degree to which an innovation is consistent with existing values
and past experiences of adapters, (3) complexity— the degree to which
an innovation is relatively difficult to understand and use, (4)
divisibility— the degree to which an innovation may be tried on a
limited basis, and (5) communicability— the degree to which the
results of an innovation may be diffused to others.
Implied in the idea of change is the opposing notion of resistance to change. This resistance may take the form of conservatism, tradition, rigidity, lethargy, complacency, immunity, or imperceptibility. Willower (45.), Guest (5.), Likert (8.), and Eichholz (48.) all discuss the various aspects of resistance to change.

The so-called Brickell Report (3.) was a study on how innovation could be expedited in New York State. The chief finding from this study was that innovation has three separate phases: (1) design, (2) evaluation, and (3) dissemination. Each of these phases is by its nature exclusive of the others, and each must be carried out separately by different agents under different circumstances. If innovation is to occur effectively, Brickell contends, it will necessitate a major statewide reorganization of the structures of education.

The single most comprehensive book dealing with innovation is *Innovation in Education*, a compilation of views from a number of writers edited by Miles (10.). The many well-known contributors to the book include Ronald Lippitt, Paul Mort, Daniel Griffith, Benjamin Willis, Martin Mayer, Theodore Bremeld, and others. The writers deal with both theory and practice as they direct their attention to the educational system as the context within which innovation occurs, the characteristics of educational innovations, innovative persons and groups, and the planning and execution of change processes.
In his book, Miles provides a working definition of innovation which suits the purpose of this study as well.

Innovation is a species of the genus 'change.' Generally speaking, it seems useful to define an innovation as a deliberate, novel, specific change, which is thought to be more efficacious in accomplishing the goals of a system. ...The element of novelty, implying recombination of parts or a qualitative difference from existing forms, seems quite essential. (10.,14)

In summarizing the many authors' contributions to the book, Miles has stated what, in the broadest sense, is also the concern of this study.

Often, much more attention is put on constructing the innovation itself than on planning and carrying out the strategy for gaining its adoption.

Yet it seems clear that for almost all innovations the process of implementation itself needs careful study, planning, and experimental work. (10.,647)

The most important idea in the book, referred to again and again by the several contributors, is the nebulous nature of innovation. "There is no single factor, in and of itself, which is highly related to adaptability" (10.,321). Innovation is an intricate process influenced not so much by one factor as by the interrelationship of several factors. This quality of innovation explains why the process relating to it is understood only to a limited degree.

A chapter by Paul Mort in the book is especially interesting in that it summarizes the overarching findings from the research that has been done in the area of educational innovation. Mort was
particularly concerned with the time factor, or lag, that exists in the spread of any innovation, and he suggested the necessity for patience in innovation.

Even though no one factor can be identified that is highly related to adaptability, there are factors in education which are influential in the change process. Demeter found that

Building principals are key figures in the (change) process. When they are both aware of and sympathetic to an innovation, it tends to prosper. When they are either ignorant of its existence, or apathetic if not hostile, it tends to remain outside the bloodstream of the school. (47,23)

According to Ross (14,567-9) in a summary of research on finance and adaptability, communities characterized by high average personal incomes and abundant financial support of schools tend to be the setting of the most innovative school systems. Also regarding the economic factor, Rogers maintains that

Among the great variety of factors related to innovativeness among schools, the best single predictor of this dimension is educational cost per pupil. The wealth factor almost appears to be a necessary prerequisite for innovativeness among public schools. (13,40)

Stufflebeam has a clear insight into the necessity for examining and defining the factors which comprise the dynamics of change.

Many changes are taking place in our schools and colleges and undoubtedly many more changes will be required and attempted in the near future. Yet, we know far less than is desirable about almost every aspect of change as it occurs in education. There is an urgent need to understand and control the change process so that needed new educational programs can be developed on sound bases, diffused rapidly and thoroughly, and implemented effectively. (51,1)
A corollary to Stufflebeam's statement is that without an adequate understanding of the change process, improvement in education will be sporadic, inefficient, and costly. For the educator who is aware and concerned, this condition presents a challenge.

Significance of the Study

As one becomes familiar with the literature in the area of marking practices and grading systems, it becomes apparent that perceptive educators are concerned with the "mess in marks." Leese claims that school marking systems are a "world of intrigue, mesmerism, and double-talk" (31,9). Melby characterizes marking systems as "destructive" (34,449). Farwell states that "of the many practices found in education today, those involved in marking and reporting appear to be most completely shrouded in confusion, misinterpretation, and lack of understanding" (22,462).

Much of the cause for the continuing concern and frustration about the shortcomings of marking systems stems from the fact that there seems to be no possible once-and-for-all solution. The situation appears to be one of "we can't live with them and we can't live without them." This investigator sees no alternative but to accept this condition, though not necessarily in the spirit of resignation; there is nothing in this condition that prevents worthwhile improvements from being made in marking practices.

It was stated previously that the writer agrees with Wrinkle that the grading system in any school should be in the most
direct possible way an extension of the curriculum. Therefore, as
the content and practices, and especially the objectives, of the
curriculum evolve in response to the influences of society, so should
the marking practices change commensurately. In this era in which
innovations in curriculum are occurring rapidly, it follows that
corresponding changes should occur in grading systems. The literature
indicates that such changes do not always occur. One may speculate
as to the reasons: many educators do not understand the direct and
necessary relationship between the curriculum and the grading system;
tradition and parental insistence have made change in grading practices
difficult to implement; apathy and lack of any pressure for change
have made the status quo comfortable; resignation to the attitude that
grades are a necessary evil to be tolerated, and that one grading
system is about as good as another.

This investigator contends that none of the above reasons,
if scrutinized from a professional point of view, is tenable. It is
further contended that marking practices can be improved in directions
indicated by many writers, and that such improvement will occur if
schoolmen are willing to experiment and innovate. While the available
literature includes suggested principles to be incorporated in
grading systems, and descriptions of marking practices that have been
incorporated in various schools, there is very little literature that
describes how change in grading systems can effectively be implemented.
It is this gap in the literature that this study is designed partially
to fill. Through identifying the characteristics associated with
innovation in grading systems as well as those associated with inertia, and by reporting practices which have been successful in implementing improvements in grading systems, the investigator intends to present information hitherto unavailable in this form regarding the dynamics and procedures related to improvement of marking practices.

Limitations

The basis for the selection of a study sample was Stufflebeam's Ohio Education Innovations Survey. Therefore, the limitations of that study delimit this study as well. This includes the fact that Stufflebeam did not receive a 100 percent response to his study. Those school districts which did not respond to his study could not be included in this one.

Of the 1084 public high schools in Ohio in 1964-1965, only 80 were selected to participate in this study. Of those 80 schools, 66 responded to the questionnaire. The effect of these factors on this study data is not assumed to be large, but is nevertheless indeterminable.

The respondents to the questionnaires and the interviewees were exclusively principals of the high schools in the study group. The responses they gave necessarily reflected an administrative point of view. Responses from other school groups would have varied as the perceptions of those groups vary.
Plan of the Study

The first chapter of the study has described the background of the problem, stated the problem and the intent of the study, and presented a review of the related literature.

The second chapter will review the evolution of marking practices from 1770 to the present day. It is the purpose of this chapter to examine the various marking practices as extensions of the predominant mode of education of the time.

The third chapter will describe the procedures with which the investigator conducted the study. The several procedures include the use of (1) the Stufflebeam study (Ohio Education Innovations Survey) to select the study sample, (2) Ohio State Department of Education publications to obtain statistical information, (3) a questionnaire to survey 80 high schools in Ohio, and (4) personal interviews with ten principals from study sample schools to obtain their views relating to the problem.

The fourth chapter will present the data obtained from the questionnaire in 14 sections, the first 13 of which correspond to the 13 hypotheses, and the last of which is a free response. The data are displayed in a series of tables, and a description accompanies each table. Also included are the decisions of the investigator regarding acceptance or rejection of each hypothesis.

Chapter V will present the supplementary data that support the chief body of data obtained from the questionnaire. These data include the statistical information from the state publications,
a sampling of modern marking practices reflected in report cards
and statements of marking policy submitted by the principals, and
the investigator's impressions from his interviews with the princi-
pals.

Chapter VI, the final chapter, will present a discussion
of the findings from the several sources, the conclusions of the
investigator, recommendations relating to means for changing marking
practices, and recommendations regarding desirable directions in
which change should occur.
CHAPTER IX

REVIEW OF PAST MARKING PRACTICES

Secondary school marking practices, as we know them today, have undergone few basic changes in the past two or three generations. Change has occurred less extensively in this area than in most other aspects of education. Nevertheless, marking practices have not always been the familiar ABCDF or the 100 point scale. These were adopted after a century or more of using a variety of methods for indicating achievement. A brief review of past marking practices provides a basis for understanding how our present methods evolved, why they have changed so little in the past half-century, and what changes in our present practices would be desirable.

1770-1826

The early forerunner of grading as it is now performed was the public examination. Kandel provides an account of this.

In the secondary schools, the academies, and later the high schools, examinations were public and were conducted quarterly or annually by committees. Thus in 1770 the Trustees of the Philadelphia Academy 'considering that nothing can contribute more to keep up a laudable Emulation and Diligence among the youth educated in this Seminary than regular and stated examinations of the different schools and bestowing Premiums and other marks of approbation on the most deserving Scholars' appointed four days in the year for such examinations. There were no
fixed practices. Examinations were held every quarter or every term, or annually as well as at other fixed dates, or semi-annually, or even monthly. (6.,23)

A further report by Kandel elaborates somewhat on the visiting committee as an agent in conducting public oral examinations. This particular statement describes the practice current in the Boston secondary schools in 1789.

Such examinations were, of course, oral, and were conducted either by the teachers themselves or under the supervision of the visitors or by the visitors themselves. Mass examinations could, in fact, not be conducted, since instruction in the main was individual, each pupil recited to his teacher, and advancement or promotion from one stage to another depended on the teacher's opinion of the pupil's readiness. (6.,22)

By 1826, the method of evaluation of students had begun to change somewhat, as indicated by an article in the American Journal of Education.

At the close of every month, the boys in each apartment undergo a rigid examination in all the studies of that month. This is conducted by the principal, with whom only the first class remain permanently, in the presence of their particular teacher, and such other instructors of the school as find it convenient to attend. These monthly examinations are sometimes attended by the sub-committee of the school, and are open for parents, and any other persons interested. If any class, or any individuals do not pass satisfactory examinations, they are put back and made to go over the portion of studies in which they are deficient till they do pass a satisfactory examination. The rank of each scholar and his seat for the succeeding month are determined by this examination; unless an account of places for each recitation of the month has been kept, in which case they are determined by a general average. They boy at the head of the first
division of the first class is monitor for the month. (39.,266)

That grading practices employed in the early 19th century were varied is attested to by a description of the evaluative method used in the Boston High School for Girls in 1827.

The government of the school is vested in a set of books in which is recorded an accurate and minute account of every scholar's performances, deportment, absence, and tardiness; and at the end of each quarter, she is advanced to a higher, or degraded to a lower, section or seat, as this record shall appear in her favor or against her. The whole business is regulated by fixed principles, that are well understood; and every individual is, literally speaking, the artificer of her own rank, which is affected by every exercise she performs, and by every error she commits, either in recitation or conduct. Everything depends upon numerical calculation; and, were it expedient, the school might be classed by the scholars themselves.

A Credit is given to every member of the school, for each regular recitation which is performed in a correct and satisfactory manner.

A Check is given to such as fail their lessons. Thus, in every recitation, each pupil receives either a check or a credit.

Merits are awarded for correct and orderly deportment, for excellence in the usual exercises of school, and for voluntary labors.

Forfeits are incurred by neglecting to attend to required exercises, at the proper time, and in a proper manner, by making appeals without sufficient grounds, and by slight irregularities of conduct.

In making the quarterly records of the school, by which the rank of scholars is determined, every check cancels one credit, or two merits, and every misdemeanor, five merits. (20.,206)

The principal of the school, the author of this article, also describes his role in evaluation.

To be sure the teachers do their duty faithfully, I regularly review the scholars in all
their studies. The books of the school are divided into convenient and proper stages; and a section cannot go forward to a new stage, before they have passed thorough and critical examination in the last. In this way, I eventually attend to all the studies of the school in person. The salutary influence of these reviews, both upon the teachers and their pupils, will be readily conceived. The pupils, on their part, are anxious to go on fast, that the lower sections may not pass by them; and the teachers are as anxious that they should go on well, that they may not incur the mortification which must result from an unsatisfactory review. (20, 209)

In the half century or so covered by the references cited thus far, it is clear that evaluation of students was changing just as the character of the schools was changing. During the 1700's, there were but relatively few secondary school students, and teaching was done largely on an individual basis, each child moving along in his studies at a rate appropriate for him. Also because of the small number of students, an examination of each student by the school trustees or the visiting committee was feasible. This method of evaluation existed for at least two reasons. First, the quality of instructor that was available was not always extremely high, and the responsible citizens hesitated to place unlimited faith in him. The townspeople therefore reserved the privilege of confirming the teacher's judgment regarding which students were prepared to pass on to more rigorous studies. Second, public examination of students was a means of evaluating the skill of the teacher.

In the early 1800's, the number of secondary school students was increasing rapidly. Individual instruction of each pupil by the
teacher became an impossibility. As a result, the Lancasterian system of monitory study came into vogue. This system utilized advanced students in the capacity of tutors, and so saved in the expense of several additional instructor's salaries. It is this method that is described in the two articles from the American Journal of Education dating from the mid-1820's. The concept of the graded school had not yet been incorporated, but students were assigned to "apartments." These were small groups of students who moved along together through progressively more rigorous stages as a unit. At the end of each month, examination of the students supplemented by consideration of their skill in recitations and their deportment provided the basis for determining whether the student should move on with his "apartment" or be placed in a group involved in studies a step behind.

While these schools were still open to the "school sub-committees" and to parents and citizens during the times of examinations, it is significant that the promotion of students was by 1826 largely in the hands of the instructors and particularly the principal. Also significant is the fact that no longer could one School Committee serve all the schools in the larger cities. Sub-committees were appointed in these cities and assigned responsibility for a certain number of schools. However, even these sub-committees took less and less part in the actual examination of students, and eventually confined their function largely to judging the practices of the school officials.
1826-1900

The growth of schools, both in size and numbers, heralded the end of the era of all but token outside involvement directly in school activities. As a result, other means of evaluating students evolved. The written examination was the most widely adopted device. It did not immediately replace the public examination, but eventually the public examination faded from use altogether, a victim of progress and expediency. In 1847, however, according to Page, public examinations were still in evidence, if not without objection.

It is now the usage of all our schools to have public examinations,—generally at the close of a term, or portion of a term,—in order to test, in some measure, the industry and skill of the teacher, and the proficiency of the pupils. ... Still public examinations as frequently conducted, are not without their serious objections.

(11.,290)

...I think too much stress is at present placed upon showy exhibitions and celebrations, and that objections and dangers attend examinations, as frequently conducted. I would not recommend altogether their discontinuance. I would rather urge that the teacher, by his inflexible honesty, should make them a fair representation of the actual conditions of his school...(11.,294)

According to Page, then, who was one of the important educational commentators of the mid-19th century, public examinations became something of an "open-house" to show off the school rather than a significant instrument of evaluation of either student or teacher. By 1845, the Boston School Committee, involved in the process of conducting an evaluation of the public schools, was resigned to the fact that it was
'...impossible to do anything like justice on the (oral) examination, for Committee-men could not be found who could give the time to examine over 7,000 children.' ...The Subcommittee accordingly decided to introduce a written examination 'in addition to the usual mode of oral examination.' (6.,25)

The written examination was indeed accepted widely and enthusiastically near the middle of the century. The following report describing the Public High School in Chicago in 1857 verifies this fact:

Besides frequent oral reviews in the different branches of study, at the close of each term the several classes are subjected to a written examination of all the general topics to which they have attended, and no pupil is advanced to a higher class till he has fully established his claim to the new position. These written reviews are among the most successful means that can be employed for securing thoroughness and accuracy of scholarship. Several topics are written distinctly on the blackboard, and the students are required to expand them as fully and accurately as possible. Each pupil is seated by himself, and furnished with pen and paper; but receives no assistance, direct or indirect, from either teacher or textbook. This mode of examining a class accomplishes at least three important objects at the same time. It affords a thorough test of the pupil's knowledge of the subject; it is one of the best methods of cultivating freedom and accuracy in the use of language; and it furnishes a valuable discipline to the pupil's mind, by throwing him entirely on his own resources. (44.,266)

As early as 1847, Horace Mann, anticipating the era of scientific evaluation, noted in his Common School Journal that

We venture to predict, that the mode of examination, by printed questions and written answers, will constitute a new era in the history of our schools. There is a variety of reasons which
give it a decided superiority over any and all other methods... (7, 494)

Horace Mann's foresight regarding the role that written examinations would play in education was quite accurate. Written examinations became a major factor in evaluating students in their academic work, and they also became the basis for the whole standardized testing movement begun by Rice in 1898. But written examinations were not to be the end-all in the evaluation and marking of students; their shortcomings were realized all too soon. By 1887, in fact, Cincinnati had ceased to employ written examinations at the completion of coursework as a factor in determining promotion of students.

School board policy adopted for the 1887 school year read as follows:

There shall be no stated examinations for the promotion of pupils in the several grades of the district and intermediate schools, but the pupils in these grades shall be promoted and classified primarily on their proficiency in the several branches of the course, as shown by the teachers' estimates of their daily work...

(52, 25)

Emerson White, Superintendent of Schools in Cincinnati, and a leading educational administrator of the time, discussed written examinations relevant to his decision to eliminate them as a factor in student promotion.

When the written examination was first introduced as a basis for promotion, special stress was laid upon the final examination; i.e., the examination held at the time promotions were made. The work actually done by the pupils during the term or year was set aside practically, and their promotion was made to depend on the percent of correct answers given by them to the questions proposed. (52, 32)
These several uses of examination results have been the source of bitter jealousies and rivalries between schools and teachers. They have perverted the best efforts of teachers, and narrowed and grooved instruction; they have occasioned and made well-nigh imperative the use of mechanical and rote methods of teaching; they have occasioned cramming and other vicious methods of study; they have caused much of the overpressure charged upon the schools, some of which is real; they have tempted both teachers and pupils to dishonesty; and last but not least, they have permitted a mechanical method of school supervision. (52, 33)

Many a teacher's scrapbook contains the series of questions presented in previous examinations, and their use in preparing pupils for the next ordeal is not an unknown art. Moreover, teachers are not slow in learning the 'question hobbies' of the superintendent, and in turning this information to good account. The fact has contributed not a little to the increasing percentages in successive examinations which more than one superintendent has accepted as gratifying evidence of marked progress in the schools. (52, 34)

Cincinnati's move to the elimination of examinations was not followed widely by other school systems at the time, but it did serve to emphasize the shortcomings of examinations. Prior to this time, promotion of students in many high schools was contingent solely on examination results. This practice was hastened into obsolescence by the poor press accorded the written examination during the period of the 1880's.

Another practice that existed as early as the 1880's and continues in use up through the present time is that of exemption from semester examinations for those students whose work is of high quality during the semester. Harlow asked in 1888, "Is it not, however, the
prevailing opinion that a record of daily work is the most reliable index of the standing of a pupil?" (24.,322), and then described how Syracuse High School "for some time past" had allowed examination exemptions. This means, of course, that examinations were certainly not the sole basis, nor even the most important basis, for determining the course grade, or for promotion.

During the earliest years of secondary education discussed in this chapter, i.e., the 1770's, the marking of students was simply a pass-fail procedure as determined by an oral examination by the schoolmaster and/or the visiting committee. By the 1820's, there was an attempt to evaluate the student on a continuing basis and in several aspects as indicated in the Boston High School for Girls through the use of credits, checks, merits, and forfeits. As written examinations gained in prominence, student achievement was determined through the total number of points accrued on examinations and recitations. After 1850, the use of scales was begun. For some years, however, there was little agreement as to what scale should be used. Payne commented on this situation in 1875.

It is of considerable importance to decide on a scale of marks. Some prefer 100, some 10, and others 5. The first seems to allow too great a latitude; the teacher is in danger of forming a loose estimate of values. The last is, perhaps, too limited in its range, not elastic enough to represent all the variations in quality which recitations are likely to present. (12.,165)

Because of the inconsistencies in marking practices from school to school, colleges paid little attention to records of high
school marks, but rather examined each student who sought entrance. Then, in the 1870's, the University of Michigan and Northwestern University began admitting students without examination from certain approved high schools. Other colleges soon followed suit, and before long, colleges traded lists of approved high schools. This cooperation between colleges resulted, in the 1880's, in the formation of the first of the Associations of Secondary Schools and Colleges. Among the early innovations instituted through the influence of the newly founded organizations was an increase in the consistency of the marking pattern from school to school. The pattern agreed upon was the 100-point scale on which either 70 or 75 points was considered to be passing. This same basis continues to be used by many schools and many teachers today, although a conversion to some symbol is usually made from the scale before the mark is officially recorded.

The appraisal of students as it was conducted just prior to the turn of the century is the topic of a statement by Williams.

Under such circumstances (i.e., the rote teaching methods of that time) student appraisal was a simple and very direct process. Teachers had a roll book which contained the names of the students, with a space for every school day in which to record the daily achievement of each student. The teacher would call upon a student to recite and then would record in the book an estimate of the faithfulness with which the student had learned the lesson. These estimates were usually recorded on a scale of 10 for 'perfect.' At the end of the term the average of these marks would be computed, the results of examinations would be averaged (sometimes weighted), and these averages would be added and another average computed, the result being expressed as a percentage on a
scale of 100 for 'perfect.' Not infrequently this term average was worked out to three or four decimal places. The result constituted an appraisal of the students' total success in learning what he was expected to learn. This form of student appraisal was wholly in terms of what students had learned in school of subject matter and proper conduct. The procedure operated as an excellent device to select those who demonstrated fitness to be promoted into the work of the next year or into some institution of higher grade. It also served to discourage those who could not, or would not, fit themselves into the prescribed pattern. The real appraisal was, therefore, whether or not the student had been pulled, hauled, squeezed, and forced into a predetermined and rigid pattern of learning and conduct. In this respect student appraisal could be, and often was, used as a basis for judgment of the success of the teachers. The major purpose was wholly integrative in nature—to impress a uniform stamp of culture upon all students.

(17.,314-15)

1900-1930

At the turn of the twentieth century, a new psychology, heralded by men such as John Dewey, G. Stanley Hall, and Edward L. Thorndike, began to influence education. No longer was empiricism and intuition sufficient cause for innovation in education. New concepts such as the intelligence quotient and individual differences, new learning theory to replace the naive Herbartian approach, and new statistical methods all were developed and became cogent factors in determining the directions in which educational change occurred. No aspect of education was exempt from scrutiny, and no aspect escaped criticism. Marking practices were no exception.

The greatest criticism of school marks was the extreme
variability that existed in any comparison of sets of marks, and hence an inherent degree of meaninglessness in marks. The extent of variability astounded the researchers, for, as Starch reports in 1918,

Until a decade ago, no one questioned either the validity or the fairness of these measurements (i.e., school marks given by teachers). It was tacitly assumed that marks were almost absolutely correct, or very nearly so, a fact attested by the surprisingly common practice of marking to the fractional part of a point even on a 100 percentage basis. (15.3)

Max Meyer conducted the first research in 1908 on the lack of conformity of school marks, but Daniel Starch was to make the greatest contribution in understanding this phenomenon. He began research on this topic in 1912, and culminated the work in 1918 with his book *Educational Measurements*.

One of the myths disproved by Starch and others early in this century, and yet a belief that continues to be held by a great many teachers and laymen alike, is that the grading of essay tests results in greater variability, and less reliability, than the grading of so-called objective tests. Researchers consistently found approximately the same degree of variability regardless of the type of examination being graded.

In a research study in 1913 that was sophisticated for its time, Starch was able to identify four factors responsible for variability in marking: These included—

1. Differences among the standards of different schools
2. Differences among the standards of different teachers
3. Differences in the relative values placed by different teachers upon various elements in a paper
4. Differences due to pure inability to distinguish between closely allied degrees of merit. (40.,630)

The factor contributing most to variability is the last mentioned, while the third, second, and first contribute less in that order.

The work of Meyer and Starch has been carried on right up to the present time by interested researchers. As one considers the content of the many articles dealing with variability, the most significant aspect is the remarkable extent of variability in all marking practices. A resume of the findings of the dozens of studies of variability must include the following conclusions:

1. Variability in marking exists to a much greater degree than most educators imagine.

2. While the factors that contribute to variability can be identified and in many cases reduced in their effect, they are by their nature impossible to eliminate.

3. Variability in grading can be reduced within any given situation by a common agreement on standards and methods of grading.

Perhaps the greatest influence which Starch had was in the area of grading scales. The hundred point scale on which a mark expressed a percent of perfection was the prevalent procedure for grading, and its efficacy was not questioned prior to 1913. In 1913,
Starch did question it with a research study, and in 1918 he reviewed his findings as follows:

If we are attempting to evaluate a paper by a scale of 100, 99, 98, 97, 96, 95, etc., we are attempting to make finer distinctions than we are capable of. The mind does not discriminate with any degree of certainty by a single judgment between a paper of grade 85 and another of grade 86. If the second is appreciably better, it more likely ought to have a grade of 90. Such small distinctions would have validity only if a set of papers were graded by ten teachers or by the same teacher ten times. The situation is analogous to estimating the width of a room in inches when it should be estimated in feet. Estimates in terms of large units, of course, do not have greater absolute accuracy, but they will have greater uniformity. (15.,11)

On the basis of experiment, and using mean variation and probable error as a statistical basis, Starch concluded that the most satisfactory marking system should be a scale of five steps: "A, or excellent, 7 percent of students; B, or superior, 24 percent; C, or average, 38 percent; D, inferior, 24 percent; E, or unsatisfactory, 7 percent" (15.,13).

There is little doubt that no one man has done more to change the grading pattern in American schools than Starch. The five-step scale was not an original idea with Starch, but he gave it great impetus by providing it with a scientific basis. He also associated the concept of normal distribution, or the "curve," with teacher marking, as may be observed in his distribution of symbols above. This latter device continues to be used widely today even though it has been the subject of controversy since its inception.
As early as 1911, F. W. Johnson commented on the lack of uniformity in grades and the need for uniform standards as follows:

While with individual teachers slight differences in the use of grades may represent different degrees of effectiveness or differences in actual abilities of pupils, no such reason could possibly be assigned for the remarkable variations shown (in this study). In general it may be assumed that wide variations are due to the lack of a uniform standard of grades employed. (27.,20)

Johnson was only the first of numbers of educators to be concerned about bringing uniformity into grading practices. Starch, in 1913, says, "One suggestion (to improve reliability of grades) would be the adoption by all schools of some uniform marking system" (41.,681).

In 1917, G. E. Maxwell claimed that "...if the grading of students was to possess validity and command respect it must depend on some more objective standard and thus exhibit greater uniformity in results, at least among the teachers of the same institution" (33.,114).

These are but a few representative statements that appear in the literature beginning in 1911 and extending through the 1920's. The adoption through this period by many schools and teachers of the ABCDF grading scale and the normal curve as a guide to the distribution of grades did help in achieving a degree of uniformity, but no really adequate solution to this problem has yet been found.

Closely related to the attempt to bring uniformity to grading practices was the increasingly pressing need to define the meaning of grades. In 1919, Sherwin Cody (4.,50) suggested dual grades, one for absolute achievement, and one for relative improvement.
One mark, he claimed, cannot adequately express both, nor is it fair to the student to try. In 1923, Ben Wood commented on the topic as follows:

If we are to grade at all, they ought to be such that they can be used intelligently; the only way they may be made to serve intelligent use is to reach a measurable agreement on what they are intended to indicate. Grading in high school and college should become more a matter of describing facts exactly and dispassionately and less a vehicle for the personal opinions and moralistic anxieties and predilections of instructors. ... If it is desirable to make a record of the effort put forth by the student in each subject, this record should be so separate and distinct in its definition as not to be confused by any chance with the actual achievement index in each subject. If it seems expedient for each teacher to pass upon the character, personality, fitness for life, or general intelligence of the student, these marks should be distinctly defined and kept separate from all other marks. (18,116)

In 1927, Charles Lindsay related the problem of the meaninglessness of grades to the lag in applying psychological principles to marking practices and to inadequate teacher training.

...the writer has attempted to point out, not only that there is a general lack of uniformity of opinion among teachers as to what marks mean, but that the present training of teachers does not seem to change their views with regard to what they mean or should mean; that mental individual differences are practically ignored in marking pupils, frequently resulting in grave injustice to them, and in sending them out of our schools feeling that they cannot succeed in anything. This sense of failure is largely responsible for the danger of defeating the aim of education itself, which is so obviously lost sight of by the teacher when she marks the pupil; that our practice in marking lags far behind our most advanced theory and knowledge of individual differences.
It is not the writer's purpose to suggest a remedy for this situation, but it would appear that more stress in teacher training should be put upon marks and their meaning. It would appear too, that there should be more correlation between our present knowledge of mental individual differences and our practice in marking, and that the significant aim of education should not be endangered by burial under a routine duty, the meaning of which is not understood. (32.,416)

Sidney Pressey saw the meaninglessness of grades from another perspective. He was among the first persons who became aware of a particular shortcoming in marking, the nature of which he commented on in the following excerpt from a 1927 article:

...the writer is convinced that there are problems in marking systems which are not yet generally recognized. In particular it would seem desirable that the marking system should take account of objectives and should express in some way progress toward objectives - should express to a child the extent to which he is reaching a standard... The amount of it seems to be that the concept of objectives has hardly yet got into educational procedures at all - and not even into many departments of educational theory. (37.,737-8)

William Wrinkle took this idea regarding objectives, and Wood's contention that a single symbol can communicate but a single meaning, and in 1947 made them the basis for his book, Improving Marking and Reporting Practices in Elementary and Secondary Schools. Little has ever been done, however, in terms of implementing the suggestions of Wood, Pressey, and Wrinkle to correlate marking symbols with specific objectives and with single meanings in order to endow the marking symbol with a greater potential for communicable meaning.
The potential harm inherent in school marks was understood well enough in 1909 by E. A. Kirkpatrick to prompt him to suggest that marks not be given at all in some subjects.

If agreement can be secured (that grading is inaccurate at least in certain subjects) upon a few subjects what objection can be offered to removing them from the influence of the marking system? Let courses of study in high schools and colleges be arranged with the understanding that one or more of those belonging to the cultural group shall be among the required studies for every student, but that no marks shall be given in such subjects, although credit will be given for the time spent on such subjects in granting degrees of all kinds. ...whatever the value of the (marking) system, it should not be held to in cases where it has the least value and may do the most harm. (30.,198-9)

The same suggestion of having no grades at all is made again in 1926 by Emmanuel Northrup.

Is it really necessary or wise to grade students? Why is it essential that we know the relative standing of the student in his classes? What is the advantage of our knowing, even in a general way, the grade of a student in his class? Is not this the only essential thing: that we know that the student is able to take up the next subject in a course and do the work in a fitting way. (36.,28)

The two statements reproduced here are taken from but two of dozens of articles which suggest the elimination of grades. The major reasons given are that grades are meaningless anyway, they are

\[2\text{This article in 1909 was written prior to the time when Starch showed marks in all subjects to be inaccurate.}\]
unfair, undemocratic, and potentially harmful to students, and they cause the teacher many hours of unnecessary clerical drudgery.

Throughout the 'teens and twenties, article after article suggested new marking devices and grading systems. Included were ideas about weighting grades, deriving coefficients, various distributions based on the normal curve, a simple pass-fail system, assigning symbols to percent ranges, assigning numerical points to letter symbols, use of percentile, grading separately for attendance, quality, and quantity, and many others. For the most part, however, these represented mechanical changes and gimmicks, and the scientific approach to grading begun by Starch and his contemporaries was lost sight of. The problems of marking were, and continue to be, profound problems, rooted in behavioral psychology, our democratic educational philosophy, curriculum development, and learning theory. Not one major marking problem was solved by the superficial innovations that were suggested and implemented, then perished, in that brief span of twenty years.

1930-1965

Just as Starch was the most prominent name associated with marking practices in the early years of this century, William Wrinkle has been the most influential person in this field since the mid-thirties. He was, during the thirties, the director of the laboratory school at Colorado State College, Greeley, Colorado, and so had more than the usual freedom to experiment and innovate. During those years, at the height of the progressive education movement,
Wrinkle employed a number of means of student evaluation, including student self-evaluation, cooperative student-teacher evaluation, teacher-written evaluation, checklists, profiles, as well as several different sets of symbols. None of them, Wrinkle reports, was really successful in the sense that it was a panacea for the really fundamental marking problems.

Wrinkle's contribution to the whole business of the grading of students has a negative dimension in that he can tell us better what not to do than specifically what to do. It was his conclusion in 1947 that marking problems can never be solved on a broad education-wide basis, for marking problems are by their nature specific in each school or school system. Marking problems, further, are not problems separate and distinct from all other problems, but are one category of curriculum problems, communications problems, and public relations problems. Because each secondary school or school district is peculiar with regard to its setting, its curriculum, its students, its faculty--its whole school personality--the improvement of marking problems must be accomplished at the immediate source. While broad principles and suggestions may be applied, no detached solution is possible.

Since 1947, the date of Wrinkle's book, there has been nothing really significant suggested on the topic of improving marking practices. Wrinkle provided us with the truism related above--a condition not entirely understood prior to that time--and with a set of guidelines for designing a grading system in any particular school.
Hundreds of articles have been written about grading systems since, and undoubtedly some have been informative and helpful, some even creative. But none has had a substantial impact such as did Starch's articles in the years following 1912-1913.

Starch made the statement in 1918 that "The current movement for measuring school products is one of the three or four most important fields of investigation in the scientific study of educational problems" (15.,1). At that time, he was able to add "Very material progress has been made in the past half dozen years..." And not since 1918 would a similar statement again have had a basis in fact. There seems no reason to assume that "measuring school products" is any less important today than it was fifty years ago. Yet, as one considers the several various fields of education,—curriculum, administration, learning theory—dozens of books have been written about each, and there are dozens of highly knowledgeable persons specifically associated with each. In the broad field of educational evaluation, it must be conceded there are several excellent works, but these generally are concerned with standardized tests, sophisticated testing procedures, and statistical treatment of test results. There is no book, save Wrinkle's 120-page treatise of 1947, on the topic of secondary school grading systems and teacher marking practices. And even that book was not limited to the secondary schools.

At this point one wonders what the nature of the next important change in grading practices will be. In the nineteenth
century, breakthroughs came in the form of symbols as marking devices, written examinations, and the hundred-point grading scale. In the early twentieth century, the five-point grading scale with its ABCDF symbols, and the normal distribution curve represented substantial progress. This investigator prefers not to attempt to specify what the next significant advance may be. However, he does feel reasonably certain he can identify two factors, and possibly a third, that will have an influence in determining the nature of that change. The first two factors are the direction of curriculum change and modern principles of psychology. The third is automation.
CHAPTER III

DESIGN OF THE STUDY

Key to the Design

The purpose of this study was to investigate the dynamics of change in marking practices and grading systems in secondary schools. In order to investigate this area, it was necessary to identify schools in which recent changes in marking practices had taken place, and, for the sake of comparison, schools in which no such change had occurred.

It would have been possible to survey on a random basis any number of schools in order to determine which had or had not made such changes. This investigator, however, attempted to be more discriminating in determining what schools might be appropriately included in the study by making the assumption that the change in marking practices and grading systems is one dimension of educational innovation. Innovativeness and lack of innovativeness provided the key, then, to identifying those schools which were more likely to have updated grading practices as opposed to those schools more likely to exhibit complacency or resistance in adopting changes in traditional marking practices.
Selection of Schools Included in
the Study Group

In order to identify those schools which may acceptably be considered to be innovative as opposed to those schools which are non-innovative, this investigator used the findings of a study done in 1963-1964 under the direction of Daniel Stufflebeam at The Ohio State University (51.). Stufflebeam surveyed all the school districts in Ohio to determine which were using innovative practices. A brief description of that study, including the research procedures he used, his criteria for judging innovativeness, and those findings pertinent to the present study, is attached as Appendix A of this study.

The source of the innovations listed in Stufflebeam's study was identified only as a certain school district rather than individual schools. Since the majority of school districts include only one high school, there was no problem identifying the appropriate high school in those districts. But in the large city school districts, where there are many high schools, it was impossible to determine in which schools the innovative programs were most prevalent. The investigator, therefore, visited six of the larger cities of Ohio which reported extensive innovative practices in Stufflebeam's study, and interviewed either the assistant superintendent or director of secondary education in each. These men described the grading systems in their cities, and specified which two or three schools would most appropriately fit the purposes of this study.
Stufflebeam's study was designed to identify innovative practices and innovative school districts. The screening panel used in that study rated each program reported from each school district on a six-point scale, the lowest or sixth level of which was described as "not innovative and clearly not worthy of further study." Therefore, any school district that was not awarded a rating above the sixth level for a reported curriculum innovation on the secondary level was considered to be non-innovative and any high school in the district was considered to be a non-innovative high school for the purpose of this study.

After all the non-innovative schools had been identified, the investigator visited the Ohio State Department of Elementary and Secondary Education in order to verify that the schools identified as non-innovative through Stufflebeam's study were in fact non-innovative, and to establish which 40 of the hundred or so schools listed would, in the opinion of experts there, best suit the purposes of this study. Also, in the case of the one non-innovative district that included more than one high school, personnel from the Ohio State Department of Education helped in selecting the specific schools to be included in this study.

Through the procedures described, the investigator chose a total of 80 high schools to be included in the total study group. The study group included two sub-groups of 40 schools each, one innovative, the other non-innovative. All schools in the study group included the upper three senior high grade levels, although many of the schools included the junior high grade levels as well.
The Design and Transmittal of the Questionnaire and Cover Letter

The significance that a study such as this one might have is as dependent on the careful design of the questionnaire as upon any other single factor. With this in mind, the investigator constructed the questionnaire very deliberately. A reproduction of the questionnaire is included as Appendix B of this study.

The statement of the problem in the study proposal included thirteen hypotheses to be tested. In order to approach the problem of testing these hypotheses most directly, the questionnaire was set up in fourteen parts, each of the first thirteen corresponding to one of the hypotheses, with the fourteenth as a final free response. Each of the first thirteen parts included only statements to be checked by the respondent in preference to requiring him to write out a response.

The nature of the several parts of the questionnaire is such that each item could not be structured in an exactly similar manner to every other. Some are designed to elicit a yes-no response, others require a ranking, and still others request that the respondent check appropriate statements. Insofar as was possible, a consistency was maintained where there were similar items.

An effort was made to make the statements throughout the questionnaire as simple, clear, and direct as possible. Two high school principals read the final draft of the questionnaire and neither could find objection with nor suggest further revision to the instrument.
The final copy of the questionnaire was typed on an executive-style electric typewriter in order to produce the clearest and handsomest instrument short of actual printing. Finally, the copies were run on heavy, high-quality mimeograph paper.

It was believed by the investigator that an effective cover letter would be an important factor in eliciting a large response to the questionnaire. The cover letter, therefore, was given considerable attention. It included an opening paragraph designed to arouse the respondent's interest, a middle section describing the study and requesting copies of grading policies and report cards, and a final statement promising the respondent an abstract of the results and appealing to his professional conscience. As with the questionnaire, an executive-style typewriter was used. The letter was run on mimeograph paper on which was imprinted an Ohio State University College of Education letterhead. The investigator signed each letter personally, and Dr. Hugh Laughlin's signature was placed on the mimeograph stencil. The cover letter is attached as Appendix C to the study.

Each questionnaire and cover letter were mailed in a large brown manila envelope. Included in the envelope was a second stamped self-addressed manila envelope to be used to return the questionnaire and other requested materials.

The questionnaires were sent to the principal of each school in the total study group. It was felt that these men in their focal positions would be better able to answer questions related to changes
in marking practices and grading systems in their individual schools than any other particular group; also the selection of respondents was restricted to persons having the position of principal in order to make the responses from each sub-group of schools comparable.

Analysis of Responses to the Questionnaire

A large part of the data obtained from the questionnaire in this study are in the form of proportions indicating the frequency with which a particular characteristic occurs within each of two sub-groups of secondary schools, one sub-group considered to consist of innovative schools, the other, non-innovative. The proportions were then compared to determine if there appeared to be a statistical difference between the two sub-groups with respect to that characteristic.

The test used in this study was taken from Wallis and Roberts' Statistics--A New Approach (16). It is designed to compare two sample proportions. The formula for the test is

$$K = \frac{(bc - ad - \frac{n_1 \neq n_2}{2})}{\sqrt{n_1n_2 (a+b)(c+d)}}$$

where $K$ is the standard normal variable measured in units of standard deviation, and the other symbols have meanings as shown in Table 1.
TABLE 1

EXPLANATION OF THE MEANING OF THE SYMBOLS IN THE K FORMULA

<table>
<thead>
<tr>
<th>Sample</th>
<th>Occurrences</th>
<th>Non-Occurrences</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>a</td>
<td>b</td>
<td>n₁</td>
</tr>
<tr>
<td>Second</td>
<td>c</td>
<td>d</td>
<td>n₂</td>
</tr>
<tr>
<td>Total</td>
<td>a ≠ c</td>
<td>b ≠ d</td>
<td>n₁ ≠ n₂</td>
</tr>
</tbody>
</table>

For any value of K, a corresponding probability may be obtained which represents the frequency with which the standard normal variable is exceeded.

The null hypothesis has been employed in this study as the basis for comparing the two sub-groups. The null hypothesis states that there is no difference between the two sub-groups of schools with respect to the characteristics being tested. There are two possible alternatives to this. One is that the characteristic appears in the innovative sub-group significantly more often than in the non-innovative sub-group. The second alternative is the opposite of the first.

The test of two sample proportions described above is a "one-tail" test. This test is appropriate in this study even though there are two alternative hypotheses because one of the alternatives is immediately precluded by the way the data are used in the formula. In each test, the larger proportion, regardless of the sub-group from which it is derived, is set down as the second sample. The test
then indicates whether the proportion from the second sample is significantly larger than the proportion from the first.

The value for K in any given instance of the test may fall on either side of the midpoint of the normal curve depending upon which proportion is larger. It is an arbitrary matter to decide which tail represents which sub-group. In this study, a larger proportion from the innovative sub-group results in a positive K value in the upper tail, and a larger proportion from the non-innovative sub-group results in a negative K value in the lower tail.

When testing a null hypothesis, there is a possibility of two types of error, that of rejecting the null hypothesis when it is true, and that of accepting the null hypothesis when it is not true. In no statistical test involving sampling is there absolute certainty; risk of error in sampling is always present. In some studies risk of such error must be kept to a minimum because of the critical nature of the risk. In this study, no such critical condition exists. The investigator, therefore, has chosen the .10 level of confidence as being the most appropriate for interpreting the data he has obtained.

The K value is a measure of the amount of variation, or difference, in the proportions being compared. The larger the variation, the greater the difference. Where one finds little or no difference between the two proportions, the null hypothesis is considered to describe best the situation. When the difference is large, the null hypothesis is rejected, and one or the other of the alternative hypotheses is accepted. By utilizing the .10 confidence
level and converting to the K-scale the investigator is able to
describe the critical zones in terms of K values. At the midpoint
of the normal curve where there is an equal, or .50, probability of
a characteristic falling into either sub-group, the K value is 0.0.
The .10 level of confidence provides a critical zone on the K-scale
from -.253 to .253 in which the K value supports acceptance of the
null hypothesis. Rejection of the null hypothesis and acceptance of
whichever is the appropriate alternative hypothesis occurs in the
critical zones on the K-scale extending below -1.282 in the lower tail
and above 1.282 in the upper tail. The non-critical zones, then,
would include -.254 to -1.281 and .254 to 1.281.

Data from Ohio State Department
of Education Publications

In order to obtain a clearer picture of the schools and
school districts included in the study group, the investigator went
to various publications of the Ohio State Department of Education.
The number of pupils in each school was obtained from the 1964-1965
Educational Directory. Property value per pupil, instructional cost
per pupil, and total expenditure per pupil were obtained from the
1964-1965 Costs Per Pupil in Average Daily Membership in Ohio's City,
Exempted Village and County School Districts. Average secondary
teachers' and secondary principals' salaries in each school district
were obtained from the 1964-1965 Administrative, Instructional and
Non-Instructional Salaries in Ohio. These data provided certain
interesting comparisons in their own right, and they were also used
to help the investigator select the ten schools whose principals he visited.

Interviews with Principals

It was the contention of the investigator that while the questionnaires would provide factual information that could be statistically tested, the dynamics of change in grading systems includes a human relations dimension that lies beyond statistical analysis. In order to include this dimension in the study, and to supplement the skeleton of information available from the questionnaire with a more thorough understanding of the relationships, the influences, and the subtleties involved, he visited ten of the principals who had returned completed questionnaires in order to question them in depth.

The one firm criterion that was adhered to in choosing the schools to visit was that the ten schools visited be divided evenly between the innovative and non-innovative sub-groups. At the time the study was proposed, it was considered that the matching of schools in as many characteristics as possible between the two sub-groups would provide interesting comparisons, but the extent of the differences between the schools in each sub-group made such matching unfeasible. Instead, the investigator attempted to choose as many different kinds of schools within each sub-group as possible in order that the greatest number of different sorts of schools would be represented. The factors used to differentiate schools within each of the subgroups were type and location of the community, size of the school, expenditure per pupil, and location of the school.
The personal interview with each of the principals was semi-structured. In each tape-recorded session, the investigator encouraged the principal to comment freely on the nature and climate of the community, the characteristics of the school, the student body, and the staff, and his impressions and perceptions on a variety of topics including students' course marks. These interviews were intended to answer two questions: (1) Why do changes occur more readily in some schools than in others? and (2) How are changes in marking practices implemented? The investigator attempted to answer the first of these questions through being sensitive to significant differences in conditions and attitudes between schools in each of the sub-groups. Those five schools in the innovative sub-group were counted upon heavily to provide some answers to the second question. The investigator was interested, for instance, in changes in marking practices that are externally initiated as well as those internally initiated, and teacher initiated changes as well as those administratively initiated.

Summary

In this chapter there have been described the procedures used to investigate the dynamics of change in marking practices and grading systems in secondary schools. The investigator made the assumption that innovative schools would tend to exhibit change in marking practices more readily than non-innovative schools, and that comparing the two different types of schools would provide
significant contrasts. A study group of 80 schools was selected using Stufflebeam's Ohio Education Innovations Survey. The total study group included two sub-groups of 40 schools each, one innovative, the other non-innovative.

A questionnaire was prepared and sent to the principal of each of the 80 schools. The questionnaire was designed to accentuate the contrast between the two sub-groups.

Supplementary information was obtained by the investigator in three ways. Information relating to size and financial condition was obtained from Ohio State Department of Education publications. General information about grading practices and policies was available in marking-related materials submitted by principals. Finally, the personal insights and perceptions of the principals were obtained in ten tape-recorded interviews.
CHAPTER IV

PRESENTATION AND DISCUSSION OF RESPONSES TO THE QUESTIONNAIRE

Procedure for Analyzing Responses to the Questionnaire

The questionnaire consisted of fourteen sections, with each of the first thirteen designed to elicit information relating to a corresponding hypothesis. The final section was an open-ended question included for the purpose of obtaining suggestions for possible means of changing grading practices.

Each of the thirteen hypotheses is listed on the following pages, accompanied by the corresponding section from the questionnaire. For each of the sections, the responses are listed in tabular form including statistical analysis where this is appropriate. Also included is sufficient explanation to clarify the basis for the investigator's decisions regarding the acceptance or rejection of each hypothesis.

The investigator's decisions to accept or reject the hypotheses were based on the comparison of two sample proportions from which the value of the standard normal variable (K) was derived. The .10 confidence level was utilized to set up ranges or zones of K-values as follows: -.253 to .253—accept the null hypothesis; -1.282 or less—reject the null hypothesis and accept the alternative
hypothesis that there are significantly more occurrences of the characteristic in the non-innovative schools than in the innovative schools; \( \neq 1.253 \) or more—reject the null hypothesis and accept the alternative hypothesis that there are significantly more occurrences of the characteristic in the innovative schools than in the non-innovative schools; \(-0.254 \) to \(-1.281\) and \(0.254 \) to \(1.281\)—non-critical ranges in which no decision is made.

Section fourteen on the questionnaire does not have a corresponding hypothesis as does each of the other sections. Responses to it were free statements written by the respondents. The investigator has attempted to summarize the statements into categories for the purpose of presenting an overall picture of principals' suggestions for designing and implementing change in marking practices. Also, certain of the more pertinent statements have been selected to be included verbatim in this study.

Analysis of Data from the Questionnaire

Hypothesis No. 1.—There will be no significant difference between innovative and non-innovative schools in the use of unusual marking practices.

The statement in Section I in the questionnaire is as follows:

Indicate which, if any, of the following marking practices are employed in your school.

The responses to Section I from all the questionnaires are summarized in Table 1. Considered as a whole, the null hypothesis
<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Number of Responses from 33 Innovative Schools</th>
<th>Number of Responses from 33 Non-Innovative Schools</th>
<th>Normal Variable (K Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Dual grades are included on reports to parents, e.g., a grade for competitive achievement and a grade for achievement based on ability.</td>
<td>4</td>
<td>1</td>
<td>.931</td>
</tr>
<tr>
<td>2.</td>
<td>An unusual set of symbols, or an unusual arrangement of symbols, is used, e.g., subscripts ((B_d; B_2)) or fractions ((B)).</td>
<td>8</td>
<td>3</td>
<td>1.320</td>
</tr>
<tr>
<td>3.</td>
<td>The report to parents includes a graph representing a profile of student achievement.</td>
<td>0</td>
<td>0</td>
<td>---</td>
</tr>
<tr>
<td>4.</td>
<td>Different symbols are used in academic and non-academic courses, e.g., use of ABCDF in an academic course and S-U in a non-academic.</td>
<td>3</td>
<td>5</td>
<td>-.378</td>
</tr>
</tbody>
</table>
TABLE 2—Continued

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Number of Responses from 33 Innovative Schools</th>
<th>Number of Responses from 33 Non-Innovative Schools</th>
<th>Normal Variable (K Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>In addition to a symbol grade (letter or number) the report includes a statement written by the teacher regarding student achievement or character.</td>
<td>7</td>
<td>5</td>
<td>.319</td>
</tr>
<tr>
<td>6.</td>
<td>In addition to a symbol grade (letter or number) the report includes a series of statements regarding student achievement or character to be checked by the teacher.</td>
<td>9</td>
<td>14</td>
<td>-1.033</td>
</tr>
<tr>
<td>7.</td>
<td>A written statement or checklist unaccompanied by any symbol grades constitutes the formal report made to parents.</td>
<td>0</td>
<td>0</td>
<td>---</td>
</tr>
<tr>
<td>8.</td>
<td>Self-evaluation by pupils is a factor in determining the course mark.</td>
<td>0</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>9.</td>
<td>Parent-teacher conferences are an integral or necessary part of the grading system.</td>
<td>5</td>
<td>7</td>
<td>-.319</td>
</tr>
</tbody>
</table>
Table 2—Continued

<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Number of Responses from 33 Innovative Schools</th>
<th>Number of Responses from 33 Non-Innovative Schools</th>
<th>Normal Variable (K Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.</td>
<td>The only grade on the report to parents is simply pass-fail, or satisfactory-unsatisfactory.</td>
<td>2</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td>11.</td>
<td>Supplementary progress reports or warnings of possible failure are sent to parents at any time that such a report is deemed appropriate.</td>
<td>27</td>
<td>26</td>
<td>0.0</td>
</tr>
<tr>
<td>12.</td>
<td>One or more of the following types of standardized tests are employed as a basis or factor in providing information appearing on reports to parents:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>achievement</td>
<td>11</td>
<td>10</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>intelligence</td>
<td>6</td>
<td>11</td>
<td>-1.126</td>
</tr>
<tr>
<td></td>
<td>aptitude</td>
<td>4</td>
<td>8</td>
<td>- .957</td>
</tr>
<tr>
<td></td>
<td>personality inventory</td>
<td>2</td>
<td>4</td>
<td>- .428</td>
</tr>
<tr>
<td>Number</td>
<td>Item</td>
<td>Number of Responses from 33 Innovative Schools</td>
<td>Number of Responses from 33 Non-Innovative Schools</td>
<td>Normal Variable (K Value)</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>13.</td>
<td>A grading period other than the usual one of six weeks duration is used as follows:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>monthly</td>
<td>0</td>
<td>0</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>9 weeks</td>
<td>3</td>
<td>3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>semester</td>
<td>3</td>
<td>2</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>irregular</td>
<td>0</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>reports to parents are on a rotating basis, rather than to all parents at once</td>
<td>0</td>
<td>0</td>
<td>---</td>
</tr>
<tr>
<td>14.</td>
<td>Marking practice not listed above other than the common ABCDF or a scale based on 100.</td>
<td>4</td>
<td>0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
appears to be supported by these data. As one examines the responses to the first eleven items, a simple count shows that from the 33 schools in each sub-group there were 65 responses from innovative schools and 64 responses from non-innovative schools. Also from the first eleven items, there are four in which innovative schools exceed non-innovative schools, four in which non-innovative schools exceed innovative schools, and three in which the responses are identical. In each of the sets of four, there is an instance of the difference being only one response. For only one item does the innovative sub-group exceed the non-innovative sub-group by an amount which lies within the .10 level of confidence. In no instances does a non-innovative school exceed an innovative school by that amount.

In Item 12, dealing with using standardized tests in the grading procedure, non-innovative schools are shown to employ the practice somewhat more often than innovative schools do. For each of the latter three types of test—intelligence, aptitude, and personality inventory—the non-innovative schools are found to employ them about twice as frequently as innovative schools. Nevertheless, the overall use of these tests is sufficiently infrequent that the difference between the two sub-groups never reaches the .10 level of significance.

In Item 13, which deals with grading periods, the response to each choice is small, and the difference in the number of responses from each sub-group of schools for the choices involved is never greater than one.

---

3 See page 123 for the investigator's commentary on this item.
In Item 14 principals were requested to describe any practices not included in the list preceding it. Two principals mentioned using a five-point scale to evaluate grades in advanced classes. A third principal referred to a special report form used for pre-vocational students. The fourth principal listed the pass on effort grade and the conditional grade, both described in Chapter V of this study. This last principal also mentioned that absence was reported to parents by class period rather than by day.

Hypothesis No. 2.—There will be no significant difference between innovative and non-innovative schools in the use of differentiation of grades in enriched, advanced placement, accelerated, or honors courses or classes.

There are two parts to the question designed to test this hypothesis. This is the first part of Section II from the questionnaire.

Does the curriculum in your school include enriched, advanced placement, accelerated, or honors courses or classes for the very able students (in addition to the regular college preparatory courses)?

**TABLE 3**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Innovative</td>
<td>10</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>Innovative</td>
<td>30</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

\[K = 4.683\]
The data in Table 3 provide a K value which indicates overwhelmingly that there is a significant difference between innovative and non-innovative schools in the frequency with which enriched, advanced placement, accelerated, or honors courses or classes are offered. Thirty innovative schools indicated they did so, while only ten non-innovative schools made a similar response. These two values, then, become the base for determining the extent to which differentiation in grading has occurred in the schools of each sub-group that offer these special courses.

This is the second part of Section II from the questionnaire.

If you do have enriched, advanced placement, accelerated, or honors courses, do you in some way differentiate grades (e.g., weighting of grades or use of different symbols) in these courses as opposed to the regular courses?

TABLE 4

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Innovative</td>
<td>4</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>Innovative</td>
<td>19</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

K = .925

Table 4 includes the responses from those schools which offer the specified types of courses, and indicates the number of these schools that employ differentiation in grading in these courses. The K value obtained (.925) when the two proportions are compared falls
in the non-critical zone of the upper tail. Expressed as a percent, the comparison is 63 percent of innovative schools that differentiate grading in able student courses to 40 percent of non-innovative schools. Evidence from the data is insufficient to accept or reject the null hypothesis. At best, the results indicate a weak tendency in favor of rejection of the null hypothesis.

Hypothesis No. 3.--There will be no significant difference between innovative and non-innovative schools in the use of differentiation of grades in slow-learner or remedial classes.

There are two parts to the question designed to test this hypothesis. This is the first part of Section III from the questionnaire.

Does the curriculum in your school include slow-learner or remedial classes or courses for students unable to meet the more rigorous standards in regular classes?

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Innovative</td>
<td>19</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Innovative</td>
<td>30</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ K = 2,816 \]

The situation in Table 5 is similar to that in Table 3. The proportion of innovative schools offering remedial or slow-learner courses or classes is significantly higher than the corresponding proportion of non-innovative schools as indicated by the extremely
high K value obtained (2.816). Thirty innovative schools and 19 non-innovative schools reported offering these courses.

This is the second part of Section III from the questionnaire.

If you do have slow-learner or remedial classes or courses, do you in some way differentiate grades in these courses as opposed to the regular courses?

TABLE 6

THE NUMBER OF SCHOOLS IN EACH SUB-GROUP REPORTING THE USE OF DIFFERENTIATION OF GRADES IN SPECIAL COURSES FOR SLOW-LEARNERS

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovative</td>
<td>12</td>
<td>17</td>
<td>4</td>
</tr>
<tr>
<td>Non-Innovative</td>
<td>8</td>
<td>11</td>
<td>14</td>
</tr>
</tbody>
</table>

K = 0.0

In Table 6, which uses as a basis for comparison 30 innovative schools (one of which did not answer the second part of Section III) and 19 non-innovative schools, the K value is 0.0, indicating that there is no statistically significant difference between the two sub-groups. The non-innovative sub-group holds a slight percent edge, 42 percent to 41 percent, over the innovative sub-group. The null hypothesis, stating that there is no difference in innovative and non-innovative schools in the use of differentiation of grades in slow-learner classes, must be accepted.

Hypothesis No. 4.—There will be no significant difference between innovative and non-innovative schools in the number of schools with a suggested philosophy of grading.
Section IV from the questionnaire was stated as follows:

Does your school have a written philosophy pertaining specifically to grading practices that teachers are asked to consider in assigning marks?

**TABLE 7**

NUMBER OF SCHOOLS FROM EACH SUB-GROUP REPORTING THE OCCURRENCE OR NON-OCCURRENCE OF A SUGGESTED PHILOSOPHY OF GRADING

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovative</td>
<td>20</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Non-Innovative</td>
<td>21</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

**K = 0.0**

The data included in Table 7 are seen to present a situation similar to those in Table 6. The proportion of schools in each sub-group reporting having a written philosophy is so similar that there is no statistically significant difference. The K value is 0.0 and the probability of the occurrence of the characteristic in either group is .50. The null hypothesis, therefore, must be accepted.

Hypothesis No. 5.—There will be no significant difference between innovative and non-innovative schools in the number of schools having a written policy on grading which stipulates that teachers shall subscribe to certain standards and/or employ certain practices.

Section V from the questionnaire was stated as follows:

Does your school have a written statement of its policy on grading which stipulates that teachers shall subscribe to certain standards and/or employ certain practices (other than
simply a brief statement indicating that teachers will use a certain common set of symbols)?

**TABLE 8**

**NUMBER OF SCHOOLS FROM EACH SUB-GROUP REPORTING THE OCCURRENCE OR NON-OCCURRENCE OF A WRITTEN POLICY ON GRADING**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Innovative</td>
<td>16</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>Innovative</td>
<td>19</td>
<td>12</td>
<td>2</td>
</tr>
</tbody>
</table>

**K = .218**

From Table 8 it is seen that 29 non-innovative schools responded, 16 affirmatively, and 31 innovative schools responded, 19 affirmatively. Expressed as percents, 61.5 percent of non-innovative schools exhibit the characteristic as compared to 61.2 percent of innovative schools. As in Tables 6 and 7, the K value is in the mid-range, and there is no statistically significant difference between the two proportions. The null hypothesis must be accepted.

Hypothesis No. 6.--There will be no significant difference between innovative and non-innovative schools in bases employed for determining students' grades.

Section VI from the questionnaire was stated as follows:

Which of the following is the basis used by most of the teachers in your school to determine students' grades?
<table>
<thead>
<tr>
<th>Number</th>
<th>Item</th>
<th>Number of Responses from 33 Innovative Schools</th>
<th>Number of Responses from 33 Non-Innovative Schools</th>
<th>Normal Variable (K Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>A competitive basis which compares each student's achievement with the average achievement of the class as a whole.</td>
<td>21</td>
<td>19</td>
<td>$f_{.252}$</td>
</tr>
<tr>
<td>2.</td>
<td>Comparison of student's achievement with a standard selected or established by the teacher.</td>
<td>17</td>
<td>17</td>
<td>0.0</td>
</tr>
<tr>
<td>3.</td>
<td>Comparison of a student's achievement with his ability to achieve.</td>
<td>13</td>
<td>9</td>
<td>$f_{.783}$</td>
</tr>
</tbody>
</table>
The data obtained in response to Section VI are in rather close agreement between the two sub-groups. For the first two items in the table, the K values obtained, \( \phi 252 \) and 0.0, respectively, indicate that the null hypothesis should be accepted. The K value obtained for the third item, \( \phi 783 \), lies near the center of the non-critical range, and shows no clear tendency for either acceptance or rejection. On the whole, the data provide a stronger argument for accepting the null hypothesis than rejecting it, even though the third item is not so clear-cut as the first two, particularly the second.

Hypothesis No. 7.--There will be no significant difference between innovative and non-innovative schools in what is considered to be the primary function of course marks.

Section VII as stated from the questionnaire is as follows:

Rank with the numbers 1, 2, 3, 4, and 5 the order of importance of the following functions of the grading system in your school.

TABLE 10

RANK ASSIGNED TO EACH OF FIVE POSSIBLE FUNCTIONS OF GRADES BY THE SCHOOLS IN EACH SUB-GROUP

<table>
<thead>
<tr>
<th>Item</th>
<th>Rank Assigned by 33 Innovative Schools</th>
<th>Rank Assigned by 31 Non-Innovative Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inform the student and parents of the quality of the students' work</td>
<td>27 2 4 0 0</td>
<td>26 1 3 1 0</td>
</tr>
<tr>
<td>Stimulate the student to greater effort in his learning activities</td>
<td>6 7 10 10 0</td>
<td>1 16 11 3 0</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Item</th>
<th>Rank Assigned by 33 Innovative Schools</th>
<th>Rank Assigned by 31 Non-Innovative Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 13 11 8 0</td>
<td>3 10 10 8 0</td>
</tr>
<tr>
<td>Diagnose student strengths and weaknesses in order to provide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>for more informed counseling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 10 8 14 0</td>
<td>1 4 8 17 1</td>
</tr>
<tr>
<td>Provide a record of the student's achievement useful to colleges</td>
<td></td>
<td></td>
</tr>
<tr>
<td>or employers, transfer to other high schools, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0 0 0 0 33</td>
<td>0 0 0 1 30</td>
</tr>
<tr>
<td>A disciplinary measure in cases of misbehavior, or as a deterrent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>to misbehavior</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The response to Section VII as recorded in Table 10 shows that 26 of 31 non-innovative schools, or 84 percent, rank the first mentioned item as the primary function of course marks, while 27 of 33 innovative schools, or 82 percent, do the same. The two ratios are so nearly equal that the K value resulting from their comparison is 0.0. This indicates that there is no significant statistical difference in the two ratios and the null hypothesis is accepted.

The other values included in the table can be seen by inspection to indicate a relatively close relationship between the two sub-groups insofar as the rating of the importance of the various functions of grades is concerned. Most significant of these other values are the 33 and 30 schools from the innovative and non-innovative
groups, respectively, that rated the disciplinary function of grades as lowest. The one principal who did not rate it lowest rated it next-to-lowest. There is, then, close agreement in each sub-group on what are the most and the least important functions of grades.

Hypothesis No. 8.—There will be no significant difference between innovative and non-innovative schools in the position (i.e., area of responsibility) of the person who provided the impetus for change in the marking practices or the grading system.

There are two parts to the question designed to test this hypothesis. This is the first part of Section VIII of the questionnaire.

Has there occurred in your school during your tenure as principal, or since 1958, any change in your marking practices?

TABLE 11
NUMBER OF SCHOOLS FROM EACH SUB-GROUP REPORTING THE OCCURRENCE OR NON-OCCURRENCE OF A RECENT CHANGE IN MARKING PRACTICES

<table>
<thead>
<tr>
<th>Sample</th>
<th>Yes</th>
<th>No</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Innovative</td>
<td>19</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Innovative</td>
<td>28</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

K = 1.371

The first part of this question is a critical part of the study. An assumption was made by the investigator when choosing his study group that innovative schools would be more likely to display changes in marking practices than would non-innovative schools. Twenty-eight of the 33 innovative schools reported a change, while
19 of the 33 non-innovative schools reported a change. A comparison of these two proportions results in a K value of 1.371. This value lies within the .10 confidence level, and the investigator can, therefore, assume that there is a significant difference between innovative and non-innovative schools with respect to the frequency of recent change in marking practices.

The first part of this question also identified the schools in each sub-group that had changed marking practices since 1958, or within the tenure of the present principal. This information was important in the second part of Section VIII, and in the several succeeding questions, for these were the schools which could describe how the changes had been carried out.

This is the second part of Section VIII as stated in the questionnaire.

If there has been a change in your marking practices, who provided the initiative for this change?

Table 12 lists the number of responses from 28 innovative schools and 19 non-innovative schools as they indicated the persons responsible for initiating change in marking practices. K values resulting from the comparison of corresponding proportions from each sub-group are listed. Only in the case of the Director of Curriculum does a significant difference occur. This difference in favor of the innovative schools is probably due to the fact that the innovative schools are more frequently situated in large school districts in which the position of Director of Curriculum or Director
### TABLE 12

**PERSONS OR GROUPS INITIATING CHANGE IN EACH SUB-GROUP OF SCHOOLS**

<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools Reporting Marking Changes (28)</th>
<th>Non-Innovative Schools Reporting Marking Changes (19)</th>
<th>All Schools Reporting Marking Changes (47)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Distribution Responses</td>
<td>Number of Distribution Responses</td>
<td>Number of Distribution Responses</td>
</tr>
<tr>
<td></td>
<td>Percent Responses of Responses</td>
<td>Percent Responses of Responses</td>
<td>Percent Distribution Responses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Normal Variable (K Value)</td>
<td></td>
</tr>
<tr>
<td>School Board</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Supt. or Ass't.</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Supt.</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Dir. of Curric. or Dir. of Sec. Ed.</td>
<td>13</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Principal</td>
<td>7</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td>Principal</td>
<td>8</td>
<td>33</td>
<td>9</td>
</tr>
<tr>
<td>Counselor</td>
<td>6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Curriculum Consultant</td>
<td>2</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Teacher Comm., Tchr. Wksp. Group</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Parent-Teacher Comm., Group, or Org.</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Parent Group</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Student Group or Body as a Whole</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>39</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>101</td>
<td>100</td>
</tr>
</tbody>
</table>
of Secondary Education is more likely to exist. The percent distribution of responses as seen in the table shows a relatively close agreement between the two groups of schools.

The last two columns in the table include the total responses from all schools, and the overall percent distribution. The principal, with 33 percent of the responses, is seen to be the most frequent initiator of change in marking practices, with teacher committees or workshop groups, at 31 percent, running a close second. The superintendent, with 14 percent, is also seen to be fairly active in this area. The Director of Curriculum, where that position exists, appears to play a reasonably active role in revising marking practices. Neither the board of education, nor the counselors, nor the curriculum consultant, nor the students are particularly effective agents of change in marking practices.

Hypothesis No. 9.—There will be no significant difference between innovative and non-innovative schools in the stated reasons for making a change in marking practices.

Section IX from the questionnaire was stated as follows:

If a change in marking practices has occurred in your school, why was this change made?

Table 13 lists four specific reasons for changing marking practices. The first item, dealing with variability in grades, elicited the smallest response of the four, but the K value derived from a comparison of corresponding proportions indicates that non-innovative schools mention this reason significantly more often.
<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools Reporting Marking Changes (28)</th>
<th>Non-Innovative Schools Reporting Marking Changes (19)</th>
<th>All Schools Reporting Marking Changes (47)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent Distribution of Responses</td>
<td>Percent Distribution of Responses</td>
<td>Percent Distribution of Responses</td>
</tr>
<tr>
<td>Reduce variability in grades</td>
<td>2</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>37</td>
<td>19</td>
</tr>
<tr>
<td>Increase objectivity of grades</td>
<td>7</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>Provide for differentiation of grades in enriched and/or slow learner classes</td>
<td>54</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>Bring marking practices into line with the objectives of the curriculum</td>
<td>14</td>
<td>50</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
<td>53</td>
</tr>
<tr>
<td>Other reasons</td>
<td>3</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>
The second item, concerning objectivity of grades, provides no such significant difference at the .10 confidence level between the two sub-groups, although there is a weak indication that non-innovative schools employ this reason more frequently.

The third item, relating to differentiation of grades in exceptional students' courses, is the most often mentioned by the innovative schools, and least mentioned by the non-innovative schools. There is a clear-cut significant difference favoring the innovative schools in this case. This is to be expected, for the innovative schools are much larger on the average, and include more extensive curriculum offerings. Exceptional students' courses are found far more frequently in them.

Item 4, which deals with aligning marking practices with curriculum objectives, was mentioned by about half the schools in both groups.

A total of five schools checked the "Other Reasons" choice. The reasons given were to adopt symbols suitable for use by data processing equipment, to align the grading system with that commonly used by other schools, to make the report to parents more meaningful, to adjust to a nine-week grading period, and to attempt to reduce the incidence of student failures.

In Items 1, 3, and 4, the derived K values are statistically significant but each falls within a different critical zone. For each considered as a single entity one may accept or reject the null
hypothesis, but when the items are considered as a group, no decision either to accept or to reject is possible, nor can a trend be detected.

Hypothesis No. 10.--There will be no significant difference between innovative and non-innovative schools in the source of the idea for the suggested changes in marking practices.

Section X from the questionnaire was stated as follows:

If a change in marking practices has occurred in your school, what was the source of the idea that suggested the change that was made?

As one observes the first five columns in Table 14, it is apparent that the two sub-groups of schools have responded in a rather similar manner. Comparison of six of the corresponding proportions from the two sub-groups results in a K value which is in the critical zone for acceptance of the null hypothesis. There are three instances of the K value falling within non-critical zones, and only one, that referring to in-service education programs, in which innovative schools includes a statistically significant larger proportion of occurrences. Outright acceptance of the null hypothesis is precluded by the four instances of K values falling outside the critical zone for acceptance, but qualified acceptance of the null hypothesis appears to be in order.

The last two columns represent the overall picture provided by all schools that have implemented changes in marking practices.

The suggestion for change in marking practices emanates 23 percent of the time from faculty committees, and only slightly less frequently from the principal. Ranging in the 8 to 10 percent range of distribution
<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools Reporting Marking Changes (28)</th>
<th>Non-Innovative Schools Reporting Marking Changes (19)</th>
<th>All Schools Reporting Marking Changes (47)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent Number of Distribution Responses of Responses</td>
<td>Percent Number of Distribution Responses of Responses</td>
<td>Normal Variable (K Value) Number of Responses Distribution of Responses</td>
</tr>
<tr>
<td>Professional journal or recent prof. book</td>
<td>4 7</td>
<td>1 3</td>
<td>.502 5 6</td>
</tr>
<tr>
<td>Faculty committee</td>
<td>12 22</td>
<td>8 24</td>
<td>0.0 20 23</td>
</tr>
<tr>
<td>Teachers' Wksp.</td>
<td>0 --</td>
<td>1 3</td>
<td>0.0 1 1</td>
</tr>
<tr>
<td>In-service education program</td>
<td>7 13</td>
<td>1 3</td>
<td>1.369 8 9</td>
</tr>
<tr>
<td>Obsv. of or infor. re a change made in another school system</td>
<td>6 12</td>
<td>3 9</td>
<td>.105 9 10</td>
</tr>
<tr>
<td>Cen. office staff</td>
<td>4 7</td>
<td>4 12</td>
<td>-.087 8 9</td>
</tr>
<tr>
<td>Consultant</td>
<td>2 4</td>
<td>0 --</td>
<td>.454 2 2</td>
</tr>
<tr>
<td>Principal</td>
<td>10 19</td>
<td>8 24</td>
<td>-.137 18 21</td>
</tr>
<tr>
<td>Counselor</td>
<td>3 6</td>
<td>2 6</td>
<td>0.0 5 6</td>
</tr>
<tr>
<td>Teacher</td>
<td>3 6</td>
<td>4 12</td>
<td>-.600 7 8</td>
</tr>
<tr>
<td>Other source</td>
<td>3 6</td>
<td>1 3</td>
<td>--- 4 5</td>
</tr>
<tr>
<td></td>
<td>54 100</td>
<td>33 99</td>
<td>--- 87 100</td>
</tr>
</tbody>
</table>
of response are observation of other schools, in-service education programs, central office staff, and teachers. The other sources, including counselors and professional journals at 6 percent each, and consultants and teacher workshops at 2 percent and 1 percent, respectively, have less impact as sources of ideas for change.

The four other sources mentioned by the responding schools were (1) the necessity for adopting symbols appropriate for data processing, (2) state and national principal's meetings, (3) a committee on pupil personnel policies, and (4) students.

Hypothesis No. 11.—There will be no significant difference between innovative and non-innovative schools in the procedures for implementing changes in marking practices.

Section XI from the questionnaire was stated as follows:

If a change in marking practices has occurred in your school, what means were used to implement this change?

From the first five columns of Table 15 it is seen that within the first three choices there is a large amount of variation, and in the latter three very little. In only one of the first three, that referring to the principal, is the K value sufficiently high to be considered statistically significant, and to support rejection of the null hypothesis. In each of the latter three choices, the K values fall within the critical range for acceptance of the null hypothesis. There is stronger evidence, overall, for acceptance than for rejection of the null hypothesis, but tentative rather than clear-cut acceptance is indicated in this case.
<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools Reporting Marking Changes (28)</th>
<th>Non-Innovative Schools Reporting Marking Changes (19)</th>
<th>All Schools Reporting Marking Changes (47)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent Number of Distribution Responses of Responses</td>
<td>Percent Number of Distribution Responses of Responses</td>
<td>Percent Number of Distribution Responses</td>
</tr>
<tr>
<td>Directive from the central office</td>
<td>10 21</td>
<td>4 13</td>
<td>14 18</td>
</tr>
<tr>
<td>Directive from principal's office</td>
<td>10 21</td>
<td>2 7</td>
<td>12 15</td>
</tr>
<tr>
<td>Vote by, or agreement by, the faculty</td>
<td>6 13</td>
<td>7 23</td>
<td>13 17</td>
</tr>
<tr>
<td>Study project within the school or school system</td>
<td>10 21</td>
<td>8 27</td>
<td>18 23</td>
</tr>
<tr>
<td>Meetings or workshops to inform or orient teachers</td>
<td>8 17</td>
<td>5 17</td>
<td>13 17</td>
</tr>
<tr>
<td>Public relations prog. to inform or orient parents</td>
<td>3 6</td>
<td>2 7</td>
<td>5 6</td>
</tr>
<tr>
<td>Other means</td>
<td>1 2</td>
<td>2 7</td>
<td>3 4</td>
</tr>
<tr>
<td></td>
<td>48 101</td>
<td>30 101</td>
<td>78 100</td>
</tr>
</tbody>
</table>
The last two columns in the table show that study projects are used 23 percent of the time to change marking practices, more by a fair margin than any other choice. Ranging from 15 to 18 percent are four choices including (1) directives from the central office, (2) vote by, or agreement by, the faculty, (3) teachers' workshops, and (4) directive from the principal's office. Public relations programs were used only 5 percent of the time.

In the category of "Other," three schools indicated they had used procedures not suggested in the first six choices. Examination of them, however, indicated that each of these may appropriately be included in the first category above.

Hypothesis No. 12.—There will be no significant difference between innovative and non-innovative schools in the persons who actively supported changes in marking practices or the grading system.

Section XII from the questionnaire was stated as follows:

If a change in marking practices has occurred in your school, which persons or groups do you feel were actively supportive of the change?

The first five columns in Table 16 comparing corresponding proportions from the two sub-groups give an overall picture of similarity between the sub-groups. Of the ten items in the table, four result in K values within the critical range for acceptance of the null hypothesis, five fall in the non-critical range, and only one falls in the critical range for rejection of the hypothesis. There is more evidence for acceptance of the null hypothesis than for
### TABLE 16

**PERSONS OR GROUPS SUPPORTIVE OF CHANGE IN MARKING PRACTICES AS INDICATED BY EACH SUB-GROUP OF SCHOOLS**

<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools</th>
<th>Non-Innovative Schools</th>
<th>All Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reporting Marking</td>
<td>Reporting Marking</td>
<td>Reporting</td>
</tr>
<tr>
<td></td>
<td>Changes (28)</td>
<td>Changes (19)</td>
<td>Changes (47)</td>
</tr>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Normal</td>
</tr>
<tr>
<td>Number of Distribution Responses of Responses</td>
<td></td>
<td>Normal Variable (K Value)</td>
<td>Number of Distribution Responses</td>
</tr>
<tr>
<td>Board of Education</td>
<td>11</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>Superintendent</td>
<td>21</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>Central office staff (Curric. Dir., etc.)</td>
<td>13</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Principal</td>
<td>25</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>Counselors</td>
<td>20</td>
<td>16</td>
<td>26</td>
</tr>
<tr>
<td>Teachers as a whole</td>
<td>22</td>
<td>18</td>
<td>35</td>
</tr>
<tr>
<td>A committee or certain group of teachers</td>
<td>4</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Parent-teacher group</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Parent group</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Students</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>123</strong></td>
<td><strong>73</strong></td>
<td><strong>196</strong></td>
</tr>
</tbody>
</table>
rejection, but because of the one item which supports rejection and
the five that provide no decisive evidence, acceptance is on a tenta-
tive basis.

The total of all persons or groups who supported change in
marking practices is represented in the last two columns. It is seen
there that principals, with 20 percent of the percent distribution,
are the most frequent supporters, and that the superintendent and
the teachers, with 18 percent each, both are quite active supporters.
A gap of 5 percent exists below the 18 percent level, but the counselors,
with 13 percent, the board of education, with 10 percent, and the
central office with 9 percent are all reasonably active. Less
active are teacher committees, students, parents, and parent-teacher
groups.

Hypothesis No. 13.—There will be no significant difference
between innovative and non-innovative schools with respect to the
persons who actively opposed changes in marking practices or the
grading system.

Section XIII from the questionnaire was stated as follows:

If a change in marking practices has occurred in
your school, what persons or groups resisted or
actively opposed the change?

The most striking characteristic indicated in Table 17 is
the large number of schools making no response. Apparently few
persons or groups resist change in marking practices. A second
characteristic noted in the table is that there is no statistically
significant difference between the two sub-groups. Therefore, the
null hypothesis is accepted without qualification,
<table>
<thead>
<tr>
<th>Item</th>
<th>Innovative Schools</th>
<th>Non-Innovative Schools</th>
<th>All Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reporting Marking</td>
<td>Reporting Marking</td>
<td>Reporting</td>
</tr>
<tr>
<td></td>
<td>Changes (28)</td>
<td>Changes (19)</td>
<td>Marking</td>
</tr>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Number</td>
</tr>
<tr>
<td></td>
<td>Number of</td>
<td>Number of</td>
<td>Distribution</td>
</tr>
<tr>
<td></td>
<td>Distribution</td>
<td>Distribution</td>
<td>of Responses</td>
</tr>
<tr>
<td></td>
<td>Percent</td>
<td>Normal Variable</td>
<td>Percent</td>
</tr>
<tr>
<td></td>
<td>Number of</td>
<td>Number of</td>
<td>Distribution</td>
</tr>
<tr>
<td></td>
<td>Responses</td>
<td>Responses</td>
<td>of Responses</td>
</tr>
<tr>
<td>Board of Education</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Superintendent</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Central office staff</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(Curric. Dir.,etc.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principal</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Counselors</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Teachers as a whole</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>A committee or certain group of teachers</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Parent-teacher group</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Parent group</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Students</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Section XIV from the questionnaire.—The statement of Section XIV as it appeared in the questionnaire is as follows:

Will you comment briefly on what changes you would like to make in the marking practices or grading system of your school? How would you implement such changes?

In the following two tables, the investigator has attempted to summarize the responses obtained, and to develop a "feeling" for the concerns and attitudes of principals regarding change in grading practices. The categories used were arbitrarily conceived to accommodate the responses.

**TABLE 18**

**CHANGES THAT PRINCIPALS WOULD LIKE TO EFFECT IN THEIR PRESENT MARKING PRACTICES**

<table>
<thead>
<tr>
<th>Description of the Change</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispense entirely with grades</td>
<td>8</td>
</tr>
<tr>
<td>Satisfied with present grading practices</td>
<td>7</td>
</tr>
<tr>
<td>Establish weighted grades for honors and/or slow-learner classes</td>
<td>6</td>
</tr>
<tr>
<td>Would like to make some change but do not know what</td>
<td>5</td>
</tr>
<tr>
<td>Transfer a student to different levels of the same course to indicate achievement</td>
<td>3</td>
</tr>
</tbody>
</table>

From Table 18 it is seen that eight principals would like to dispense with grades rather than continue to try to improve grading practices. Almost as many principals are satisfied with the grading systems they have. Actually, in analyzing the written comments of the principals, and in talking with some of them personally, the
investigator finds more similarity in the first two items than is readily apparent. The first group feels that grades do not do what they are designed to do, and the practice of grading should be discontinued. The second group admits grades are essential in the educative process we presently employ. Many feel, however, that change would be more for the sake of change rather than because any one grading system is demonstrably superior to all the others.

Only six principals mentioned anything about weighting grades. This was rather surprising to the investigator, but it is partly explained by the fact that a substantial number of the schools included in the study already use some form of weighting.

Five principals indicated they would like to implement an improved method of grading, but did not know what direction to take. Again, there is a similarity in the attitudes expressed here as in those represented in the first two items in the table. The suggestion to improve grades is frequently met with feelings of inadequacy, frustration, and resignation.

Three principals suggested that, as a substitute for the traditional assigning of marks, students might simply be shifted between different levels of the same course. This suggestion presumes a large high school in which as many as five levels in some subjects are available. It was also pointed out that the non-graded high school (no class distinctions) would be particularly appropriate for this purpose.
TABLE 19
MEANS WITH WHICH PRINCIPALS PROPOSE TO CHANGE MARKING PRACTICES

<table>
<thead>
<tr>
<th>Description of the Means</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal, with authorization from higher offices, should implement the change</td>
<td>15</td>
</tr>
<tr>
<td>Faculty committee should be the most influential change agent</td>
<td>13</td>
</tr>
<tr>
<td>The community should be involved in the change</td>
<td>4</td>
</tr>
<tr>
<td>City-wide study should be employed</td>
<td>1</td>
</tr>
<tr>
<td>Local professional association should be the most influential change agent</td>
<td>1</td>
</tr>
</tbody>
</table>

The questionnaire was answered by principals only. This fact presents a bias in the sort of results obtained in this study. The response shown for the first item in Table 19 possibly is influenced by that bias, for it indicates that the principals themselves should be the chief instigators or implementers of change in grading practices. Almost as many principals, however, designate a faculty committee as the most appropriate agent to carry out the change. Four principals felt the community would have to be prepared or its approval solicited if a change in grading practice were to meet with success.

The following are several statements from the principals, reproduced verbatim:

Move to ungraded classes. No idea how to bring about such a change in a large city system.
At the present time we have a committee of teachers and administrators studying the grading system: (1) to have a grading system that is the same 1-12; (2) time between grading reports to parents; (3) weight to be given accelerated courses.

Changes in grading systems would always grow out of teacher review and study. Our only concern with grading is that we may not be giving just consideration to the talented student who is in an enriched class. We do not presently weight this grade in computing class rank. We possibly should have a distinctive mark for honors work. We possibly should also have a mark to indicate a student has not achieved satisfactorily, but has expended maximum effort and deserves credit.

It would be a blessing if grades were not required by universities and parents. As we refine our marking system (school-wide) we would coordinate this with the Asst. Supt. in charge of curriculum.

Eliminate letter grades—rely more on achievement tests to find if there is any advance in knowledge.

Would like to determine weight of semester exams. Standardization of grading has been a bit difficult for us recently with having a large teacher turn-over. Any changes in our grading will be brought about through faculty discussion and agreement.

We still use the percent figure—would like to use the letter grades. I don't feel a teacher can grade as close as 1% tolerance. Give a broader method of reporting. The teachers are working toward this change through committees.

I would like to have grading on a satisfactory or unsatisfactory basis. We have grouping slow, average, fast and also have special classes in the math, science and the social studies area. We should not penalize students with honor and merit roll lists as we have found students avoiding certain classes in order to get on them.

To implement this change in grading, I believe it would have to be planned whereby the teachers and public would be informed thoroughly as to the reasons for this type of grading. This would take approximately a year before the grading would go into effect.
De-emphasize grading so as to reduce competition among students for honors and places on scholastic honor rolls. Record grades in numerical values, doing away with the controversial ABCDF system. Increase grading period from 6 to 9 weeks, and eliminate semester and final examinations.

I would like a P mark for low ability students that do not perform above basic reading, writing, and arithmetic and fail in regular academic mark. Transcript P would represent passing only.

We are interested in a grading system that reflects student achievement with his ability to achieve. This could be done by using a symbol beside the grade to designate the grouping he is in: Academically talented - average - remedial. It would be a terrific job educating this community to this type of grading system.

We are starting a "slow-learner" class. It will be our aim to instill upon our faculty the necessity of differentiating grades on these students. We will have in-service training on this subject.

Since our college-bound students will be exposed to the pressures of A's, B's, C's, and D's, 2.00's and 1.8's, and the like, when they arrive in college, we feel that they should have some acquaintance with this prior to college. For others where emphasis is not placed on grades to such an extent we would be pleased to simply indicate "progress." To implement would require a strong public relations program to overcome a half century of tradition, habit and custom.

The purpose and philosophy of marks are the least understood and most controversial. Frankly, I would like to see an end to marking as we traditionally know it and institute achievement based on levels. The ungraded approach interests me. In order to implement such a change, it would require "selling" to both parents and colleges (for college-bound).

With a three-track program we have three elements of marking. Would like to see an S or U used to indicate achievement. Class rank stands in our way. A parent-teacher-central office staff is going to study marking and reporting.
I should like to have a range of grades rather than specific figures. Thus I should prefer a marking system of A-B-C-D-F. Credit for the work for the year would hinge on scores on terminal and hopefully standardized tests. This would not be credit towards graduation, for students would be graduated from the school only on the completion of the accomplishment of certain minimal essentials, but credit and permission to proceed to the next stage in that specific subject. This would abolish rank in class, of course, a vicious artifact that I find casting a malignant influence over the entire high school program. How to implement all this? First, by working with and in the staff to gain understanding and acceptance of such a point of view. From there, we would expand our ideas into the community. Certainly at least three or four years would be needed to engineer such a change.

We have studied our grading system on several occasions but we have never had enough common agreement to warrant a change. God knows we need something. Hope you find the answer.

Summary

In Chapter IV, the data obtained from the questionnaires were organized into 14 sections, the first 13 of which corresponded to a hypothesis being tested, and the last of which summarized the free responses of the principals.

In each of the first 13 sections, the data were subjected to a statistical test to determine if a significant difference at the .10 level of confidence existed between the two sub-groups for the characteristic being tested. In six cases the null hypothesis was accepted outright, in five cases it was accepted with reservation, and in two cases no decision was possible. In no instance was the null hypothesis rejected.
In the last section of the chapter the investigator has summarized the principals' free responses, and has reproduced several of them verbatim for the reader's inspection.
CHAPTER V

PRESENTATION AND DISCUSSION OF SUPPLEMENTARY DATA

Relationship of School Size and Certain Financial Measures to Change in Grading Practices

The investigator established through the procedures described in Chapter III a total study sample of 80 schools evenly divided into two sub-groups, one innovative and the other non-innovative. For each of the 80 schools, certain data relating to size and financial condition was obtained in order that a comparison could be made between the two sub-groups.

Questionnaires were sent to each of the 80 schools, and 33 schools from each sub-group returned completed and usable questionnaires. From these questionnaires it was determined that 28 innovative schools reported making a recent change in grading practices, while 19 non-innovative schools made a similar report. To obtain as much contrast as possible, the investigator established a smaller group within each of the sub-groups. One is described as the 28 innovative schools that made grading changes, and the other as the 14 non-innovative schools that did not make grading changes.

Table 20 includes the four groupings described above. The values in the table are averages for each of the measures listed.
<table>
<thead>
<tr>
<th>Description</th>
<th>Number of Pupils in School</th>
<th>Property Value Per Pupil</th>
<th>Instruction Cost Per Pupil</th>
<th>Instruction Cost Percent</th>
<th>Total Expenditure Per Pupil</th>
<th>Secondary Principal Salary</th>
<th>Teacher Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovative schools (40)</td>
<td>1467</td>
<td>15605</td>
<td>318</td>
<td>70.6</td>
<td>450</td>
<td>11307</td>
<td>6882</td>
</tr>
<tr>
<td>Innovative schools that made grading changes (28)</td>
<td>1456</td>
<td>16994</td>
<td>331</td>
<td>71.0</td>
<td>467</td>
<td>11452</td>
<td>6995</td>
</tr>
<tr>
<td>Non-innovative schools (40)</td>
<td>507</td>
<td>10797</td>
<td>238</td>
<td>65.5</td>
<td>363</td>
<td>8632</td>
<td>5549</td>
</tr>
<tr>
<td>Non-innovative schools that did not make grading changes (14)</td>
<td>492</td>
<td>10558</td>
<td>231</td>
<td>65.6</td>
<td>353</td>
<td>8194</td>
<td>5481</td>
</tr>
</tbody>
</table>
It is seen in the table that certain relationships do appear to exist between the measures included and innovativeness, and particularly innovativeness in the form of grading changes.

A striking difference is immediately apparent with regard to school size. The innovative schools are seen to be, on the average, nearly three times as large as the non-innovative schools. Within each of the sub-groups, there is little difference in size between the schools that made and did not make changes in grading practices.

In all of the financial measures included in Table 20, the innovative schools reflect a substantially higher level than do the non-innovative schools. In property value per pupil, they are nearly one and one-half times higher; in instruction cost per pupil and total expenditure per pupil, they are about one and one-third times higher; in the percent of total expenditure used for instruction, they are 5 percent higher; in secondary principals' salaries, they are nearly one and one-third times higher; and in teachers' salaries, they are one and one-fourth times higher.

In each of the financial measures, it is seen that the innovative schools that made grading practice changes are slightly higher than all innovative schools taken together. Also, all non-innovative schools taken together are higher in each category of financial measure than the group of 14 non-innovative schools that did not make grading practice changes, with one exception. The exception is in the category of percent of total expenditure used for instruction, but the difference in the two figures is only .1 percent.
In four of the six financial measures, a trend is quite consistently followed, and it is partially observed in the sixth as well. It may be expressed this way: The more innovative a school is, particularly if it is innovative in grading practices, the higher the measures of the school's financial condition; the less innovative a school, and particularly if it is not innovative in grading practices, the lower the measures of the school's financial condition.

Impressions from Interviews with Principals

The purposes for conducting the interviews with ten principals (five from innovative schools that had made or were making changes in marking practices and five from non-interactive schools that had not made and were not making such changes) were described at some length in Chapter III. Briefly, the purposes were to find possible reasons why some schools effected changes in their grading systems while others did not, and to determine the means that were used to change the grading systems. Actually, to maintain that the reasons for change were found would be an overstatement. The investigator did find certain differences between the schools from the two groups, but to determine how significant each difference is as it relates to the dynamics of change in grading systems is beyond the scope of this study.

In the interviews, the principals in nearly every school said something about school size and various economic factors. These were presented and discussed in the first section of this chapter.
Mention is made of this at this point in order that absence of further discussion of these factors in this section is not interpreted by the reader as a glaring omission. There is little doubt that school size and financial conditions are cogent and pervasive factors in any consideration of innovativeness.

Principals in non-innovative schools indicated in several cases that they had difficulty recruiting and retaining good teachers. They cited as reasons lower pay levels and greater distances from large population centers. No complaint regarding teacher competence was made by a principal in an innovative school.

Problems of a basic nature were found more often to be a preoccupation with principals in non-innovative schools. These concerns included recent consolidation and subsequent tensions, rapid growth with limited facilities, and a low level of financial support from the community. The principals in innovative schools were more prone to mention as concerns improvement of programs and instruction, lowering the rate of drop-outs, and offering additional honors courses.

Involvement of the staff in program planning and curriculum revision was an area discussed more frequently by principals of innovative schools. One explanation may be that the innovative schools, by reason of their larger size and more extensive curriculums, are compelled to do more planning and revision. But several of the principals in non-innovative schools made comments that indicated they tended to subscribe to a do-it-yourself philosophy, and to administrative fiat.
A fact the investigator feels may possibly be significant is that none of the non-innovative schools he visited had an assistant principal. The principal alone was responsible for the entire spectrum of administrative functions. Two of the innovative schools had three administrative assistants each, one had one assistant principal, and two had none. One of the latter was planning to add an assistant principal next year.

Finally, the principals of the innovative schools seemed more knowledgeable regarding the philosophy and practices associated with school marks. This may be largely explained by the fact that each of these men has recently been involved with changing the marking practices in his school, while the principals in the non-innovative schools had had no such recent experience.

Each of the principals of the innovative schools commented on the means that were used to bring about a change in grading practices. The principal of a large city high school said a curriculum council of about thirty people including teachers, administrators, and PTA members suggested the change. It was tried experimentally at one high school, and when it worked successfully, was made a city-wide practice. A second principal from a high school in a medium-sized city indicated that an experiment in using subscripts with letter grades was being tried at a junior high school in the system, and if successful would be used at the senior high. The whole plan, however, was conceived by the administrators in the district.

A third principal in a large suburban high school devised a system for weighting grades, but submitted the plan to the staff
for their consideration before adopting it. A fourth principal at a medium-sized suburban high school related that grading had been a topic at two staff meetings this year. Out of these meetings came agreement that the primary purpose of grades is "aiding and abetting the process of learning," and that a student's failing grade must be considered a failure on the part of the teacher as well. A fifth principal in a small suburban high school stated that he with the help of a mathematics teacher, was in the process of devising a procedure for weighting grades. He admitted that when he completed the plan and the superintendent accepted it, he intended to incorporate it as school policy.

Information on Marking Practices Obtained from Grading Materials Submitted by the Respondents

Principals were requested in the cover letter to include with their completed questionnaires copies of reports to parents, grading policies, and descriptions of studies on grading they may have conducted. Over half of the principals did comply with this request, and from materials examined the investigator obtained considerable information. From the many items submitted, he has attempted to select and describe in this study those which represent a modern approach to marking and which may be of particular interest to other educators.

Most of the schools from which information was received used as a basis for the grading system the common ABCDF symbols to mark academic achievement. Many schools supplement these five
symbols with one or more additional symbols. Several schools use
"H" as the highest mark in an honors or advanced placement class
as a means of being more equitable to the superior student taking
more difficult class work. Also used fairly often is a mark for
"pass on effort," usually a "P." This mark is felt to be necessary
by the schools that use it in order to recognize effort even though
achievement is low, to refrain from failing students whose best
interests would not be served by failure, and to avoid diluting the
regular "b" grade.

Plus and minus marks are used with the letter symbols by
some schools when marking achievement. Most of these use them to
indicate a tendency to a higher or lower mark but do not award a
point value when averaging grades.

One school reported using a "condition mark" at the end of
the first semester. The mark is given a student who has not completed
the course requirements at that time, but who demonstrates good
attitude and effort. The understanding upon receiving a conditional
mark is that the work of the second semester will determine what
credit is to be granted for the first semester.

Cumulative marking was reported by one high school in lieu
of the usual independent marks given each grading period in a class.
The student is in this way informed each time grades are presented
exactly what his current grade status is up to and through the
grading period immediately past.

Evaluation of achievement is done in one way or another.
every secondary school. Most schools also evaluate in one or more other areas, including effort, citizenship, and need. In no case did the investigator find an instance in which evaluation occurred in all four areas on the same report to parents.

Where marks for effort or citizenship are awarded, there usually is an explanation of their meaning on the report form. Typical on a five-level scale is (1) Excellent; (2) Commendable; (3) Satisfactory; (4) Unsatisfactory - needs improvement; (5) Very poor - must improve. Three-level scales usually use symbols such as E, S, and U, or simply plus, no mark at all, and minus. Two-level scales usually use S and U. Effort was marked in one case by including the grade the teacher thought the student should get as compared with the grade he earned. Marks for effort sometimes appear as subscripts or exponents to the achievement grade, but more often are reported in a separate column.

Some schools indicate on the report form what citizenship means. Typical is this four-part description: (1) The student's ability to get along with others; (2) the student's attitude; (3) the student's conduct; (4) the student's service to the school. In this case it is intended that the ABCDF symbols are to be used, although one might question whether an "F" in citizenship is an appropriate mark.

An indication of student need, or suggestions for improvement, was included on several report forms. These consisted of a numbered list of statements, from which the appropriate number was
selected. A range of from nine to fourteen statements was found on the report forms examined by the investigator. One representative list includes these: (1) Attending all classes regularly and promptly; (2) preparing all assignments on time; (3) producing a higher quality of work; (4) listening to and following directions; (5) respecting rights and property of others; (6) responding to constituted authority; (7) refraining from unnecessary talking; (8) reporting to class with necessary supplies; (9) consulting with the teacher about a complex of factors is recommended to the parent. A blank space is sometimes provided to write in any statement not included in the list.

The statements of philosophy, policy, and procedure of some schools on grading are decidedly brief. There are cases in which only the latter category is treated, and sometimes that is little more than a description of the symbols to be used. Other schools treat the topic elaborately, and go into considerable detail even to the point of providing example calculations of grade averages.

Statements of a philosophy of grading vary from school to school, but usually include reference to the purpose of grades in that school, the subjective nature of teachers' marks, and the necessity for handling marking fairly and professionally.

One school presented a rationale for the grading system it employed in the form of reasons for accepting the system. The reasons are these:

1. It will simplify and facilitate the assigning, averaging, and posting of grades.
2. It will provide a means of differentiating grades to accommodate grouping.
3. It will provide a fairer method of determining class rank.
4. It will provide for the low achiever who is really trying.
5. It will provide a simplified system which can be more easily accommodated by modern record keeping systems.
6. It will provide for changes of policy in regard to honor roll, examination requirements, etc., without either the philosophy or the mechanics of the system having to be changed.

Most schools produce some statement of policy or guidelines regarding grading; some quite brief, others rather lengthy and detailed. A list of topics covered in these statements includes establishing defensible standards, normal distribution of marks, the use of special marks, marking in the light of excused and unexcused absence, student's right in demanding an explanation of a mark, the use of unsatisfactory progress reports, grounds for failing a student, a reminder that grades are not a disciplinary device, and other similar aspects of grading.

The weighting of grades was frequently reported by those schools having honors, advanced placement, or enriched courses or classes, and/or slow-learner or remedial courses or classes. Weighting has been done in two ways. One way is to restrict the number of high or low grades that are given in a particular course. The other way is to assign different point values for the same symbol in different classes. These two ways are not mutually exclusive, and some schools combine them in their grading systems.
Two schools have established five levels of instruction for the purpose of grade weighting. In one school the levels are as follows: Level Five, Advanced placement; Level Four, Honors courses; Level Three, Regular courses; Level Two, Attenuated courses; Level One, Special classes. An "A" grade in Level Five course is valued at seven points, a "C" grade five points. An "A" grade in a Level Three course is valued at five points, and an "F" grade is valued zero points wherever it occurs. In the same school an approximate distribution of letter marks is suggested. This distribution advises that "A's" and "B's" should constitute at Level Five 85 percent of all grades, at Level Four 45 percent, at Level Three 25 percent, at Level Two 20 percent, and at Level One 15 percent. It is also suggested that "C's" should constitute at Level Five 15 percent of all grades, at Level Four 45 percent, at Level Three 50 percent, at Level Two 60 percent, and at Level One 70 percent. This school provides an example, then, of a combination of the two methods of weighting.

A few schools used a factor method of assigning different point values to grades. The most common form of this method is that in which the point equivalents of grades in honors courses are multiplied by 1.2. In some cases, slow-learner course grade-points are multiplied by a .8 factor, but this occurs less frequently.

Two secondary schools reported having conducted studies of grades or grading practices. One of these was a survey of the grades given by different teachers over a period of time. Its object was to bring grades more into line from teacher to teacher.
The second study was carried out in a large school system rather than an individual secondary school. It was noted in the system that 4 percent of the students in the college preparatory program in one high school failed at least one course, while 32 percent of a similar group failed a course in another high school. The two groups were similar in ability and achievement, and should have fared equally well with high school grades. To investigate the situation further, a committee of teachers, principals, supervisors, and directors was selected to conduct a research study. The study is described as follows:

Groups were matched in each school on the basis of intelligence test scores and standardized achievement test scores in English and mathematics; then total grade point averages were compared. When mathematics grade point averages were compared for the groups matched on IQ and mathematics achievement, it was found that the average GPA varied from a low of 2.02 (C = 2.00) in one school to a high of 3.02 (B = 3.00) in another. When the matched groups were compared in English, it was found that the average GPA varied from 1.91 (D = 1.00) in one school to 2.87 in another. When it was noted that the average scores for the groups were above IQ 120, were at the 80th percentile when compared to national norms in English, and were at the 99th percentile in mathematics, these differences in average GPA's from school to school became all the more difficult to explain. (50,2)

This study was the basis for a revision of the city-wide grading system.

Automated grading with the use of a computer-printer was reported in use in many high schools, particularly those high schools in larger school systems. At least two schools indicated on the
questionnaire that a recent change in their marking symbols had been
effected to accommodate computerized grading. Also, two schools
reported that teachers recorded pertinent grading information on a
mark-sense card that was translated by the computer into a grade report. Such a device does not relieve the teacher of any of the
work or responsibility associated with determining students' grades,
but it does expedite considerably handling the grades after the
teacher has recorded them.

The traditional grading period in secondary schools is six
weeks. However, in the statements of grading policies and procedures
submitted by the principals, it was seen that some schools have
adopted the nine-week grading period. The advantage of this is self-
evident as one considers the amount of effort teachers must expend
and the emotional strain many teachers experience at each grading
period.

Information submitted by two of the schools indicated that
they completed profile sheets primarily for the use of college
registrars. The information on the profile sheet from one of the
schools included a brief statement of pertinent facts about the
community and school, useful information about class levels, grade
distribution of the senior class, seventh semester median grade
averages for senior students, and class average scores on a number of
standardized tests. The profile sheet, according to one principal
whose school produced one, was appreciated by college registrars as
an aid in interpreting the transcripts from that school.
Summary

In Chapter V the investigator discussed the supportive information obtained from Ohio State Department of Education publications, from personal interviews with ten principals, and from marking-related materials submitted by the principals.

Innovative schools are shown to be much larger and to rate higher in financial measures than non-innovative schools. Those innovative schools that have effected a change in marking practices rank higher in the financial measures than the average of all innovative schools. Those non-innovative schools that have not made changes in grading practices rank lower in the financial measures than the average of all innovative schools.

From the interviews, it was apparent that principals from non-innovative schools were more concerned than their counterparts from innovative schools at the essential level of maintaining the educational program. Principals of innovative schools spoke more frequently of expanding and improving the program.

An examination of marking-related materials submitted by principals indicated that there is little that is unique being done with student marking in Ohio's secondary schools. Nevertheless, some schools do employ reasonably modern practices, and a number of these were described.
CHAPTER VI

DISCUSSION OF FINDINGS AND RECOMMENDATIONS

Statement of the Position of the Investigator

One principal commented on a questionnaire, "God knows we need something. I hope you find the answer." That principal will be disappointed in this study if he expects to find "the answer" in it. The position taken by the investigator as he undertook the study was that there is no ideal grading system appropriate for all secondary schools, no panacea for all marking problems. Nothing he found in the course of the study suggested this position should be altered. He still feels that it is the obligation of each educator to decide what in his judgment is the best grading system for his particular situation. Furthermore, a well-informed educator will be in a position to make an intelligent judgment.

This position of the investigator is reflected in the design of the study and implies certain uses the study may have. These are (1) to encourage the educator to review and to revise if necessary present grading practices, (2) to inform the educator what modern grading practices are presently in use, (3) to apprise the educator of those factors which appear to be related to changes in marking
practices, and (4) to advise the educator of means which may be used to effect a change in grading practices.

Conclusions from Review of Past Marking Practices

During the years between 1915 and 1935, many more or less mechanical innovations were implemented, but generally were of little value in improving the marking situation. This consistent pattern of failure, or at best limited success, helped to make clear that, for the most part, esoteric knowledge and systematic scientific research, as exemplified by Starch's efforts, will provide the sort of information that will be useful in improving marking practices. We need to know, for example, more about how to measure the acquisition of attitudes and insights in addition to factual material; we need to be able to define curriculum objectives in a manner that facilitates measurement of the extent to which the student achieves them; we need to understand more about motivation and the impact of success and failure experiences; we need to improve the curriculums to provide in our secondary schools the experiences that are truly useful and important for the students; and we need to train teachers in the more effective use of school marks. This is not to say that we cannot improve the marking system until we know all there is to know about the several aspects of evaluating student achievement. It does mean that the number of important improvements that may occur in marking practices is to a large degree contingent upon, and therefore limited by, the extent of parallel improvements in a number of
other areas of the total educative process, and the transfer of those improvements to marking practices. As one considers the nature of the improvements that are necessary, the task of selecting and assimilating those improvements relevant to improving marking practices, the communication necessary to disseminate promising suggestions, and the resistance to adoption of innovation at the level of practice because of tradition, inertia, and pressure from the community and from colleges, one then understands why grading practices have changed so little while the need for change has long been obvious. The situation, briefly stated, is that improving marking practices appears deceptively simple; it is, in fact, a highly complex undertaking.

Conclusions from Responses to the Questionnaire

Throughout this section of the discussion related to the responses to the questionnaire, it is well if the reader bears in mind that the responses are those of principals. This must be considered to be a possible source of bias in the findings. And the investigator, himself a principal, may be less sensitive to this bias in some instances than another reader.

Innovativeness was the key used by the investigator in his approach to the whole area of grading. Thirteen hypotheses were formulated, each stating that there would be no significant difference in innovative and non-innovative schools with regard to certain aspects of their marking practices. It was the opinion of the investigator as he commenced the study that the null hypothesis
would be rejected in the majority of cases, and the alternative hypothesis (i.e., a significant difference does exist between innovative and non-innovative schools with regard to certain aspects of their marking practices) would be accepted. A tally of conclusions derived from the data, however, includes no rejection of any null hypothesis. Six of the null hypotheses were accepted without reservation, five were accepted with reservation, and two were neither accepted nor rejected. Table 21 presents a resume of the decisions regarding acceptance and rejection of each null hypothesis.

**TABLE 21**

RESUME OF DECISIONS REGARDING ACCEPTANCE OR REJECTION OF THE HYPOTHESES

<table>
<thead>
<tr>
<th>Hypothesis Number</th>
<th>Hypothesis Description</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Use of unusual marking practices</td>
<td>Accept</td>
</tr>
<tr>
<td>2.</td>
<td>Differentiation in marking in able students' courses</td>
<td>Undecided</td>
</tr>
<tr>
<td>3.</td>
<td>Differentiation in marking in slow-learning students' courses</td>
<td>Accept</td>
</tr>
<tr>
<td>4.</td>
<td>Have a suggested philosophy of grading</td>
<td>Accept</td>
</tr>
<tr>
<td>5.</td>
<td>Have a written policy on grading</td>
<td>Accept</td>
</tr>
<tr>
<td>6.</td>
<td>Bases for determining students' grades</td>
<td>Accept w/res.</td>
</tr>
<tr>
<td>7.</td>
<td>What is considered to be the primary function of marks</td>
<td>Accept</td>
</tr>
<tr>
<td>8.</td>
<td>Person initiating the change in marking</td>
<td>Accept w/res.</td>
</tr>
<tr>
<td>9.</td>
<td>Reasons for making a change in marking</td>
<td>Undecided</td>
</tr>
<tr>
<td>10.</td>
<td>Source of the idea for a change in marking</td>
<td>Accept w/res.</td>
</tr>
<tr>
<td>11.</td>
<td>Procedures for implementing change in marking</td>
<td>Accept w/res.</td>
</tr>
<tr>
<td>12.</td>
<td>Persons supporting change in marking</td>
<td>Accept w/res.</td>
</tr>
<tr>
<td>13.</td>
<td>Persons resisting change in marking</td>
<td>Accept</td>
</tr>
</tbody>
</table>
There is one important point that needs to be clarified in order that the reader does not obtain an incorrect impression from the summary of the results of the tests of the hypotheses. Three of the sections on the questionnaire consisted of two-part questions. The purpose of the first part of the first two of these questions was to establish the number of schools in each sub-group that offered special courses in which differentiation of grades might be appropriate. For Section II, responses indicated that 30 innovative schools offered special classes for very able students, while only ten non-innovative schools offered such classes. For Section III, responses indicated that 30 innovative schools offered special classes for slow learners, while only 19 non-innovative schools offered such classes. In other words, innovative schools are far more likely than non-innovative schools to offer special courses for either the very able or slow-learner students. It should be noted, however, that the hypotheses corresponding to each of these sections was concerned only with whether differentiation of grades occurred where these special kinds of courses were offered.

Section VIII was the last of those having two parts. In the first part of it, the principal was asked if any change in marking practices had occurred recently in his school. Of the 33 innovative schools responding, 28 indicated there had been; of the 33 non-innovative schools responding, 19 indicated there had been. A comparison of the two proportions produced a K value of 2.175. This value lies at the .02 confidence level. In other words, there are
98 chances out of 100 than an actual difference exists between innovative and non-innovative schools for the characteristic in question, and only two chances out of 100 that the observed difference is due to chance alone. Very clearly, then, innovative schools are more likely to produce examples of change in grading practices than are non-innovative schools.

A review of the responses to Section I brings the investigator to the conclusion that disappointingly few schools report employing the grading practices listed. Most of the practices listed on the questionnaire are considered to be modern ones, updating the traditional practices. The investigator does not contend that all schools should employ all the practices. In fact, certain ones preclude certain of the others. He is of the opinion, however, that it is an unfortunate commentary that so few schools use these practices. Only in the case of supplementary progress reports or failure warnings does the majority of schools indicate use of the practice. And only for one other item was a response given by over 25 percent of the schools, that being the item dealing with a series of statements regarding student achievement to be checked by the teacher.

The investigator was able to check specifically into two of the four schools that indicated the only report to parents is simply pass-fail, or satisfactory-unsatisfactory. In both of these schools the conventional ABCDF symbols were used in the normal way. The principals evidently misunderstood the intent of that item on
the questionnaire. There is a doubt in the investigator's mind whether the other two schools that responded similarly actually do as they have indicated. It seems more likely that a misunderstanding occurred in those cases as well. This contention is supported by Terwilliger's data from a survey of marking practices in 129 secondary schools which indicated that no schools used the pass-fail method of reporting to parents (42.,10).

The investigator finds himself incredulous at the response to Item 12 in Section I dealing with standardized tests. His impressions from his other sources of information cause him to doubt the relatively high response here. It appears that many of the respondents may have read the first line of the item which states, "One or more of the following types of standardized tests are employed," without reading the remaining portion of the sentence which continues "as a basis or factor in providing information appearing on reports to parents."

From Sections II and III on the questionnaire may be obtained information regarding the frequency with which special courses for the very able and the slow learners are offered, and the extent to which differentiation of grades in those courses occurs. The investigator conducted a survey of 83 Ohio high schools in 1963 from which comparable information was obtained. Table 22 presents these data for the sake of comparison.
TABLE 22

PROPORTION OF SECONDARY SCHOOLS IN OHIO THAT OFFER SPECIAL COURSES FOR VERY ABLE AND SLOW-LEANERS, AND THE EXTENT TO WHICH DIFFERENTIATION OF GRADES OCCURS IN THOSE COURSES: 1963 AND 1966

<table>
<thead>
<tr>
<th>Item</th>
<th>Year</th>
<th>1963</th>
<th>1966</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schools indicating their curriculums included special courses for very able students</td>
<td></td>
<td>69%</td>
<td>61%</td>
</tr>
<tr>
<td>Schools indicating their curriculums included special courses for slow-leaner students</td>
<td></td>
<td>64%</td>
<td>74%</td>
</tr>
<tr>
<td>Schools indicating differentiation of grading in courses for very able students</td>
<td></td>
<td>39%</td>
<td>58%</td>
</tr>
<tr>
<td>Schools indicating differentiation of grading in courses for slow-learners</td>
<td></td>
<td>53%</td>
<td>41%</td>
</tr>
</tbody>
</table>

The body of data from which the figures in the table are derived is too scant to claim a trend, although further investigation into this topic might be fruitful. What is indicated here is that in 1966 a slightly smaller proportion of high schools have honors-type courses, but a higher proportion of those having them differentiate the grading. Conversely, in 1966 a larger proportion of high schools have slow-learner-type courses, but a smaller proportion of those having them differentiate. These results seem to be in line with the feelings expressed by several principals in the course of interviews. Several were concerned about being fair to students doing the more difficult honors work, and yet hesitated to penalize the slow students for their inability to meet the demands of normal classes.
The data above were obtained only from high schools in Ohio. Similar information was obtained by Terwilliger in 1963 from 129 high schools from all parts of the country chosen to be representative of all high schools. His commentary on his data was that Seventy-five percent of the schools in our sample report that tracks are used in certain subjects but only 25 per cent have formalized policies for grading in subjects where tracks exist. A variation of the same problem is the weighting of grades in tracked subjects for computation of class rank. Only 9 per cent of the schools in this study report giving special weight to honors and advanced placement courses. There is an urgent need for careful study of this situation. (42.34) Clearly, then, in 1963, Ohio high schools placed greater emphasis on weighting of grades in different levels of courses than do schools generally throughout the country. Presumably, judging from the sizable gap between Ohio high schools and all American high schools in 1963, Ohio high schools continue to exceed the average of all high schools with regard to the frequency with which weighting of grades occurs.

From Sections IV and V in the questionnaire it may be determined that 62 percent of the schools responding claimed to have a written philosophy of grading, and that 58 percent had a written statement of policy. In the materials submitted along with the questionnaire, there were far more statements that qualified as policy than as a philosophy. Even so, the documents that were analyzed varied from the very brief to the very extensive and no one pattern appeared to predominate.
This study has separated philosophy and policy. Terwilliger (42, 34) had a similar category in his study except that he described it only as a "written policy on grades." His study sample was selected to be representative of all schools throughout the country. His findings were that 52 percent of the schools in his sample had a written policy on grades. His data were gathered in December, 1963, whereas the data in the present study were gathered in March, 1966. Even considering this time difference, it would appear that Ohio ranks above the average of all schools in the nation in this category.

In Section VI on the questionnaire, it is noted that a competitive basis is the most popular used to determine students' grades. At any time such a practice is used, there is the danger Wrinkle warns of concerning the unfair competition of unequals. Grouping, establishing special courses, and weighting of grades are all means used to make the competition more fair. Nevertheless, avoiding competitive marking in favor of achievement-ability marking or individual achievement marking is more compatible with a democratic philosophy of education. However, there are practical problems of implementation of such a basis for grades.

Principals indicated in their responses to Section VII that they were in close agreement in their opinion that the primary function of marks was to inform the student and parents of the quality of the student's work. As one considers that grades have been shown to be relatively poor informants of any specific information, and that
they operate more effectively as a motivator than in any other
function according to Wrinkle (19.), one may then have some concerns
about the way grades are used. Principals also agree closely that
grades have no justifiable use as a disciplinary measure. There was
a time not so many years ago when grades were very much a disciplinary
device. Perhaps the greatest progress in the whole area of grading
in recent decades has been the effective elimination of this con-
notation from grades.

Sections VIII through XIII deal with the dynamics of
change in grading practices, the dimensions of which include
initiators, reasons, sources, means, supporters, and resisters. The
study data relating to each of these dimensions of change are pre-
sented in Chapter IV, and the interested reader may refer there to
determine what the factors of each of these dimensions are, and the
relative importance of each factor. But, disregarding individual
factors and dimensions and focusing on the larger pattern, the most
significant single finding is that the principal is the most important
agent in the overall change process. The data indicate that he is
the most frequent initiator and supporter of change in marking
practices. He also is frequently an idea source as well as a
director of change. He rarely resists change in marking practices.
These findings support Demeter's conclusion referred to earlier in
this study that "Building principals are key figures in the (change)
process."

Also significant as one observes the larger picture is the
extent to which faculty committees are involved in change in grading practices. Here again, however, it is usually the principal who appoints the faculty committee, and he may possibly be its most influential member. But include the principal or not, faculty committees are seen to be a frequently involved agent in changing grading practices, ranking only slightly behind the principal as an initiator of change, and just ahead of him as a source of ideas.

The faculty is also intimately involved with the implementation of a change in grading. If one takes into account three of the procedures for implementing a change—(1) vote by, or agreement by, the faculty; (2) study project within the school or school system; and (3) meetings or workshops to inform or orient teachers—then the faculty is involved in more than a passive way in 57 percent of the cases. It is interesting to note also that principals rate teachers almost as supportive of change in grading practices as they rate themselves.

With regard to resistance to change, it appears that almost nobody actively objects to change. The principals reported that only a few teachers and very infrequently a counselor wants to continue "the old way." The principals did not indicate that parents resisted change, although McKeen is quoted in the first chapter of this study claiming that "unjustified parental confidence" in the traditional grading practice is one of the primary reasons for lack of more change. The evidence from this study is much too brief to conclude that McKeen is mistaken in his assumption, but it does present an interesting hypothesis for further investigation.
In response to Section IX on the questionnaire, slightly over one-half of the principals indicated that the reason for a change in marking practices was to bring those practices into line with the curriculum. Not once, however, did any principal specify, either in his written comments on the questionnaire or during the personal interviews, exactly what was meant by that response. This investigator was remiss, he feels, in not pursuing this topic further at the time of the interviews. This investigator is reasonably certain, however, that Wrinkle's interpretation of that statement—making marking practices an integral part of the curriculum rather than an adjunct to it—was not what the principals had in mind.

In retrospect, the investigator recognizes three shortcomings in the set of hypotheses he formulated. Underlying the first criticism of the hypotheses is the fact that, contrary to what the investigator anticipated, none of the hypotheses were rejected. Still, the investigator is of the opinion that there may be important differences between innovative and non-innovative schools to which the hypotheses, as stated, were not sensitive. The addition of one or more hypotheses, with corresponding sections on the questionnaire, designed to investigate the nature of the changes implemented in each sub-group would have been a valuable addition to the study. If the investigator is correct in his opinion, then the null hypotheses that would have dealt with the nature of such changes would have been rejected.

The second shortcoming relates to the previous discussion
In this chapter regarding the three two-part sections. The first part of each of these questions implies an hypothesis. In each case the implied hypothesis has been tested and the K value corresponding to it displayed in a table, just as the stated hypothesis it preceded. Because these implied hypotheses were critical to the study, they should have been included in the study as stated hypotheses. Each, incidentally, would have been reported as a rejected null hypothesis.

The final shortcoming concerns the fact that the latter six hypotheses, Hypotheses 8 through 13, compare innovative and non-innovative sub-groups of schools in the extent to which certain aspects of the dynamics of change were reported. In the investigator's opinion, this extended the comparison of the two sub-groups beyond the point where such a comparison could make a contribution to the study. The first seven hypotheses dealt with what grading practices were employed, and in these a comparison between sub-groups was appropriate. In the latter six hypotheses, the investigator's more urgent concern was how changes in grading practices had occurred wherever they had occurred, and not in which sub-group the school was included. With this notion in mind, the investigator added a third section to each of the tables in Sections VIII through XIII, which presented combined data from both sub-groups of schools. From this section of the tables the more pertinent relationships were derived.
Conclusions from Discussion of Supplementary Data

The generalization from Chapter V summarizing the analysis of financial measures was as follows: The more innovative a school is, particularly if it is innovative in grading practices, the higher the measures of the school's financial condition; the less innovative a school, and particularly if it is not innovative in grading practices, the lower the measures of the school's financial condition. This conclusion is in agreement with the findings of Ross and of Rogers referred to in Chapter I. However, a close analysis of certain data does present one inconsistency—or at least, qualification. The generalization above stresses the importance of economic factors as an indication of innovativeness. And yet, 19 of the 33 schools judged to be non-innovative with regard to curriculum change did report a change in grading practices. So while Rogers states that "The wealth factor almost appears to be a necessary prerequisite for innovativeness,"⁴ changes in grading practices seem not to be as dependent on that factor as other sorts of innovations.

From the interviews with principals, it was noted that their concerns included difficulty recruiting good teachers, recent consolidation, rapid growth and limited facilities, and lack of an

⁴It seems appropriate to note here that the increasing number of government-sponsored curriculum programs being incorporated in the less wealthy school districts in recent years tends to counter Rogers' claim. But the very basis for Rogers' statement (perhaps overstatement) is in part the basis as well for the government's initiative.
administrative assistant. Upon closer inspection, all of these concerns are symptoms of the lack of adequate financing of the educational program. The opinion is expressed often enough that money alone isn't the answer to the problems of education, but the evidence that this investigator finds both in the literature and from the present study indicates that money certainly helps.

From the materials related to the grading system submitted by the principals, it was observed that no school included evaluation in all four areas mentioned on a composite of all materials. It is doubtful that more than an isolated few public school educators today accept the premise once popular that education is wholly concerned with the acquisition of subject matter. The role of the school in our society has evolved to the point where it is now generally conceded to have a concern with the whole intellectual and social development of the child. When acquisition of subject matter was the singular objective of the schools, a single percent grade or letter symbol was much more adequate than it is now. In order to describe the child's progress more meaningfully and more completely in the many dimensions of child development with which the school is concerned, evaluation in several categories is necessary. Those presently used by many schools are achievement, effort, citizenship, and suggested improvement. Even these all together sometimes are inadequate in exceptional instances. The written notice and the parent conference are appropriate in such cases. Unfortunately, the latter two practices are too time-consuming to use extensively, for they would be preferable even in non-exceptional cases.
General Findings

There is change occurring in marking practices in the high schools in Ohio—more change, from indications available to this investigator, than in schools throughout the nation generally. Seventy-one percent of the schools responding to this study reported that some change in the grading system had occurred since 1958, although most of this change occurred within the past three or four years. But even with change occurring frequently, occasionally a school is still found that uses the 100-point grading scale discredited by Starch a half-century ago.

The change that is occurring is entirely of an accommodation type. There do exist many well-defined grading systems, many elaborate ones, many with much to recommend them. But nowhere is a revolutionary departure to be observed. The reasons for this have been alluded to in other parts of this study, and they are understandable. Nevertheless, it is a discouraging commentary. Terwilliger recognized this problem in the comments on his study when he stated that

The introduction of a standard set of separate character and work habits might help... (where there is a variety of information to be transmitted). Only 57 percent of the schools in our sample currently use such ratings. Schools that do not have a separate marking system for these aspect of performance force their teachers to incorporate them with their subject matter grades. (42, 34)

It seems a fair conclusion from the tenor of the principals' responses to this study that grading practices are their sincere
concern. Principals indicate that teachers very commonly share this concern as well. And all this is to be expected at this level where marking is a necessary and constant practice. But concern appears to stop here. Almost nobody has done work at the theoretical level with marking practices. No revolutionary grading system has appeared because no rationale involving theory, psychology, and philosophy has been devised. Not since the core curriculum integrated student evaluation, particularly self-evaluation, as an intrinsic part of the curriculum, has any major breakthrough occurred in grading. The core curriculum movement has lost its momentum, and this is unfortunate from the standpoint of student evaluation.

Grades will continue to be employed in a more or less traditional way by the large majority of schools for some time to come. This investigator, from a philosophical point of view, would prefer to eliminate students' marks, or to use a core-type teacher-student cooperative evaluation. But even within the context of our present grading practice substantial improvement could be realized by means described in this study—encouragement by the principal, involvement of the faculty, research on grades, and orientation of the parents. For example, the supposed necessity for giving grades six times a year seems exaggerated. Some schools have successfully adopted a four-grading period schedule, and a few schools report they report grades formally only once a semester. The investigator sees no reason why a once-a-semester grade similar to the college practice could not serve equally as well as three grades per semester. These characteristics should perhaps be included in such a marking plan.
(1) A cumulative record of achievement so that a student can be informed at any time (and he should be informed frequently) of his grade status;

(2) A conference with the student and notification of the parents when a student is in danger of failing, when the quality of work drops sharply, when effort lags noticeably, and when achievement is consistently not commensurate with ability.

This "improvement" certainly does not solve any one of the numerous problems of grading. It does, however, relieve the teacher of the necessity for preparing grades formally on three different occasions during the semester, and gives him the opportunity to be more directly involved with the students and parents most in need of this attention.

This study was undertaken with the assumption that innovativeness was the key to change in grading practices. The assumption was affirmed in Section VIII where significantly more innovative schools (85 percent) reported changes in grading practices than did non-innovative schools (58 percent). But in Section I, innovative and non-innovative schools reported virtually identical numbers of the "unusual" practices being employed. Also, at no point throughout the study was the null hypothesis rejected. This means that the investigator found no instance of a significant difference between innovative and non-innovative schools in the characteristics described in the hypotheses. Finally, Stufflebeam, in identifying innovations in school districts, did not mention any innovation
involving a change in grading practice. Nevertheless, 19 of the schools judged from Stufflebeam's study to be non-innovative did report a change in marking practices.

The explanation of the apparent inconsistency described above is, at least in part, that the commonest sort of change in marking practices, e.g., adding a "P" symbol to an established set of symbols, or even substituting ABCDF for a percentage scale, causes no major disruption nor necessitates difficult reorientation. It usually constitutes the first, or lowest, level of innovation suggested by Chin (page 21 of this study), the level at which innovations are most quickly and easily incorporated into the existing structure. Also, as one refers to Rogers' (page 21 of this study) five characteristics relating to the rate of diffusion of innovations, he observes that the sort of changes in marking practice described above would, indeed, diffuse quickly, sometimes almost unnoticeably.

The seeming inconsistency resolves itself into the question, "When is a change an innovation?" Perhaps innovation (loosely defined as it is), like beauty, lies in the eyes of the beholder. In any case, Stufflebeam's criteria for innovativeness were obviously not sensitive to simple changes in grading systems, and admittedly, none of the changes reported in this study represented major departures from traditional practices. It occurs to this investigator that change in marking practices is an active expression of a school's willingness to move and desire to improve, and as such may constitute a sensitive indicator to the tendency to innovate.
Recommendations

Before proceeding with his own recommendations, the investigator refers the reader to page 18 of this study where there is listed a National Association of Secondary School Principals-sponsored set of action guides for improving grading practices. This investigator is in agreement with those recommendations, and feels that they complement his own recommendations which follow.

The recommendations from this study fall into three categories as indicated by the sub-headings in this section. However, the recommendations may also be considered to fall into a second set of four categories determined by the source from which they were derived. These categories, with the numbers of the recommendations which each includes indicated in parentheses, are as follows:

1. Recommendations derived directly from or implied from the statistical analysis of the data (1, 2, 3, and 4)

2. Recommendations which were developed from the principals' free responses to Section XIV, from their statements during personal interviews, or from marking-related materials submitted by them (6, 10, 12, and 13)

3. Recommendations derived from a review of the literature (7, 8, and 11)

4. Recommendations representing the investigator's impressions and convictions obtained through the course of the study (5, 9, 14, 15, 16, and 17)
Recommendations relating to the dynamics of change

1. Any change in grading practices should have the encouragement and participation of the principal, for he has the focal position in innovation in the secondary school.

2. Any change in grading practices should involve the faculty of the school as much as possible in the inception and planning of that change.

3. At the time a change in the grading system is being implemented, thorough explanation should be provided teachers, counselors, students, and parents.

4. Provision should be made in each high school for a standing committee whose function it is to provide a continuing review of the marking system.

5. Studies of the grading system should be conducted frequently to detect weaknesses in current practice, and to provide support for suggested improvements.

6. Experimental marking practices should be tried on a small scale when this is appropriate to determine their effectiveness and acceptance before being fully incorporated into the grading system.

7. At the time the grading system is being redesigned, consideration should be given to the four functions marks serve--
(1) administrative, (2) guidance, (3) informative, and (4) motivation—in order that there is assurance that each of the functions is effectively served.

8. At the time a curriculum change occurs (e.g., an advanced placement course or a pre-vocational program is added) there should be concomitant consideration of the implications for the marking system.

9. The principal should make it a point that in his school there exists or there is to be developed a comprehensive statement of the philosophy, policy, guidelines, and procedures for marking.

Recommendations relating to marking practices

10. Evaluation of students should be conducted in at least four dimensions of the students' intellectual and social growth: (1) academic achievement, (2) effort, (3) citizenship, and (4) suggestions for improvement.

11. The competitive basis for marks employed by most teachers and most schools because of tradition and pressure from community and college should be deemphasized, and marks that are individual, related to ability, and that promote learning should be encouraged.

12. The length of each grading period should be extended to a full semester, and the number of grading periods should be diminished from the usual six per year to two per year.
13. A means of differentiating grades in exceptional students' courses should be employed in order that the very able students are not undeservedly penalized nor the slow-learner students unduly rewarded.

General recommendations

14. Colleges of education should accept the responsibility for apprising prospective high school teachers of the philosophy of marking, the purposes of marks, the findings from research on marking, modern marking practices, the inherent fallacies of marks, and the necessity for a deliberate rather than intuitive approach to marking.

15. The same attention should be given to "measuring school products" at the present time as in 1912-1918 when marking student achievement was "one of the three or four most important fields of investigation in the scientific study of educational problems."

16. Educators working at the level of secondary education should have as a long-range goal the eventual elimination of the traditional grading system with its inevitable shortcomings, and should strive to develop a system of evaluation having a defensible rationale.

17. A national committee on marking practices and grading systems should be established to enlist the efforts of creative and influential educators on behalf of the improvement of marking practices, to initiate broad research on marking practices, to
coordinate the activities of those persons or agencies contributing to improvement of evaluation in the secondary schools, and to disseminate pertinent information deriving from the above activities to interested educators.

Concluding Statement

The marking of students is, in the final analysis, a naive attempt to describe with very simple, very naked symbols an exceedingly complex pattern of intellectual and social growth in students. Such a practice can at best meet with limited success. To the extent that the practice can be improved, it should be. But improvement of traditional practices can never really solve the problems related to the "mess in marks." Only when the educative process has been redefined, either to exclude the necessity for evaluation, or to include it as an integral element, will the problems we now associate with the evaluation and marking of students be resolved.
Description of

Ohio Education Innovations Survey

In 1963, the College of Education of The Ohio State University began a study designed to identify what educational innovations were currently being employed in the curriculums of Ohio's public schools.

Daniel Stufflebeam, director of the study, stated the purpose of the study.

The Ohio Educational Innovations Survey ... was conducted in the belief that information about recent changes in Ohio public schools would be useful to teachers and administrators for identifying, studying, and adopting new programs; to educational researchers for analyzing the content and process of current change in Ohio public schools; and to educational leaders for organizing education to facilitate the change process. (51., l)

The findings from that study have been published under the title of Catalogue of Educational Changes in Ohio Public Schools.

Stufflebeam's study was carried out in four steps. The first step comprised a questionnaire survey of all Ohio superintendents and executive heads asking them to list and briefly describe new programs and procedures in their districts from approximately 1958.

Step two was a screening process of the responses to the questionnaire. At least three judges with professional education experience rated each reported activity on a 1 to 6 scale as follows: (1) clearly innovative, (2) maybe innovative; (3) new national program,
(4) not innovative, but worthy of follow-up, (5) not enough information to decide the merits of the program, and (6) clearly not worthy of further study.

In the third step of the study, a second questionnaire was administered for the purpose of gathering additional data about each program that received a rating higher than six from the professional panel. This second questionnaire was sent to that person who had been designated by the superintendent or executive head as the person most knowledgeable about that particular program.

The final step in the study was evaluation of the responses to the second questionnaire. Four categories were used in rating the reported programs: (1) The approximately thirty programs most worthy of follow-up, visitation, and on-the-scene study; (2) the approximately 170 programs which on the basis of the information included in the questionnaires were judged worthy of summary description in the catalogue; (3) programs in addition to the 170 programs judged to merit inclusion in the catalogue listing of titles; and (4) programs inappropriate for listing in the catalogue. Evaluation of the questionnaires, and designation of each into one of the four categories, was the responsibility of an eight-member survey team made up of highly regarded, experienced educators. The criteria employed by the survey team in deciding upon the ranking of programs were these:

1. Uniqueness: May be unique in either or both of two ways, i.e., conceptualization and implementation.
2. Clarity of Objectives: The school must have implemented the program with a definite goal or goals.

3. Care in Planning: A maximum of rational preparation for a new program is deemed desirable.

4. Relevance: A program dealing with a critical problem area is to be given advantage over one of questionable relevance to current educational problems.

5. Applicability: A program should apply either directly or in modified form to many schools.

6. Provision for Evaluation: There must be provision for some formal means, developed by the school or the district, by which the new practice is to be evaluated.
Questionnaire—High School Marking Practices

**Name of school for which this questionnaire is prepared** _________________________

**Note:**
1. The technical term "marking" and the more casual term "grading" are used interchangeably in this questionnaire.
2. If in responding to any item or group of items on this questionnaire you feel that a brief explanation will provide valuable clarification, you are encouraged to include this.
3. The information and opinions you include will be treated in an entirely professional and confidential manner. The report which is made of this study will not include the name of any individual school or administrator.

I. Indicate with a check in the space provided which, if any, of the following marking practices are employed in your school.

1. _____ Dual grades are included on reports to parents, e.g., a grade for competitive achievement and a grade for achievement based on ability.

2. _____ An unusual set of symbols, or an unusual arrangement of symbols, is used, e.g., subscripts (B₁, B₂) or fractions B
d

3. _____ The report to parents includes a graph representing a profile of student achievement.

4. _____ Different symbols are used in academic and non-academic courses, e.g., use of ABCDF in an academic course and S-U in a non-academic.

5. _____ In addition to a symbol grade (letter or number) the report to parents includes a statement written by the teacher regarding student achievement or character.

6. _____ In addition to a symbol grade (letter or number) the report includes a series of statements regarding student achievement or character to be checked by the teacher.

7. _____ A written statement or checklist unaccompanied by any symbol grades constitutes the formal report made to parents.

8. _____ Self-evaluation by pupils is a factor in determining the course mark.

9. _____ Parent-teacher conferences are an integral or necessary part of the grading system.

10. _____ The only grade on the report to parents is simply pass-fail, or satisfactory-unsatisfactory.

11. _____ Supplementary progress reports or warnings of possible failure are sent to parents at any time that such a report is deemed appropriate.
12. One or more of the following types of standardized tests are employed as a basis or factor in providing information appearing on reports to parents:

___ achievement
___ intelligence
___ aptitude
___ personality inventory

13. A grading period other than the usual one of six weeks duration is used as follows:

___ monthly
___ 9 weeks
___ semester
___ reports to parents are on a rotating basis, rather than to all parents at once

14. Check this space if you employ any marking practice not listed above other than the common ABCDF or a scale based on 100. Please describe it briefly on the back of this page.

II. Does the curriculum in your school include enriched, advanced placement, accelerated, or honors courses or classes for the very able students (in addition to the regular college preparatory courses)? Yes  No

If yes, do you in some way differentiate grades (e.g., weighting of grades or use of different symbols) in these courses as opposed to the regular courses? Yes  No

III. Does the curriculum in your school include slow-learner or remedial classes or courses for students unable to meet the more rigorous standards in regular classes? Yes  No

If yes, do you in some way differentiate grades in these courses as opposed to the regular courses? Yes  No

IV. Does your school have a written philosophy pertaining specifically to grading practices that teachers are asked to consider in assigning marks? Yes  No

V. Does your school have a written statement of its policy on grading which stipulates that teachers shall subscribe to certain standards and/or employ certain practices (other than simply a brief statement indicating that teachers will use a certain common set of symbols)? Yes  No

VI. Which of the following is the basis used by most of the teachers in your school to determine students' grades?

1. ___ A competitive basis which compares each student's achievement with the average achievement of the class as a whole

2. ___ Comparison of student achievement with a standard selected or established by the teacher
3. ____Comparator of a student's achievement with his ability to achieve
4. ____Other than the preceding three. Will you please describe briefly on the back of this page?

VII. Please rank with the numbers 1, 2, 3, 4, and 5 the order of importance of the following functions of the grading system in your school.

_____Inform the student and parents of the quality of the student's work
_____Stimulate the student to greater effort in his learning activities
_____Diagnose student strengths and weaknesses in order to provide for more informed counseling
_____Provide a record of the student's achievement useful to colleges or employers, transfer to other high schools, etc.
_____A disciplinary measure in cases of misbehavior, or as a deterrent to misbehavior.

VIII. Has there occurred in your school during your tenure as principal, or since 1958, any change in your marking practices? Yes____ No____ If yes, who provided the initiative for this change?
(Check one or more of the following as is appropriate)
1. ____School board
2. ____Superintendent or Assistant Superintendent
3. ____Director of Curriculum or Director of Secondary Education
4. ____Principal
5. ____Counselor
6. ____Curriculum consultant
7. ____Teacher committee or teacher workshop group
8. ____Parent-teacher committee, group, or organization
9. ____Parent group
10. ____Student group or student body as a whole

IX. If your answer was "yes" in VIII above, why was this change made?
1. ____Reduce variability in grades
2. ____Increase objectivity of grades
3. ____Provide for differentiation of grades in enriched and/or slow-learner classes
4. ____Bring marking practices into line with the objectives of the curriculum
5. ____Other reasons. Please explain on the back of this page.

X. If your answer was "yes" in VIII above, what was the source of the idea that suggested the change that was made?
1. ____Professional journal or recent professional book
2. ____Faculty committee
3. ____Teachers' workshops
4. ____In-service education program
5. Observation of or information regarding a change made in another school system
6. Central office staff
7. Consultant
8. Principal
9. Counselor
10. Teacher
11. Other source. Please list on the back of this page.

XI. If your answer was "yes" in VIII above, what means were used to implement this change? (Check one or more of the following as is appropriate)
1. Directive from the central office
2. Directive from principal's office
3. Vote by, or agreement by, the faculty
4. Study project within the school or school system
5. Meetings or workshops to inform or orient teachers
6. Public relations program to inform or orient parents
7. Other means. Please add on the back of this page.

XII. If your answer was "yes" to VIII above, which persons or groups do you feel were actively supportive of the change?
1. Board of Education
2. Superintendent
3. Central office staff (Curriculum Director, etc.)
4. Principal
5. Counselors
6. Teachers as a whole
7. A committee or certain group of teachers
8. Parent-teacher group
9. Parent group
10. Students

XIII. If your answer was "yes" to VIII above, what persons or groups resisted or actively opposed the change?
1. Board of Education
2. Superintendent
3. Central office staff (Curriculum Director, etc.)
4. Principal
5. Counselors
6. Teachers as a whole
7. A committee or certain group of teachers
8. Parent-teacher group
9. Parent group
10. Students
XIV. Will you comment briefly on what changes you would like to make in the marking practices or grading system of your school? How would you implement such changes?
APPENDIX C
January 17, 1966

Dear Mr.----:

Are you, like so many thoughtful educators, dissatisfied with the traditional and archaic marking practices most schools employ? Have you wondered what improvements might be made, and how? These concerns are the topic of a doctoral study currently being conducted at The Ohio State University.

The enclosed questionnaire is an instrument designed to provide certain information required for research in this area. It is deliberately brief and will require a minimum effort on your part. Through this means we hope to encourage the highest possible number of responses. The nature of the study is such that the number of participants included cannot be large; we therefore must depend heavily on receiving a response from each principal selected to participate.

Also important to our research into this topic are several other items which you have readily available. Will you include, along with your questionnaire, the report card you use at the senior high school level, a copy of any written policies you may have regarding your marking practices, and the findings of any studies or surveys on marking practices you may have made in your school? A stamped addressed manila envelope is included for your convenience in transmitting these items.

It is anticipated that the findings of this study will be of general interest to school administrators. At the completion of the study, we intend to send an abstract of the findings to each participant. Meanwhile, your cooperation as a fellow professional is very necessary and much appreciated.

Can we receive the requested items from you at your earliest convenience?

Sincerely,

Richard H. Kindsvatter
Graduate Student in Education

Hugh D. Laughlin, Professor of Education
The Ohio State University
BIBLIOGRAPHY

Books


**Periodicals**


34. Melby, Ernest D. From an untitled paragraph, Phi Delta Kappan, Vol. 46 (May, 1965), 449.


Other Sources


