INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

75-7258

STELLER, Bonnie J., 1928-
THE MARKING PROCEDURES USED BY PUBLIC SCHOOL TEACHERS IN THE STATE OF MICHIGAN.

Michigan State University, Ph.D., 1974
Education, theory and practice

Xerox University Microfilms, Ann Arbor, Michigan 48106

THE MARKING PROCEDURES USED BY PUBLIC SCHOOL TEACHERS IN THE STATE OF MICHIGAN

By

Bonnie J. Steller

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services, and Educational Psychology

1974

ABSTRACT

THE MARKING PROCEDURES USED BY PUBLIC SCHOOL TEACHERS IN THE STATE OF MICHIGAN

By

Bonnie J. Steller

This study dealt with the question of establishing the status quo in the State of Michigan regarding the procedures used by teachers when assigning marks to students as indicators of their performance on a portion of the school curriculum. The intention was to define differences in the relationships between thirteen student characteristics and the teachers' personal characteristics, situational factors including subject area and grade level taught, the teachers' attitudes regarding the appropriate goals for education and the functions of marks, and other procedures that are associated with marking.

A total sample of 1022 teachers representing 140 school districts and 511 schools were randomly selected to receive a mailed questionnaire. A response rate of 70.15 percent was obtained.

Essentially canonical correlation analyses and chi square tests of independence were utilized for the statistical analysis of the data.

The results indicated that teachers can be categorized into two groups on the basis of whether the marks they assign are based upon objective information or upon subjective estimations. The teachers' attitudes regarding the appropriate goals for education and the usefulness of marks to teachers were found to be the basic determinants of the relative importance associated with each of the student characteristics. Approximately equal numbers of teachers reported that effort and test marks were of greatest importance. The majority of teachers reported that they base marks on a combination of subject matter mastery and the students' growth. However, those teachers who favored objective marks were inclined to prefer to base marks on relative performance compared to others or a combination of relative performance and subject matter mastery.

A profile of the two groups of teachers demonstrated that male teachers more frequently reported that the marks they assign are derived from objective information. In this group were also more younger teachers and teachers of upper grade students and academic subject areas. These teachers reported that the basic goal of education should be to provide students with a sound framework of substantive knowledge. The other group of teachers included primarily older female teachers of either lower grade level students or secondary nonacademic areas. These teachers reported that the

primary function of education should be to socialize the children rather than to instill knowledge.

Additional findings revealed that the majority of school districts provided a dual marking system with reports issued to parents and students at six week intervals. The majority of teachers reported that they were satisfied with the present marking system which for many elementary teachers included parent-teacher conferences.

An important concern was raised as a result of this study. Teachers, for the most part, were found to be incapable of defining precisely those tasks that are involved during the process of assigning marks. It, therefore, is apparent that not only do the resultant marks lack reliability but also that teachers, for this reason, cannot defend or explain the assigned marks.

ACKNOWLEDGMENTS

I would like to thank the members of the faculty at Michigan State University who contributed to my overall education. Particularly, I would like to thank the members of my doctoral committee who contributed their time and knowledge to assist me in the design and execution of this research. Included on this committee were Dr. Maryellen McSweeney, Department of Educational Psychology, Statistics and Research Design; Dr. Andrew Porter, Office of Research Consultation; Dr. Charles Eberly, Office of Evaluative Services; Dr. Ellen Strommen, Department of Psychology; and Dr. Robert Ebel, Department of Educational Psychology, Measurement and Evaluation. I would like to offer a special thanks to Dr. Ebel, the committee chairman, without whose guidance this research would have been much more difficult.

I would also like to thank Paula R. Knepper for providing both psychological encouragement and physical assistance during this period of research.

My family also deserves a thanks for their encouragement and patience.

Also a special thanks to the typist without whose services this research would have remained known only to its author.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

THE MARKING PROCEDURES USED BY PUBLIC SCHOOL TEACHERS IN THE STATE OF MICHIGAN

The practice of assigning marks to students, as indicators of their academic and/or social performance, has become an established tradition at all levels of education from kindergarten through graduate school. However, despite the universality of the practice, there has been an associated, continuing dissatisfaction, expressed both publicly and privately, regarding particularly the reliability of marks and the question of whether they are accurate indicators of students' abilities.

In 1918 when E. L. Thorndike made his statement that "whatever exists at all exists in some amount," he was referring to the belief that achievement, just like other characteristics of man, can be measured quantitatively. A further addition to this statement was advanced in 1939 by William A. McCall, who added the following: "Anything that exists in amount can be measured."

The marking of student performance is at least partially based upon the above stated assumption. Each teacher is entrusted with the responsibility of providing at specified

intervals a mark or group of marks that is an accurate and reliable indicator of each student's performance relative to a portion of the school curriculum. Each mark is then recorded on a permanent file, which is an accumulated record of every mark assigned by teachers during the student's entire school career and is referenced by college admissions officers and future employers for many years. These marks are intended, also, to provide immediate and useful knowledge to the student and his parents concerning current academic performance.

One fallacy, however, becomes obvious when comparisons are made between marks assigned a student. Are marks assigned by different teachers comparable? The answer, of course, is expected by the student and his parents to be "yes."

Background of the Problem

In the second decade of this century Starch and Elliott presented sound evidence concerning the lack of reliability of teachers' marks.[1,2,3] About the same time Kelly verified that not only did different teachers frequently assign different marks to the same student papers but also that specific marks held widely diverse interpretations for

---

[1]David Starch and E. C. Elliott, "Reliability of the Grading of High School Work in English," School Review, XX (1912), 442-457.

[2]Starch and Elliott, "Reliability of Grading Work in History," School Review, XXI (1913), 676-681.

[3]Starch and Elliott, "Reliability of Grading Work in Mathematics," School Review, XXI (1913), 254-259.

different teachers. Marks earned for equivalent work might vary from G, good, to F-, fair minus, within a single elementary school or from 70 percent to over 81 percent between different high schools.[4]

Two major concepts may explain at least a part of this variation in the mark assigned by different teachers under similar circumstances. The first such concept is the lack of agreement regarding the basis or framework within which marks are to be assigned as has been demonstrated by several studies. The second concept related to the lack of comparability of marks concerns the student characteristics that are represented by the marks.

Thorndike refers to this first difference as a difference in the frame of reference according to which a mark is expressed. The first frame of reference was termed performance relative to potential and the second termed mastery. The remaining frame of reference is performance relative to a group, national, school, class, or a group defined by the particular teacher as typical. The first frame of reference he rejects because teachers are unable to assess with any degree of reliability the "gap between potential and competence." Marks assigned on the basis of mastery are suitable to only a few educational situations. On the other hand, there are difficulties associated with attempts to mark relative to

---

[4]Frederick James Kelly, Teachers' Marks: Their Variability and Standardization (New York: Teachers' College Press, 1914).

some group. Local autonomy virtually prohibits the use of national or even state norms. At the same time values and philosophical issues complicate attempts to develop school norms. Many teachers utilize their own concept of the "typical class," an undefinable ideal, as a class norm. Thorndike, for these reasons, does not recommend the application of any of the alternatives.[5]

If, in fact, the major cause of the lack of comparability between marks is the absence of agreement regarding the basis on which marks are to be assigned, policies at either the school level or district level, specifying the basis to be utilized by all teachers, would decrease the difference. Yet Terwillinger found variability not only between schools, with the majority of policies stating the basis as either absolute mastery of subject matter or achievement with regard to ability, but also between departments and teachers within these schools. While over 40 percent of the teachers reported that they based marks on achievement with regard to ability, mathematics, foreign language, and business teachers frequently reported absolute subject matter mastery as a basis. Art and physical education teachers reported a preference for achievement with regard to class and self

---

[5]Robert L. Thorndike, "Marks and Marking Systems," Encyclopedia of Educational Research, edited by Robert L. Ebel (New York: The Macmillan Company, 1970), pp. 759-766.

improvement was the basis for marks assigned by agriculture, home economics, and driver's education teachers.[6]

Apparently even though a policy is established at the school level, teachers do not feel bound to follow these recommendations and, therefore, marks remain not comparable even within those schools which have established policies regarding the basis upon which marks are to be assigned.

The other concept responsible for the lack of comparability of marks concerns the student characteristics that are represented by the marks. Palm and Herron both reported that some teachers intend to consider such factors as effort and attitudes when assigning marks, while others do so unconsciously.[7,8] Similar findings were made by Owen and Proffitt who both showed that many diverse concepts such as accuracy, regularity, the ability to apply information, subject matter mastery, and the ability to study independently might be included in a single mark.[9,10]

---

[6]James S. Terwillinger, "A Survey of Secondary School Marking Practices and Policies," National Association of Secondary School Principals, L (March, 1966), 1-37.

[7]R. R. Palm, "A Proposed Secondary School Uniform State Marking System," Minn. J. Ed., XV (1935), 213-215.

[8]J. S. Herron, "How Teachers Rate Their Pupils," Elementary School Principals Bulletin, VIII (1929), 235-239.

[9]M. M. Proffitt, Grading in Industrial Schools and Classes, Cir. 28 (Washington, D.C.: U.S. Office of Education, 1929).

[10]W. B. Owen, "Making the Grade," Nation's Schools, XVIII (August, 1936), 21-22.

Terwillinger found these variations to be largely accounted for by the subject area which the teacher teaches. Over 50 percent of the studied teachers reported that they made use of weighting formulas when assigning marks. Business teachers allocated the greatest weight to daily work; mathematics and science teachers allocated the most weight to test marks; language arts and foreign language teachers allocated the greatest weight to projects, while others gave equal weight to daily work, test marks, and projects. Significant differences were also found between teachers of academic and non-academic subject areas in their consideration of behavior, attendance, and punctuality.[11]

In an attempt to encourage teachers to limit the student characteristics, that are included in marks, to cognitive factors, dual marking systems are frequently employed. Such a dual system provides for the assignment of separate and distinct marks to represent affective characteristics in addition to the usual academic marks.

In an investigation of the effects of single and dual marking systems on the academic marks of gifted sixth grade pupils, Housdorff and Farr reported no associated differences, thereby, concluding that a dual system provides no advantage over a single mark.[12]

---

[11] Terwillinger, loc. cit.

[12] Henry Housdorff and S. David Farr, "The Effects of Grading Practices on the Marks of Sixth Grade Gifted Children," J. Ed. Res., LIX (December 1965), 169-172.

Similar conclusions were drawn by Halliwell and Robitaille, when they reported a significant correlation of .68 between assigned academic marks and a subjective mark representing effort. If a dual system is to increase the reliability and validity of academic marks, teachers must be provided with adequate methods for the assessment of affective factors.[13]

The above mentioned evidence demonstrates that variations do occur among teachers in regard to the bases upon which marks are assigned and the student characteristics that are considered when assigning marks. Yet little is known about what factors are related to the various procedures used by teachers when they assign marks to students.

The majority of studies dealing with the question of how teachers assign marks to students were conducted during the third and fourth decades of the present century. However, since that time numerous changes have taken place which may have had a significant impact on these procedures.

Not only are teachers required to have an increased number of years of training but also the content of this training has undergone major revision. The expanded knowledge of human growth and behavior has made teachers increasingly aware of their contributions to students' development.

---

[13]Joseph W. Halliwell and Joseph P. Robitaille, "The Relationship between Theory and Practice in a Dual Reporting Program," J. Ed. Res., LVII (November, 1963), 137-141.

Subject matter expertise is no longer considered to be sufficient to produce the desired results within the students.

At the same time students are expected to acquire during their school careers a more complete and comprehensive body of knowledge than ever before. However, concerns have also been voiced that the educational processes can have serious, at times detrimental, unintended effects on the students.

These concerns have led to the increased criticism of education, a part of which has been directed toward the practice of assigning marks to students. Critics have strongly and repeatedly recommended that marks be completely eliminated from the schools.

On the other hand, the supporters of marking claim that marks can serve important functions to the teacher, the student, and the parents. Other techniques for the accomplishment of these functions are not yet available. To eliminate marks would be to create within the school an artificial situation, wherein, students are not provided the means for judging their own strengths and weaknesses, an essential capacity for adult activities.

The supporters of marking have proposed numerous suggestions regarding changes which would, they believe, improve marking procedures and practices. However, these suggestions have dealt, for the most part, either with mechanical considerations of form or with theoretical considerations.

General Purpose of this Study

The purpose of this study is to examine the current situation in the State of Michigan in regard to the framework within which marks are assigned and the composition of marks. Although no direct attempt is made to replicate the earlier studies conducted during the third and fourth decades of the twentieth century, this study deals with many of the same issues and, therefore, references will be made to these studies when appropriate.

These issues deal with such topics as the comparability of marks, the objectivity versus the subjectivity of marking, and the influence of teachers' personal characteristics and situational factors to the framework within which marks are assigned and the composition of marks.

Such information as is provided will be useful in a number of contexts. The first and most important context is in the design of alternatives to marking or in the modification of existing systems. As the emphasis on continuing life-long education increases, in conjunction with the requirements of increased technological knowledge and skills, more and more often there is a need for reliable indicators of individual's strengths and weaknesses in regard to academic and technical skills. Even with the acknowledged low reliability of marks, college admissions officers must base at least a significant portion of their decisions on marks, as other means are lacking.

In the context of schools or school districts such information will be useful in the establishment of guidelines and the determination of policies regarding the assignment of marks. Although institutional policies have in the past been tried and have failed, because of the lack of quality control,[14] if properly administered, such a system can insure the comparability of marks at least within its jurisdiction.

In a further context that of the colleges of education such information can serve two purposes. The first purpose is to revitalize the study of student marking instead of its condemnation based upon incomplete evidence. The second purpose is its contribution to the training of prospective teachers and for in-service training of teachers.

In the context of the in-service teachers, such information would assist them in the formulation and definition of their own marking procedures and, at the same time, aid in their interpretation to parents and students of assigned marks.

Specific Objectives of This Study

The major goal of this study is to derive an answer that is an accurate representation of the current situation as reported by teachers in the State of Michigan to the following question: How are teachers assigning marks to students?

---

[14]Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs: Prentice-Hall, Inc., 1972), pp. 337-338.

This question will be answered by synthesizing the responses to the following associated questions:

What personal factors are significantly related to the weighting applied to student characteristics that are considered when assigning marks?

What personal factors are significantly related to the basis on which marks are assigned?

What situational factors are significantly related to the weighting applied to student characteristics considered when assigning marks?

What situational factors are significantly related to the basis on which marks are assigned?

What attitudinal factors are significantly related to the weighting applied to student characteristics considered when assigning marks?

What attitudinal factors are significantly related to the basis on which marks are assigned?

What other aspects of the marking procedure used are significantly related to the weighting applied to student characteristics considered when assigning marks?

What other aspects of the marking procedure used are significantly related to the basis on which marks are assigned?

Do teachers agree in regard to the relative importance of student characteristics considered when assigning marks?

Do teachers agree in regard to the basis on which marks are assigned?

What types of objective measures of student performance are made by teachers?

How frequently are objective measurement of student performance made by teachers?

What changes do teachers suggest regarding marking procedures and policies?

How do school districts differ in their marking policies?

Definition of Terms

Assigned marks: This refers to the letter, number, or other symbol that appears on a report card or student progress report, which is intended to provide the student and/or parents at regular time intervals with information regarding the academic progress or relative standing of the student.

Attitudinal factors: This refers to the teacher's expressed attitudes regarding the following: (1) the appropriate goals of education, (2) the functions that marks can successfully serve, and (3) suggested changes or improvements in marking procedures and policies.

Basis for Marks: This refers to the frame of reference within which marks are assigned. Three possible bases are included: (1) mastery of subject matter, (2) student growth when compared to himself, and (3) student performance relative to other class members.

Marking policies: This refers to institutional requirements within which framework the teacher must operate when assigning marks. Included are the following: (1) frequency of marking, (2) presence or absence of a dual marking system, (3) types of symbols used to indicate marks, and (4) the overall format of the reporting form.

Marking practices: This refers to the methods utilized by teachers when assigning marks to students and includes the following elements: (1) student characteristics

that are considered by the teacher when making the decision to assign a particular mark, (2) relative importance or weight assigned each of these student characteristics, (3) type and frequency of objective measures of student performance, (4) frequency of use of particular statistical techniques when assigning marks, and (5) sources of useful information in the decision-making process related to marking.

Situational factors: To be included in this category of interest are the following: (1) size of the school district, (2) size of the school, (3) urbanization level of the community, (4) percent of minority student within the school district, (5) grade levels included in the school, (6) grade level taught, and (7) subject area taught.

Personal factors: To be included are the following characteristics of the teacher: (1) sex, (2) age, (3) teaching experience as measured by the number of years taught, (4) college degrees held, (5) self estimation of the level of academic success achieved, and (6) number of college courses completed in the area of grading, measurement, or evaluation.

Size: Size both in association with schools and school districts refers to the number of teachers employed.

Student characteristics: To be included in this category are those factors that are considered by teachers

when assigning marks. These factors are as follows: (1) attendance, (2) personal appearance, (3) behavior, (4) in-class participation, (5) amount of effort put forth, (6) attitudes, (7) neatness of work, (8) homework, (9) test marks, (10) quiz marks, (11) optional work, (12) group reports or projects, and (13) individual reports or projects.

Guiding Principles and Assumptions

The following set of principles and assumptions form the framework within which this study was conducted.

1. The major function of marks is to provide a reliable and accurate indication of student performance in regard to a portion of the school curriculum.
2. Teachers want to assign reliable and accurate marks.
3. The more reliable a mark is, the more valid it is as an indicator of student performance.
4. Marks based upon objective measurements are more reliable than are marks based upon subjective estimations.
5. Measurements which are comprised of objective formatted items are more reliable under normal school conditions than are measurements comprised of other types of items.
6. A reliable mark is one which does not vary significantly for the same work between points in time or between different teachers.

Overview

This chapter has presented the problem statement as well as the objectives to be accomplished by the study.

Chapter II includes a review of the major literature in the field. Chapter III includes a summary of the procedures that were utilized. Chapter IV presents the data analysis and findings. Chapter V presents the conclusions and recommendations which were derived from the study.

CHAPTER II

REVIEW OF THE LITERATURE

The review of literature for this study included an examination of the philosophical and methodological arguments which have been presented in regard to the practice of assigning marks to students. Both the critics' and the defenders' points of view were examined in regard to these arguments in order to present an unbiased review, although the author's biases may, nevertheless, at times be evident.

This review of literature included many of the well-known writings on the subject as well as several of the lesser known articles. Because of the large number of articles and books, which have been written on the subject of marking since 1910, only a limited amount of the material has been presented in this chapter. The material presented was chosen for the most part because of its relevance to the variables, investigated by this study. For this reason, the chapter has been primarily organized according to the variables under investigation, with the exception of the section dealing with the historical background of the problem and the philosophical arguments for and against marking.

Historical Background

In the early period of education the direction of evaluation was the reverse of that taken today. The teachers were evaluated in terms of the performance of their students. Teachers were held accountable for the success or failure of their students, a justifiable procedure as long as the criteria for success were highly visible. However, as the requirements for success became complex, this method was no longer practicable.

By the mid-nineteenth century formal education as preparation for college had gained prominence among the upper class families. However, college admission depended upon the family's social status and their financial ability rather than upon the prospective student's scholastic abilities. The use of examinations was considered beneficial primarily as indicators of the student's readiness to leave the college but not in regard to admission.

With the rapid growth between 1870 and 1910 in the number of students enrolled in public elementary and high schools, new methods of informing students as to their relative performance were introduced. By 1910, however, the written reporting of student performance had gained almost universal application in the United States. Although many elementary teachers provided written descriptions of the performance of each individual student, virtually all high

school teachers used percentages or other similar markings to report the students' performance in the different subject areas. During the next few years elementary teachers also gradually adopted point systems or percentage marking procedures. Nevertheless, the search for improved methods was continued even to the present. Not only has a single "best" method not been developed, but there is doubt that such a method is feasible or possible.

Table 2.1 shows the changes which have taken place within the Philadelphia Public School System between 1913 and 1961 in regard to the assignment of marks.[1]

Many of these changes were in response to the criticisms made of marks and marking procedures, ranging from philosophical arguments to methodological and technical critiques.

Philosophical Arguments Regarding Marking

One philosophical argument against the practice of assigning marks to a student in order to convey information regarding the student's relative performance has been primarily concerned with the question of whether one individual had the right to make judgments regarding the performance of another individual.

---

[1]Howard Kirschenbaum, Sidney B. Simon, and Rodney W. Napier, Wad-ja-get? (New York: Hart Publishing Co., Inc., 1971), pp. 45-54.

TABLE 2.1.--Changes in the Assignment of Marks from 1913 to 1961 in the Philadelphia Public School System.[a]

Year	Elementary	High School
1910 to 1920	10 point A single overall mark 10 distinguished, 9 excellent, 8 good, 7 fair, 6 poor, 5-1 very poor	Percentage for each subject
1920 to 1930	5 point	Percentage equivalent 91-100 excellent, 81-90 good, 71-80 fair, 61-70 passing, below 61 deficient
1930 to 1940	3 point A outstanding, B satisfactory, C needs improvement	5 point Dropped percentage equivalent
1940 to 1950	2 point S satisfactory, U unsatisfactory plus trait checklist 4 point Added 0 outstanding, I improvement to above system	Percentage equivalent 90-100 A, 80-89 B, 75-79 C, 70-74 D, 60-69 E, below 60 F
1950 to 1960	5 point A, B, C, D, F	5 point A, B, C, D, E including grades for behavior and work habits
1961	Added checklist in reading and arithmetic to above 5 point system	

[a]Howard Kirschenbaum, Sidney B. Simon, and Rodney W. Napier, Wad-ja-get? (New York: Hart Publishing Co., Inc., 1971), pp. 45-54.

Pewitt's statement that ". . . the real, but unadmitted purpose . . . is to maintain a caste system of superiority over inferiority, educated over uneducated, teacher over student, adult over child," is an example of this point of view.[2] Complete equality was demanded for the student, insisting that the teacher had no moral authority to judge the students' performances. Increased capabilities through education or even maturation were denied. For the teacher to judge the student was to label one individual superior over another.

On the other hand, Dunivan in his rebuttal to Pewitt stated that marking does not create a caste system because the individuals do not remain permanently fixed within a given station but may advance either forward or backward. Teachers have ". . . no compulsion to use grading to prove their superiority over students." The pattern of authority of which marking is only a small part should have no detrimental effect on students, because this is a natural phenomena of all life. "Each of us has at least one person who has authority over him."[3]

From the same point of view Becknell commented that any attempt at insuring absolute equality will fail because

---

[2]Joe Pewitt, "Grading as a System of Coercion," School and Community, LVI (Feb., 1970), 20.

[3]Lindell P. Dunivan, "Grading as a System of Guidance: A Rebuttal," School and Community, LVI (May, 1970), 31.

"human beings will always build up social barriers between groups." If the role of the school is to provide the opportunity for its students to develop into good citizens, the school should create an environment representative of adult life. Because "adults are graded in everything they do," the school should provide an environment in which the students learn to be aware of their own strengths and weaknesses. Poor students should not be deceived into thinking that their work is as good as that of the A students.[4]

A satisfactory answer to either of these points of view is impossible. The school exists as a part of the surrounding society and, therefore, must operate within the same framework, including both freedoms and limitations. A society in which judgments are made and estimations of worth provided will, in most cases, support a school which utilizes similar techniques. Therefore, to criticize teachers for marking students is to criticize the similar practice with the society of which it is a part.

Another philosophical argument deals with the issue of external reward. Johnson commented that grades do not encourage students to think for themselves. They learn to memorize answers, thereby, the natural love of learning for personal satisfaction is destroyed.[5]

---

[4]George L. Becknell, "S and U Grading System," Texas Outlook, XXV (July, 1941), 18.

[5]Icie F. Johnson, "Injustice of Grades," School and Community, LIV (October, 1967), 24-25.

Such a statement requires the assumption that the amount of learning which takes place is not related to the grade received. If this is, in fact, the situation, then the problem lies not with the mark itself, but with the procedures used in the assignment of the mark or with the teaching process. Marks, which are based upon sufficient, relevant, important evidence, are valid indicators of what and how much the students have learned.[6]

On the other hand, De Zouche in 1945 presented an argument against the need of external rewards.

> Does everything under God's heaven have to be done for the sake of winning? Can it never be done for the sake of the thing itself? Life is not a matter of winning or losing. Life is a matter of living, simply, courageously, and happily with our fellow creatures. Must we be rewarded, tagged with a blue ribbon for every worthy effort we make? Even adults have to be whiddled with an award. The government gives us an E if our war production reaches a certain point. The Red Cross hands us a sticker to put in our windows to announce to our neighbors that we gave a donation. The Tuberculousis Society clips a button to our lapel for the same reason. And in St. Louis the Community Chest gives us a dear little red feather to wear in our headbands as soon as we've sent in our checks.
>
> Perhaps if we stopped giving children a gold star for brushing their teeth, we wouldn't be giving adults red feathers for contributing to the Community Chest.[7]

---

[6]Robert L. Ebel, Essentials of Educational Measurement (Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1972), pp. 313-315.

[7]Dorothy De Zouche, "The Wound is Mortal," Clearing House, XIX (February, 1945), 339-344.

The difficulty lies in the interpretation of marks as rewards. Although marks may provide either a sense of satisfaction or of insufficiency, their primary function is to report, in a factual manner, the student's level of performance. High marks should not be intended as rewards or low marks intended as punishments.

Methodological Arguments Regarding Marking

The earliest studies which pointed out the low reliability of marks were those conducted by Starch and Elliott in 1912.[8] It must be remembered, however, that until after 1920 marks were based upon essays, which under normal conditions, seldom have been shown to have acceptably high reliabilities.

Dissatisfaction with the then-commonly-used practice of assigning percentage equivalent marks led the College High School of Colorado State University at Greeley under the direction of William Wrinkle to search for improved methods of marking and reporting. Wrinkle reported six fallacies associated with marking.

1. Anyone can tell what a single A B C D F mark means.
2. A student can achieve any mark he wishes--if he is willing to make the effort.
3. The student's success in his after-school life compares favorably with his success in school.

---

[8]Starch and Elliott, op. cit.

4. The student's mark is comparable to the worker's paycheck.

5. The competitive marking system provides a worthwhile and justifiable introduction to competitive adult life.

6. School marks can be used as a means to an end without their becoming thought of by students as ends in themselves.

Although these fallacies are philosophical in nature, their involvement in the search for an improved method according to Wrinkle can lead to several practical results and findings, several of which are presented below in an abbreviated form.

1. Outcomes must be thoroughly analyzed so that meanings can be clearly stated.

2. The number of different forms should be kept to a minimum.

3. The basis for student evaluation must be specific.

4. Parents are likely to misunderstand a report which includes many details.

5. If students are favorable to an idea, parents will be favorable.

6. The summarization of reports is too big a task and not workable.

7. The scale type form is not satisfactory.

8. The check form is simpler and economical of space.

9. Detailed, elaborate cumulative records are not economical.

10. Conference plan is not a practical solution at the secondary level.

11. A five-point scale should be maintained for administrative needs.

12. Check forms become increasingly detailed and, therefore, impractical.

13. Objectives should be stated in terms of desired student behavior.

14. Teachers have difficulty in writing effective comments.

15. Best way to write comments is to tell what student did.

16. Reports to parents of detailed specific objectives unnecessary.

From these findings was developed a one page report which included a four-point mark to each of twelve general objectives and dual checks for actual and expected achievement based on outstanding, above average, average, below average, and very poor.[9,10,11]

Anderson claimed that marks are not constructive methods of communicating with parents for three reasons: (1) marks are too general, (2) marks are too impersonal, and (3) marks provide for only one-way communication.[12]

---

[9]William Wrinkle, "The Story of a Secondary School Experiment in Marking and Reporting," Educational Administration and Supervision, XXIII (October, 1937), 481-500.

[10]William Wrinkle, Improving Marking and Reporting Practices (New York: Rinehart and Co., Inc., 1947).

[11]Samuel R. Johnston, "Are There Better Ways of Evaluating, Recording, and Reporting Student Progress?," National Association of Secondary School Principals, XXXIV (March, 1950), 79, 89.

[12]Robert H. Anderson, "The Importance and Purposes of Reporting," National Elementary School Principals Bulletin, XLV (May, 1966), 6-11.

Of the same viewpoint is Melby who stated that the specific observational records usually maintained by teachers are ignored in the marking process in favor of a single meaningless symbol.[13]

These comments have dealt with the provisions and limitations set forth by school administrators in regard to the assignment of marks to students. Teachers seldom are free to make decisions regarding such matters as the number of points to be permitted, the format of the reporting form, or the frequency of reporting. Such decisions are regarded as administrative and at best will be made by a committee of teachers acting under the guidance of an administrator.

Teachers on the other hand, can wrongly affect marking in seven ways, according to Palmer. These ways are as follows: (1) by abdication, (2) by a "carrots and club" system of rewards and punishments, (3) honestly by default when a single test serves as the entire basis for a mark, (4) by being a testing zealot who advocates the survival of the fitest, (5) by changing the rules in midstream, (6) by being a psychic grader who needs no evidence, and (7) by anchoring grades on the rainbow of perfection.[14]

R. C. Little reported from a survey of teachers that, while they do desire to grade constructively in a way

---

[13]Ernest O. Melby, "It's Time for Schools to Abolish the Marking System," Nation's Schools, (1966), 104.

[14]Orville Palmer, "Seven Classic Ways of Grading Dishonestly," The English Journal, LI (1962), 464-467.

most fruitful for the growth of their pupils, the search for a single "best" procedure is likely to go on for some time.[15]

Bates proposed that the null hypothesis be tested that, "those students whose progress in reading, arithmetic, and spelling is reported by means of report cards will demonstrate no more achievement in these subjects than students who do not receive report cards."[16]

Although the philosophical arguments presented denied the existence of a simple, easy solution, for the methodological arguments, solutions are no more easily found. Kirschinbaum, Simon, and Napier in Wad-ja-get? present the advantages and disadvantages for each of the eight most frequently used alternatives to the assignment of symbols to represent the student's level of performance. These alternatives are as follows: (1) written evaluations, (2) self evaluations, (3) give marks but don't tell the students, (4) the contract system, (5) the mastery or performance approach, (6) pass/fail grading, (7) credit/no credit grading, and (8) blanket grading. When the number and severity of the disadvantages are compared to the advantages, a "best" procedure is not defined.[17]

---

[15]Ruth Coyner Little, editorial, "Whither Grading," National Education Association Journal (1947).

[16]Bea Bates, "Can We Stop the Merry-go-round of Grades?", Arizona Teacher, LV (January, 1967), 19 and 26.

[17]Kirschenbaum, op. cit., pp. 292-307.

Therefore, the methodological arguments also remain unanswered for the present. An investigation of the current situation will provide at least some direction to the search for improved practices.

Basis for Marks

At least two decisions must be made by the teacher whenever a marking procedure is selected or a mark is assigned. The first decision concerns the basis on which the mark or marks are to be determined. Three possible approches have been suggested. Marks may be assigned according to the following: (1) mastery of the subject matter, (2) performance as compared to a predefined group, usually the other class members, or (3) growth when compared to past performance or demonstrated ability. Little concurrance of opinion has been exhibited in this regard.

The arguments for and against marking in relation to others have dealt with the question of whether competition has a useful or acceptable place within the school. Wrinkle claimed that because it is common knowledge that all students can not achieve any mark, that competitive marking is unfair to the low achiever. Although adults do compete, this competition takes place primarily among equals. Also in adult life there is more cooperation than there is competition at least between unequals.[18]

---

[18]Wrinkle, Improving, pp. 36-49.

In a reaction to this point of view Schwartz, Tiedman, and Wallace stated that "the ability to achieve high marks is just one of many ways in which pupils vary and the elimination of (competitive) marking will not thereby create a Utopian institution in which all pupils work together on a basis of complete equality."[19]

In answer to fear voiced by teachers that, students who are competing only against each other and not against some absolute standard, will mutually agree to take it easy together, Maxon replied that intense competition occurs when an individual works with his peers. In the American free enterprise system competitors do not mutually agree to take it easy nor do athletes when attempting to make the team.[20]

Fine showed that while elementary teachers often avoid the use of competitive marking procedures, secondary school guidance counselors insist upon this traditional method because of the weight assigned marks by college admissions officers.[21]

---

[19]Alfred Schwarts, Stuart C. Tiedman, and Donald L. Wallace, Evaluating Student Progress in the Secondary School (New York: David McKay Co., Inc., 1957), pp. 391-392.

[20]W. B. Maxon, "Grading a Serious Matter," National Education Association Journal, XIII (October, 1964), 56-57.

[21]Benjamin Fine, "A.B.C. of Grading Puzzles Parents," New York Times Magazine, November 18, 1957, cited by Ann Z. Smith and John E. Dobbin, "Marks and Marking Systems," Encyclopedia of Educational Research, 3rd ed., edited by Chester W. Harris (New York: The MacMillan Co., 1960), pp. 783-791.

If marks are to be based upon performance relative to a predefined group, the decision must be made as to whether a national, state, school, or class group will be used. Sufficient information is seldom available for the use of either a national or state group. At the same time there is seldom sufficient cooperation at the school level to make such a comparison possible. For this reason as well as for the sake of convenience the class, either in actuality or as a vaguely ideal class, is the most-often-used basis for comparison. Yet such a group is seldom large enough to provide adequate information.

One alternative basis for marks is mastery of subject matter. This position was favored by Marshall. However, he claimed that the marks of subject matter mastery should clearly only include content areas per se not other judgments.[22]

Wrinkle also agrees that if marks are influenced by other extraneous factors, they become meaningless. He goes on to say, on the other hand, that marks which include only academic information are significant only insofar as the academic information is commonly recognized.[23]

According to Ebel the primary objective of education is to assist students in the acquisition of a command of

---

[22]John C. Marshall, "Evaluation: Does it Mean Anything?", Clearing House, XLIV (May, 1968), 535-538.

[23]Wrinkle, "The story of," op. cit.

substantive knowledge.[24] If this is the case, then a mark should represent the degree to which this objective has been attained, i.e., how much knowledge the student possesses. However, if a different objective has been selected, the basis of marks should also be adjusted.

Both Linder and Rothney maintain that the basis for marks is not as important as is the fact that the students should know and understand what basis is to be used in determining marks.[25,26]

Seldom has a recommendation been made that marks should be based upon the student's own growth as based upon past performance or capacity. Although the statement that the reliability of growth scores is extremely low, as the variance error of a difference score is the sum of the variance errors of component scores, has often been made, it is highly unlikely that this can totally explain the lack of these recommendations.

On the other hand, Melby stated that teachers should evaluate in terms of capacity and growth.[27] However, evaluation and marking need not have been simultaneously referenced.

---

[24]Ebel, op. cit., pp. 55-69.

[25]Ivan Linder, "Is There a Substitute for Teachers' Grades?," American School Board Journal (July, 1940), 25-26.

[26]John W. M. Rothney, Evaluating and Reporting Pupil Progress (Washington, D.C.: National Education Association Press, 1955).

[27]Melby, op. cit.

Student Characteristics Considered

Teachers also must decide when selecting a marking procedure or assigning a mark what student characteristics are to be represented by the mark. While the inclusion of both academic and affective information often occurs, the recommendation has been made by Marshall that grades be limited in their composition in order to make them comparable between students. A system of grading should be used that would enable one to know that an A student is in the upper group of his class, a grade of B would indicate that the student is in the second group, etc.[28]

Toliem reported that the majority of teachers include in marks certain characteristics of students, while other characteristics are included less often. Table 2.2 indicates those characteristics that were frequently reported and the percent of teachers who claimed to consider each. It should be noted, however, that several of the factors listed were previously referred to as bases for marks rather than as student characteristics.[29]

Variations in the characteristics represented by a mark occur not only in regard to the inclusion of affective

---

[28]Marshall, op. cit.

[29]William Joseph Toliem, "Factors which Affect the Decision Making of Classroom Teachers as they Evaluate Students" (Unpublished Ph.D. dissertation, State University of New York at Buffalo, 1972).

TABLE 2.2.--Factors Reported Considered by Teachers and the Percent of Teachers Reporting Each.

Factor	Percent
Progress when compared to own growth	92.1
Student interests	84.2
Class participation	82.9
Observation of student	82.9
Student-teacher conferences	80.3
Classroom tests and quizzes	77.6
Classroom behavior	75.0
Self evaluation by student	61.8
Homework	57.9
Type of group	54.0
Progress when compared to others	44.7
Reports by counselors, etc.	43.4
Observation of student plus logs	39.5
Achievement test results	32.9
Anecdotal records	19.7
Rating scales	11.8
Opinions of other teachers	6.5

factors but also in regard to the measures of cognitive skills on which the mark is determined and the relative importance attached to each.

In a study by Russell and Thalman of 335 seventh and eighth grade students, a correlation of .658 was found between average marks and the personality rating made by the teachers. The multiple correlation obtained when personality, ratings, achievement test results, and intelligence test results were correlated with average marks was .810, which indicated that 65 percent of the variance in teachers' marks could be explained by the linear relation of the three measures to the marks. When teachers' marks were categorized according to A, B, C, D, and E, significant differences were found in the personality ratings for the five groups.

Russell and Thalman concluded that:

> If the findings of this study are true, they should be recognized and either accepted or corrected in terms of our educational philosophy . . . . Serious and permanent damage to a pupil's personality can result from continued failure in school; and if the mark results from a personality conflict between the teacher and the pupil, the act is cruel and unjustified. A challenge is made to teachers to guard against prejudice and to be alert for personality problems which may cause the pupils who have them to function at a level lower than they might achieve. Recognize those problems for what they are, but avoid allowing them to appear in the disguise of a teacher's mark.[30]

---

[30]Ivan L. Russell and W. A. Thalman, "Personality: Does it Influence Teachers' Marks?," Journal of Educational Research, SLVIII (April, 1955), 561-564.

On the other hand, in an issue of What Research Says to the Teacher, a publication of the National Education Association, Ruthney recommended that the 3C's, character, cooperation, and courtesy, should be a part of every mark assigned.[31] However, clues are not provided as to how these personality traits are to be assessed. The only available instruments for personality measurement either require a highly skilled psychometrician for their administration or are subjective estimation procedures. In addition, the laws of many states, including Michigan, specifically prohibit the administration of any personality, interest, or attitude instrument without the written consent of the child's parents.

A study by Sobel demonstrated the difficulties associated with attempts to include affective factors as well as academic achievement in a mark. When marks do not represent solely the academic achievement of the students, the following results may be expected: (1) more girls than boys will receive high marks, (2) those receiving high marks would also score high on a measure of the amount of effort put forth, (3) those receiving high marks will have a better attendance record than will those receiving lower marks, and (4) those receiving high marks will be high in self confidence and have a pleasing personal appearance. Therefore,

---

[31]Ruthney, op. cit.

each mark should represent only academic achievement rather than to also include personality characteristics.[32]

Even when teachers try to provide a measure of academic performance, marks were found by Doherty to be significantly related to school attitudes, interests, effort, and citizenship,[33] while Fleming found a relationship between marks and school attitudes, chronological age, and the teacher's estimate of intelligence.[34]

Miner, in agreement with Fleming, found that variations were closely ralated to the grade level of the students being marked. Teacher assessment of student performance in the early grades was found to be actually an assessment of citizenship and/or behavior, while the assessment of secondary pupils was a measure of cognitive skills.[35]

Apparently teachers find it impossible to completely separate the affective characteristics of students from

---

[32]Frances S. Sobel, Teachers' Marks and Objective Tests as Indices of School Adjustment (New York: Teacher's College Press, 1936).

[33]William V. Doherty, "A Survey of the Evaluation of Pupil Progress in Selected Elementary Schools in Ohio" (Unpublished Ph.D. dissertation, Ohio State University, 1954).

[34]Cecile W. Fleming, A Detailed Analysis of Achievement in High School (New York: Teachers's College Press, 1925), p. 69.

[35]Betty Crowther Miner, "Three Factors of School Achievement," Journal of Educational Research, LX (April, 1967), 370-376.

their academic performance, when assigning marks. This inability can create serious problems relevant to the interpretation and comparability of marks. Because the measures of affective characteristics are subjective estimations of the degree to which the student possesses a particular characteristic, different characteristics may be included with unequal importance for different students. This fluctuation also will contribute significantly to the low reliability of marks.

Effect of Teachers' Personal Characteristics

The question has been raised as to whether certain personal characteristics of the teacher may contribute to the decisions made regarding the basis on which marks are to be assigned and the composition of these marks.

Metzner recently pointed out that many of the variations that occur in marks are due to differences in the degree to which the students accept the middle-class orientation of the school. This finding may explain why girls receive higher marks than boys and why students of higher socio-economic status receive the higher marks. He also found that teachers who come from lower socio-economic backgrounds exhibited the same patterns of approval and disapproval, as did those from higher class backgrounds. In

fact, the former showed more authoritarianism and were less sympathetic.[36]

Although Carter in 1930 reported a study by Newton, demonstrating that women teachers assigned higher marks than did men teachers,[37] a more recent study by Arnold presented contradictory findings, indicating that for fifth and sixth grade students, there was no evidence of differential marking procedures for men and women teachers.[38]

Theimer's investigation dealt with two other personal characteristics of teachers, their own grade point average and their teaching experience as measured by number of years taught. In neither instance was a significant relationship found to the marking procedures used.[39]

The personal characteristics of teachers according to these studies, do not significantly affect the procedures used when assigning marks. It would seem, however, that

---

[36]Seymour Metzner, "Teacher Bias in Pupil Evaluation: A Critical Analysis," Journal of Teacher Education, XXII (Spring, 1971).

[37]Robert S. Carter, "How Invalid are Marks Assigned by Teachers?." J. of Ed. Psy., XLIII (1952), 218-228.

[38]Richard Dean Arnold, "The Relationship of Teacher's Sex to Assigned Marks and Tested Achievement among Upper Elementary Grade Boys and Girls" (Unpublished Ph.D. dissertation, University of Minnesota, 1966).

[39]William Charles Theimer, "The Relationship between certain Personality and Educational Factors and the Grading Patterns of Secondary School Teachers" (Unpublished Ph.D. dissertation, University of Denver, 1967).

their training in the area of measurement would affect their procedures.

Although Mosely reported that 52 percent of the teachers studied had taken college courses dealing with research methodology, testing, and measurement,[40] Mayo found that such study taken after graduation did not explain the increase in measurement competency, demonstrated after two years of teaching experience. However, the completion of such study prior to graduation did result in moderate superiority in measurement competency on the part of beginning teachers.[41]

In a recent study concerning the requirements established by a random sample of 916 colleges and universities in regard to their elementary education curricula, Roeder reported that 57.6 percent of the institutions did not require the completion of a course in measurement or evaluation. Only 31.4 percent of the institutions required the completion of a course devoted solely to this topic. Table 2.3 presents the findings of this study in regard to required college courses.[42]

---

[40]Aubrey Howard Moseley, "A Study of Teachers' Perceptions of Factors Related to Educational Research" (Unpublished Ph.D. dissertation, University of Alabama, 1966).

[41]Samuel T. Mayo, Pre-service Preparation of Teachers in Educational Measurement (Loyola University, 1967), pp. 62-63.

[42]Harold H. Roeder, "Teacher Education Curricula--Your Final Grade is F," Journal of Educational Measurement, X (Summer, 1973), 141-143.

TABLE 2.3.--Nation-Wide Summary of Specific Methods Courses.

Course	No. Req. Sem. Hrs.		One-Two Sem. Hrs.		Three Sem. Hrs.		Four-Five Sem. Hrs.		Over Five Sem. Hrs.		Combined Course		Unscored		Total	
	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.	Freq.	Pct.
Evaluation	495	57.6	104	12.1	154	17.9	10	1.2	2	.2	62	7.2	33	3.8	860	100.0
Geography	343	39.9	45	5.2	277	32.2	47	5.5	76	8.8	16	1.9	56	6.5	860	100.0
Art Method	94	10.9	205	23.8	299	34.8	103	12.0	30	3.5	45	5.2	84	9.8	860	100.0
Music Method	109	12.7	203	23.6	278	32.2	118	13.7	34	4.0	37	4.3	81	9.4	860	100.0
Physical Education	126	14.7	232	27.0	65	7.6	261	30.3	39	4.5	9	1.0	128	14.9	860	100.0
"Religion"	447	52.0	8	.9	55	6.4	36	4.2	269	31.3	0	0.0	153	17.8	860	100.0

Studies such as the one by Roeder indicate that teachers often are inadequately trained in the area of marking. Even without the needed knowledge and skills teachers must make frequent and far reaching decisions, when assigning marks to students, about issues concerning which even the experts do not agree.

Harris recently published an investigation dealing with the degree of agreement among a group of measurement specialists regarding several of these issues about which teachers must make decisions. Table 2.4 presents a part of his findings concerning several of the issues.[43]

If experts cannot agree as to the "best" way to assign marks, the task of the teacher becomes that of devising marking procedures which uniquely best serve his or her needs. It then becomes surprising that marks are as comparable as they are.

Effect of Situational Factors

Studies concerning differences among the practices of teachers from schools and/or school districts of different sizes and different levels or urbanization are for the most part lacking. Because highly urbanized communities generally contain larger number of minorities, the study

---

[43]Wilbur S. Harris, "Agreement among N.C.M.E. Members on Selected Issues in Educational Measurement," J. Ed. Measure, X (Spring, 1973), 63-70.

TABLE 2.4.--Percent of Respondents Agreeing with, Disagreeing with, and Expressing Uncertainty Toward Selected Issues.

	Issue	Agree	Undecided	Disagree
3.	Essay tests are more suitable than objective tests for measuring most types of learning.	8	14	79
6.	Schools which use marks should adopt and enforce a clearly defined institutional marking policy.	80	9	11
8.	Students' academic marks should be determined solely by their academic achievement.	76	12	12
10.	The majority of teachers rely too much on their subjective judgements in marking students.	69	18	13
11.	About the same proportion of high marks should be given to classes of slow learners as to classes of rapid learners.	13	18	69
17.	Factors like attitudes and interests should be used deliberately in determining students' marks.	15	18	67
19.	The acquisition and utilization of knowledge are the primary aims of formal schooling.	74	7	19

TABLE 2.4.--Continued.

Issue	Agree	Undecided	Disagree
20. Most teachers use too few appropriate statistical techniques in evaluating and marking students.	76	14	19
22. "Pass-fail" or "credit-no credit" is more desirable than marking with three or more categories for academic classes.	22	20	58
23. "If something exists, it exists in some quantity, and therefore it can be measured."	64	15	21
29. Academic marks should be based more on achievement status than on growth or progress.	54	20	26
37. Teachers should attempt to evaluate and mark students in such areas as interest, attitudes, and motivation.	35	19	46
39. Absolute standards are more desirable than relative standards in evaluating and marking students.	37	17	46
40. Some type of numerical or letter marking system is essential to good educational practice.	45	15	40

by Fish, which concluded that when the effects of socio-economic factors were removed no relationship was found between racial background and marks, might be an indicator of a lack of relationship between urbanization level and the marking procedures used by the teacher.[44]

Another study, dealing with the urbanization factor, was conducted recently by Finn. He observed in a sample of urban and suburban elementary schools that teachers in the urban schools had slightly more experience than did those in the suburban schools, 10.7 years as compared to 9.4 years. This study also involved the grading of two essays by fifth grade teachers with additional false information, provided regarding the student's sex, race, past achievement, and intelligence. The following conclusion was reported.

> Teachers in the urban environment are relying to a greater extent upon information about the pupils in rating their products. While on the average, the suburban teachers rate all pupils' essays similarly regardless of race, sex, or ability information, the urban teachers do not. Yet the suburban teachers unlike their urban counterparts do not have Blacks in their classes.[45]

Terwillinger found differences in the marking policies and practices of schools of different sizes. Marks

---

[44]Enrica Fish, "The Relationship of Teachers' Assigned Marks among Elementary Grades" (Unpublished Ph.D. dissertation, University of Minnesota, 1969).

[45]Jeremy D. Finn, "Expectations and the Educational Environment," Review of Educational Research, XLII (Summer, 1973), 387-410.

were lowered as a disciplinary measure in 40 percent of the large schools but in less than 10 percent of the small schools was this practice used. The large schools more often made provision for a separate citizenship or deportment mark. As discussed in the preceding chapter, Terwillinger also found that many of the differences associated with marking practices were related to the subject area taught.[46] Similar findings were reported by Theimer.[47]

Although significant differences have been identified on the basis of certain situational factors in regard to marking procedures, attempts to determine specific relationship are lacking. It is not known whether these differences are consistent across schools or school districts or whether the differences are associated with particular characteristics of the situation and/or of the teacher.

Effect of Teacher Attitudes

The educator's beliefs regarding the worthwhile goals of education, according to Ebel, are important determiners of his viewpoints regarding the role of measurement and methods for its implementation.[48]

---

[46] Terwillinger, op. cit.

[47] Theimer, op. cit.

[48] Ebel, op. cit., pp. 29-41.

Much of the recent controversy regarding the failure of education to live up to its responsibilities can be traced to differences in opinion regarding the appropriate goals of education. Is the major goal of education to teach cognitive skills? Are the goals of education multidimensional or unidimensional?

Millman made the following statement, "The key task of our schools is to maximize the amount of a subject each student has 'mastered.'"[49] On the other hand Butterworth insisted that the goals of education should be to insure that every student attains a sufficient degree of self identity to make accurate, wise choices. He further stated that this is not being accomplished because "Marks have reached the point where indeed they are means to an end and, in fact, do represent, and have substituted for, the broader outcomes of education."[50]

Concurrence of opinion is noted in the statement by Smith and Dobbins that "perhaps the development of (an inproved marking) system awaits wider agreement on the goals of instruction and the prupose of marking."[51]

The most frequently stated functions of marks are as follows: (1) to motivate students, (2) to provide useful

---

[49]Jason Millman, "Reporting Student Progress," Phi Delta Kappan, LII (December, 1970), 226-230.

[50]Thomas W. Butterworth, "A Guide to the Evaluation of Pupil Progress with Particular Emphasis on Pupil Marks," ERIC No. 036 814 (August, 1966).

[51]Smith and Dobbin, op. cit., p. 789.

information to the pupil and his parents regarding academic performance, (3) to provide useful information for academic and vocational counseling, and (4) to provide information relevant to administrative functions. Less frequently mentioned uses of marks are to: (1) appraise teaching methods and procedures, (2) instruct students, (3) maintain standards of performance, and (4) provide a means of reinforcing students either positively or negatively.[52,53,54,55]

Differences, however, exist in regard to the question of whether marks are assigned in order to provide information to teachers, to students, or to parents. Bolmeier reports that marks should benefit the student not the teacher,[56] while Krause reported that the traditional grading systems best served the needs of the parents but were not

---

[52]William Clark Trow, "On Marks, Norms, and Proficiency Scores," Phi Delta Kappan, XLVIII (December, 1966), 171-173.

[53]William A. Liggett, "Are There Better Ways of Evaluating, Recording, and Reporting Pupil Progress?," National Association of Secondary School Principals, XXXIV (March, 1950), 79-89.

[54]R. C. Billett, Provisions for Individual Differences in Marking and Promotion (Washington, D.C.: U.S. Office of Education, 1932).

[55]A. D. Crooks, "Marks and Marking Systems: A Digest," J. of Ed. Res., XXVII (1933), 259-272.

[56]E. C. Babmeier, "Principles Pertaining to Marking and Reporting Pupil Progress," School Review, LIX January, 1951), 15-24.

useful to either the student or the school.[57] In the years between 1940 and 1970, according to Dionne, the principal functions of marks have changed along with viewpoints regarding the basic goals of education. In 1940 the goal of education was to provide equality of opportunity, therefore, evaluations were most useful in placement and counseling. In 1970 the goal of education was to maximize each potential and the functions of evaluation were expanded to include the facilitation of instruction, resource allocation, and curricular management.[58]

The differences in the appropriate goals of education and the function of marks can hardly be differentiated from the philosophical arguments. Therefore, concrete replies are not likely to be forthcoming. During recent years the roles assumed by education have rapidly expanded. In the early days of the public elementary and high schools the single role of the school was to teach the student those subject areas which were not likely to be learned outside the school. Gradually, however, the school has assumed many of the responsibilities for the social and citizenry learning of its students. At the same time, however, there

---

[57]Dorothy Krause, "A Study of Present Practices and Attitudes Concerning the Letter Grading System" (Unpublished Ph.D. dissertation, University of South Dakota, 1969).

[58]Joseph Dionne, Future of Testing: A Look at the Trends, ERIC 043 669 (March, 1970).

is not universal acceptance regarding this changing role, as many people believe that the school should concentrate on the successful accomplishment of its former role rather than to assume new responsibilities. Because marks are indicators of the degree to which the student has agreed to the objectives of the school, marks must reflect whatever it is that the school is trying to do. Until the role of the school is clearly defined, the role of evaluation will remain ambiguous.

Effect of Other Aspects of the Marking Procedure

A lack of consistency between teachers in regard to other aspects of the marking procedures has been shown. Gould, Ohlson, and Bangs and Greene all reported that from 50 to 75 percent of a mark is determined by daily work, from 25 to 50 percent by quiz marks, and from 25 to 33 percent by test marks.[59,60,61]

It is important then to consider the number of tests or other components that are included in the mark. According to Taylor a large number of daily recitation grades must be included if this is to be an important determiner

---

[59]Herron, op. cit.

[60]C. W. Bangs and H. A. Greene, Teachers Marks and Marking Systems (University of Iowa, 1930).

[61]George Gould, "Practices in Marking and Examinations," School Review, XL (1932), 142-146.

of the assigned mark because of the low reliability associated with such measures. In order to insure a reliable mark, the mark should not be based upon only two or three test scores, unless the tests have been proven to have high reliabilities.[62]

Numerous studies have been conducted to examine the benefits to be gained from using test items, having various formats. However, the results of these studies have been inconclusive, and often contradictory.

In a study of first grade teachers Niedermeyer and Johnson found that teachers, operating according to the mastery learning concept, preferred the use of oral measurements. This method, they claimed, had several advantages: (1) it was easier to administer, (2) it required less time, (3) it eliminated copying, (4) it provided more useful information regarding the individual child's skills, and (5) the individual attention from the teacher was beneficial.[63]

On the other hand, those studies conducted to determine whether teachers can with some uniformity of judgment mark essay type examination items have come to the

---

[62]J. Carey Taylor, The Reliability of Quarterly Marks in the Seventh Grade of Junior High School (Baltimore: Johns Hopkins Press, 1931).

[63]Fred C. Niedermeyer and Howard J. Sullivan, "Differential Effects of Individual and Group Testing Strategies in an Objective-U based Instructional Program," Journal of Educational Measurement, IX (Fall, 1972), 199-204.

nearly universal conclusion that, in fact, this is not true. The range of scores obtained in this type of experiment generally includes the maximum number of possible points as well as the minimum number possible. Lawson found this to be true even when the social background, educational preparation, cultural and industrial environment, level and type of teaching demanded, and experimental background were similar for all teachers involved in the experiment.[64]

The format of items used as contributors to marks may have a significant effect on the marks assigned as well as on the decisions made regarding such marks.

Numerous suggestions have been made regarding the way that marks are to be derived and/or distributed. Yet none of these methods has been demonstrated to have practical significance. Marshall, in his article, showed eleven distributions which were suggested between 1904 and 1924.[65] Others suggest the use of deviation scores and/or standard

---

[64]Douglas E. Lawson, "Scoring of Subjective Tests with Several Vairables Controlled," Elementary School Journal, XXXVIII (February, 1938), 450-457.

[65]Marshall, op. cit.

scores.[66,67,68,69] Yet the practical application of these procedures are not offered.

While differences in the other aspects of the marking procedure has been demonstrated, an attempt to determine whether there is a relationship between these aspects and the basis on which marks are assigned and/or the composition of these marks is lacking.

School Marking Policies

Most school districts regardless of size have reported that they use a five category marking system. However, the number of categories commonly ranges from two to fifteen.

Starch in the early part of the century supported the use of only a few categories because ". . . on the ordinary percent scale a difference of 7 percent or possibly 5 percent is the smallest that can be determined with reasonable reliability."[70]

---

[66]Geoffrey P. Mason, Studies and Reports: An Empirical Analysis of a System of Achievement Grading, ERIC 014 131 (132), November, 1965.

[67]Ralph P. Newsome, "Assignment of School Marks," Texas Outlook, XXV (August, 1941), 39-40.

[68]E. S. Thorndike and E. O. Bregman, "On the Form of Distribution of Intellect in the Ninth Grade," Journal of Educational Research, X (November, 1924), 271-278.

[69]William A. Wetzel, "The Use of the Normal Curve of Distribution in Estimating Students' Marks," School Review, XXIX (May, 1921), 373-378.

[70]Daniel Starch, Educational Psychology (New York: MacMillan, 1927).

Bramlette offered five possible benefits to be derived from the use of a two category marking system: (1) increases emphasis on learning, (2) decreases emphasis on marking, (3) encourages the poorer student, (4) forces students to evaluate themselves, and (5) encourages better attitudes in parents who want a superior child but have instead an average child.[71]

Support for a marking system having several categories came from Ebel who stated that "Regardless of the inaccuracies of the basis for grading, the finer the scale used for reporting grades, that is the more different grade levels it provides, the more accurate the grade reports will be."[72]

In a study conducted to assess the effect of adding + and - grades to the existing five-point scale, thereby, creating a twelve point scale, Philbrick and O'Donnell found the greatest effect for those students at the lower end of the grade distribution. In a significant number of cases, the former C student dropped to either a C- or D+ and the D student dropped to D-. Therefore, it was concluded that the more precise scale served to lower grade point averages

---

[71]Metle Bramlette, "Is the S and U Grading System Satisfactory or Unsatisfactory?," Texas Outlook, XXVI (April, 1942), 29-30.

[72]Robert L. Ebel, "The Relation of Scale Fineness to Grade Accuracy," J. Ed. Measure, VI (Winter, 1969), 217-221.

and, therefore, would not increase the students' self concept and motivation as had been intended.[73]

The introduction of the dual marking system, in which two separate sets of marks are assigned, one for academic performance and the other for affective considerations, was one attempt to improve marking by restricting the elements included in the academic mark.[74]

In an investigation of the effects of single and dual marking systems on the academic performance of gifted sixth grade pupils, Housdorff and Farr reported no associated differences, thereby, concluding that a dual system offers no advantage over a single mark.[75]

Similar conclusions were reached by Halliwell and Robitaille, when they reported a significant correlation of .68 between the two or more marks assigned under a dual system. If a dual system is to increase the reliability and validity of academic marks, teachers must be provided with

---

[73]Joseph L. Philbrick and Patrick I. O'Donnell, "Precision in Grading Practices--Panacea or Problem?," Journal of Educational Research, LXII (December, 1968), 173-176.

[74]John Lund, "More Truth about Marks," Journal of Education (1929), 609-619.

[75]Henry Hounsdorff and S. David Farr, "The Effect of Marking Practices on the Marks of Gifted Sixth Grade Children," Journal of Educational Research, LIX (December, 1965), 169-172.

adequate methods for the assessment of the affective factors.[76]

Other attempts to improve the reporting of student performance have been the substitutions of conferences for at least a part of the written reports. The Kentucky School Journal reported that 40 percent of the surveyed teachers used scheduled conferences. Three times as many elementary as secondary teachers used conferences. Conferences are used more frequently in the Middle and Western states than in the Southeast or Northeast. "Conferences open up new avenues of communication and are especially helpful in coping with subjective areas such as attitudes and interactions with the group."[77]

Still other attempts have used checklists to replace the former methods.

One point then can be conclusively stated in regard to school policies. Differences are found in this area just as they are found in all other sectors of the marking of students. The single "best" way has not yet been found.

---

[76]Joseph W. Halliwell and Joseph P. Robitaille, "The Relationships between Theory and Practice in a Dual Reporting Program," Journal of Educational Research, LVII (November, 1963), 137-141.

[77]Editorial, "Report Cards or Parent-Teacher Conferences: Which is Better?," Kentucky School Journal, IXL (November, 1970), 23.

Summary

As the number of students enrolled in public elementary and secondary schools increased during the latter decades of the nineteenth century and the early decades of the twentieth, there was an accompanying need for a means of written reporting of student progress and performance in regard to how well the goals of the school were being accomplished. It was in response to this need that marking practices and policies were developed.

However, despite numerous changes in an attempt to develop better methods, many criticisms of marking have been voiced.

Two primary philosophical arguments have been presented against the practice of assigning marks. One of thse arguments concerns the question of the intent of marks and the other deals with the use of marks as rewards. Marks have sometimes been construed as means of maintaining a stratified social system. It has also been claimed that marks are harmful as they destroy the natural love of learning by becoming the sought-after reward. These arguments, just as many philosophical arguments, proceed from a stated assumption, in this case, that marks are not accurate, valid indicators of student performance. This assumption is not proven by the argument but is merely to be accepted as a fact. If, in fact, marks are not accurate, valid indicators

of student performance, then the intent of marks may be questionable and marks probably should not function as rewards. However, as long as marks cannot be conclusively proven to be lacking in accuracy and in validity, the claim is not supported.

The methodological and technical criticisms have, for the most part, emphasized the need for well-defined, clearly interpretable marking procedures. Eight alternatives to the commonly used practice of assigning symbols somewhat arbitrarily to represent performance were presented as follows: (1) written evaluation, (2) self evaluation, (3) marks not known by students, (4) contract marking, (5) mastery marking, (6) pass/fail marking, (7) credit/no credit marking, and (8) blanket marking. None of these methods are without a number of advantages and disadvantages, therefore, no single "best" method has been recommended.

Support has been presented for each of the three most frequently used bases for marks. Marking on the basis of comparison to the performance of others or competitive marking has received both approval, as being a more accurate representation of adult life, and disapproval, as being unfair to the less able students. Marking based upon the students' mastery of the subject matter has been criticized as being appropriate to only a limited number

of situations and then only where specific, narrowly defined goals or objectives have been stated. On the other hand, if the primary goal of education is to assist students in the acquisition of a command of a body of substantive knowledge, then marks should indicate how much knowledge the student has attained. Attempts to mark according to the growth of the individual student have not been successful due to problems regarding the measurement of change. However, this latter position has been claimed by many to be the most defensible basis for marking.

Not only must teachers decide upon the basis which they will utilize when assigning marks but they must also decide upon the student characteristics that will be represented by the mark. There is much evidence to support the statement that even when teachers believe that the marks they assign represent purely cognitive functioning there is a high correlation between these marks and ratings made of affective student characteristics. The primary associated difficulty is that subjective ratings of affective characteristics possess low reliability either between points in time or between different raters.

Although findings do not support the statement that differences in marking procedures are associated with the personal characteristics of the teachers such as sex, age, experience, college degree, grade point average, or the

number of measurement courses taken, there is some inconclusive evidence that situational factors such as the urbanization level of the community, size of the school district, percentage of minority students, etc., may be associated with the differences. Significant relationships have been found between the grade level of the students involved and the subject area and the marking procedure used.

Evidence has been presented that teachers differ regarding their beliefs concerning the appropriate goals of education and the functions which marks serve. However, the effect which these factors have in the determination of marking procedures has not been investigated directly. Nor have attempts been made to determine the relationship between marking procedures and other aspects of the marking procedure such as format of test items, frequency of testing, or the use of particular statistical techniques.

Teachers must operate within the framework provided by the school administration when they assign marks. Therefore, because school districts are known to differ in regard to their policies for marking, it can be assumed that teachers must vary accordingly.

Two major conclusions have been derived from this review of literature. The first is that differences exist in regard to how students are marked. These differences exist at the teacher, the school and the school district

level. Evidence does not clearly support the existence of relationships between these differences and the personal characteristics of the teachers, the situational factors, the attitudes of the teachers, or other aspects of the marking procedure used. The second conclusion is that no single "best" method for the reporting of student performance has been developed.

Overview

This chapter has presented a review of the literature relevant to the topic of study. Chapter III includes a summary of the procedures that were used for data collection. Chapter IV presents the data analyses and the derived findings of the study. Chapter V presents the conclusion and the recommendations which were derived from the study.

CHAPTER III

DESIGN OF THE STUDY

The purpose of this chapter is to describe the target population, to define the sample and sampling procedures, to describe the instruments used in data collection, and to specify the statistical methodology which was utilized.

Population and Sample

The target population included those teachers who were assigned on a regular basis to a classroom in a public school within the State of Michigan for the school year 1972-73. Teachers of grade levels kindergarten through grade twelve were included. Excluded were all teachers of special education classes regardless of type. Also excluded were all substitute teachers regardless of the length of their assignment, unless a school district officially listed their names on a roster of teachers assigned to classrooms on a regular basis.

A wide variation was found within the population in regard to the total number of teachers employed by each district, ranging from a minumum of 4 teachers per district to the maximum of Detroit with approximately 11,000 teachers.

Therefore, in order to maximize the probability of the selection of school districts at or near the extremes of the distribution as well as districts within the midrange of enrollment, stratified sampling procedures were used. The variance due to the size of the school districts on the variables of interest was expected to be reduced within each strata, thereby, improving the precision of the results.

Each district within the state was allocated on the basis of the total number of teachers employed to one of five strata, four of which contained approximately equal numbers of school districts, plus one stratum that contained those nine school districts employing the largest numbers of teachers.

A three-stage sampling procedure was employed. Stage one included the selection of school districts for inclusion in the study. Stage two included the selection of schools and stage three the selection of teachers to receive the mailed questionnaires. The Michigan Education Directory for 1972-73 served as the sampling frame for the first and second stages.

Proportional allocation, based upon the fraction of the total population assigned to each stratum on the basis of size, determined the number of school districts to be sampled from each stratum. In this way, the proportion of the sample included from each stratum was equal to the proportion of the population within that stratum. An exception

was made, however, in regard to the stratum containing the 9 large districts. All of the school districts allocated to that stratum were sampled. Because the number of teachers employed by the Detroit Public School District was approximately six times that of the next largest district, Detroit was subdivided into five subdistricts on the basis of geographic proximity, as determined by school addresses. The total number of teachers within each subdistrict was approximately equal to the number for the other districts within the same stratum. Each subdistrict was sampled. In this way, a total sample was created that was an accurate representation of the State of Michigan in regard to the distribution of teachers according to the size of the districts. Table 3.1 shows the population and sample distribution according to stratum.

A simple random sample of school districts was selected from each stratum with the number of districts per stratum determined according to proportional allocation as described above.

The second stage involved the selection of six schools from each sampled district. Two of these schools were senior high schools, two were junior high or middle schools, and two were elementary schools, thereby, insuring that all levels were represented in the sample. The information within the sampling frame was assumed to be correct

TABLE 3.1.--Target Population Distribution and Sample According to Strata I through V.

Stratum	Stratum Boundaries	Number of Districts in Population	Proportion of Population	Number of Districts in Sample
I	1-50	129	.242	31
II	51-100	168	.316	41
III	101-200	124	.233	30
IV	201-1000	102	.101	25
V	Over 1000	9	.018	13*
TOTAL		532	1.001	140
		2**		
		534		

*Includes 5 subdistricts for Detroit.
**Districts for which required information was not available.

in regard to the classification of schools by level. Whenever a sampled district has only one or two schools of a specified level, those schools were automatically sampled. No elementary school, however, was sampled that contained fewer than three grade levels. In the case of a district that included no junior high or middle school, only four schools were sampled, two senior high schools and two elementary schools. Whenever more than two schools of a specified level were listed, simple random sampling was used to choose those schools that were included in the study. Table 3.2 shows the allocation by school level and the number of schools per level according to strata.

TABLE 3.2.--Sample Distribution According to Level of Schools by Strata I through V.

Stratum	Total Number of Schools	Number of High Schools	Number of Junior High or Middle Schools	Number of Elementary Schools
I	75	30	9	36
II	131	40	25	65
III	117	30	27	60
IV	124	30	38	42
V	79	26	26	27
TOTAL	526*	156	125	230

*The total number of schools (526) is not equal to the total number of high schools, junior high or middle schools, and elementary schools due to schools which included more than one level and, therefore, were included only in the total number of schools.

The third stage involved the selection of two teachers from each of the selected schools. Two teachers per school was considered to be adequate as teachers were to be the unit of analysis rather than schools or school districts. Teachers were chosen as the unit of analysis because previous studies demonstrated both the lack of district wide marking procedures and the resistance of teachers to abide by such procedures even when they were established. The sampling frames for this stage were rosters of teachers' names maintained by the selected school administrators. A

simple random sampling procedure was used to select teachers for inclusion in the study.

Data Collection and Instrumentation

Table 3.3 summarizes the source of information for each variable of interest.

The primary source was the teachers' responses to the mailed questionnaire. This instrument was developed especially for this study. Approximately 15 minutes were required for completion of the twenty-three items included in the questionnaire. One item allowed for free response, while the others offered several options each. Each questionnaire included an identification number. A copy of the questionnaire is included in Appendix A.

Table 3.4 presents the reliability coefficients of the questionnaire computed using Cronbach's Alpha for groups of items for which it was appropriate. Because a reliability coefficient would not provide useful information for the free response items used, those items were not included. Also, because a reliability coefficient computed on strictly demographic items is meaningless, these items were also excluded from the calculation. The reliability coefficients reported were based upon the particular sample of responders and, for this reason, do not imply cross validation. For these reasons, the reliability coefficients were calculated

TABLE 3.3.--Sources of Information Regarding the Variables of Interest

Variable	Source
Attitudinal factors	
appropriate goals of education	Questionnaire (8 items)
functions of marks	Questionnaire (9 items)
suggested changes	Questionnaire (free)
Marking policies	
frequency	Report form
symbols used	Report form
dual system	Report form
overall formal	Report form
Marking procedures	
basis	Questionnaire (1 item)
type objective measures	Questionnaire (2 items)
frequency objective measures	Questionnaire (3 items)
frequency of use of statistics	Questionnaire (3 items)
student characteristics considered	Questionnaire (14 items)
relative importance of student characteristics	Questionnaire (14 items)
Personal factors	
sex	Questionnaire (1 item)
age	Questionnaire (1 item)
experience	Questionnaire (1 item)
degree held	Questionnaire (1 item)
academic success	Questionnaire (1 item)
courses taken	Questionnaire (1 item)
Situational factors	
size district	Education Directory
size school	Education Directory
composition	Education Directory
urbanization	Assessment Program Report
minority level	Assessment Program Report
grade level taught	Questionnaire (1 item)
subject area taught	Questionnaire (1 item)

TABLE 3.4.--Cronbach's Alpha Reliability for Appropriate Groups of Items by Strata I through V.

Variable	Item Numbers	I	II	III	IV	V	Total
Frequency of objective measurements	9-10-11	.307	.342	.455	.431	.619	.431
Format of measurements	12-13	.362	.805	.458	.439	.333	.479
Frequency of use of statistics	16-17-18	.467	.370	.242	.470	.512	.412
Goals of education	19	.896	.871	.483	.841	.904	.799
Functions of marks	20	.840	.827	.786	.991	.863	.861
TOTAL		.759	.838	.544	.899	.813	.771

only for those items which comprise five of the measures: (1) the frequency of use of objective measurements, (2) the format used for measurement items, (3) the frequency of use of statistics, (4) the appropriate goals of education, and (5) the function of marks.

Statements included within the item pertaining to the teacher's viewpoint concerning the appropriate goals of education were divided into two categories of four statements each. Four of the statements were most likely to be rated highly by someone espousing the traditional belief and the other four were most likely to be rated highly by someone favoring the progressive or humanistic belief. Scores were totalled for the eight items included in the two categories, with those in the traditional category being assigned positive values and those in the progressive category being assigned negative values. This permitted each respondent to be placed at a point on a continuum, ranging from High Traditional to High Progressive. Those respondents who selected equivalent response for both of the categories were, in this way, assigned to an Ambivalent position.

The statements concerning the functions that marks successfully serve were separated into three categories. Three statements dealt with functions relevant to the student himself; three dealt with the provision of useful information to the parents; three dealt with functions relevant

to the teacher. As respondents were asked to indicate their position on a 0 to 5 scale, each of the three category scores ranged from 0 to 15.

In order to compute a value representing the frequency of the use of objective measures of student performance, the three items dealing with the frequency with which homework is assigned, the frequency of quizzes, and the frequency of longer tests were combined as shown in Table 3.5. The summation of the values for the three items yielded possible scores ranging from 0 to 14.

TABLE 3.5.--Scoring Procedure for Items Dealing with Frequency of Measurement

Value	Homework Frequency	Quiz Frequency	Test Frequency
0	never	never	never
1			1 or 2 times marking period
2		end of unit	end of unit
3	less than 1 or 2 weekly	less than 1 or 2 weekly	every 2 or 3 weeks
4	1 or 2 weekly	1 or 2 weekly	1 or 2 weekly
5	daily	daily	

Similarly, to obtain a value representing the frequency of use of particular statistical techniques, standard scores, percentile ranks, and a normal distribution,

responses were first combined into three overall categories, always, sometimes, and never, each category being assigned a value of 3, 2, and 1 respectively. In this way, a possible score ranging in value from 0 to 9 was obtained as an indication of the total frequency of use of these statistical techniques.

The Michigan Educational Assessment Program Report provided information regarding the five categories into which school districts were separated on the basis of the urbanization level of the community. The five categories are metropolitan, urban fringe, city, town, and rural, with 15, 24, 105, 119, and 271 school districts in the population per respective category. Similar divisions were used in this study. Table 3.6 presents the distribution of responses for this variable.

TABLE 3.6.--Distributions of Teachers' Responses by Urbanization Level According to Strata.

Stratum	Metropolitan	Urban Fringe	City	Town	Rural
I	0	4	9	11	89
II	0	7	2	35	135
III	2	36	10	97	22
IV	15	117	34	15	0
V	54	32	0	0	0
Total Number of Teachers	71	196	46	158	246

In regard to the measure concerning the composition of marks, the relative weightings assigned specific student characteristics, ranks were computed for each respondent. However, the original percentage values reported were also retained.

For the other measures obtained from the questionnaires, the categories used were those included as response options. In order that quantitative types of analyses be utilized, dummy variables were created for each level of the categorical variables. Dummy variables were assigned a value of 2 if the response belonged to the particular level; otherwise, they were equal to 1.

Report forms were analyzed according to school level on the following bases: (1) is a dual system used, (2) are teacher comments included, (3) is a check list of objectives used, (4) is a narrative report used, (5) how many times per year is it issued, (6) which of the following is used: A-B-C-D-E, S-U, or S-A-I-U, (7) is it returned to teacher signed by parent, and (8) is there a place for effort and/or conduct.

Responses to the free-response items dealing with suggested changes in marking procedures were divided into the following categories: (1) eliminate entirely, (2) narrative, (3) pass-fail, (4) conferences and reports, (5) system or school wide procedures, (6) more emphasis on academic,

(7) more emphasis on affective characteristics, (8) other, and (9) no change.

Design

Because many school boards control external research and experimentation, the first step was to secure their approval. School superintendents, as representatives of the school boards, were sent letters stating briefly the purpose of the study, its importance, the amount of time and involvement requested of the personnel in the district, and a request made for their cooperation. A request was made for a list of the names and addresses of the teachers assigned within each of the sampled schools and for an unused copy of the student progress report form used in each of these schools. Appendix C contains a copy of the original letter and a copy of the follow-up letter mailed three weeks later. Questionnaires were then mailed directly to these teachers. Appendix E includes a copy of the enclosure letter sent to these teachers as well as follow-up letters.

The response rate from superintendents was 55.1 percent for the total sample with two superintendents declining to participate; therefore, another means was needed to make contact with teachers. To accomplish this, letters were sent to the principals of the selected schools in nonresponding districts, requesting that they distribute

the accompanying questionnaires. Directions were provided specifying the method by which teachers were to be chosen to receive the questionnaires. A roster of teachers' names was to be numbered consecutively, excepting special education teachers and all other special services personnel. The accompanying questionnaires were addressed to teacher number ___. These numbers had been randomly selected from a random number table based upon the available information regarding the number of teachers per building. Appendix D contains a copy of the letter sent to the principals and the follow-up letter sent two weeks later. Table 3.7 shows the distribution of responses for both those receiving questionnaires directly and those receiving them through principals.

After the deadline for the returns from the follow-up letters, the response rate for teachers was 60.8 percent. From the remaining non-responders a 25 percent sample was chosen to receive an extra effort in an attempt to elicit their responses in order to determine whether the non-responders differed on the measures of interest from the responders. A minimum of 25 non-responders were selected from each stratum to receive the extra effort, which consisted of registered letters and telephone calls.

Table 3.8 presents the response rate for the regular mailing and for the extra effort. It is apparent that the extra effort was more successful with teachers from the

TABLE 3.7.-- Dual Sampling and Response Rates by Strata and by Total.

Strata	Number Sent to Supt.	Number received from Supt.	Percent response	Number of teachers sent from Supt.	Number of teachers response from Supt.	Percent response	Number of teachers sent to principals	Number of teachers response from principals	Percent response	Total Number of teachers sent	Total Number of teacher response	Total Percent response
I	31	14	45.2	48	38	79.2	100	66	66.0	148	104	70.4
II	41*	23 (+ 1 refusal) (6 teachers)	56.1	144	106	73.6	112	73	60.9	256	179	70.0
III	30	20	66.7	142	99	69.7	92	68	74.0	235	167	71.4
IV	25*	14 (+1 refusal) (10 teachers)	56.0	84	61	72.6	142	120	84.5	226	181	80.0
V	13(9)	4	44.0	34	20	58.8	124	66	53.2	158	86	54.5
Totals	140	75	55.1	452	324	71.6	570	393	69.0	1022	717**	70.15

*Indicates the number of superintendents who replied and declined to participate plus the number of teachers who would have been included in the study.

**This total includes response from each sampled school district with the exception of those indicated in column two.

TABLE 3.8.--Responses from Regular Mailing and from Extra Effort by Strata and by Total

Strata	Total Number of Teachers	Number of Early Responses	Percent Early Responses	Total Number Receiving Extra Effort	Number of Responses from Extra Effort	Percent Response from Extra Effort
I	148	93	62.9	25	11	44.0
II	256	165	64.5	25	14	56.0
III	234	139	59.4	28	28	100.0
IV	226	161	71.3	27	20	74.0
V	158	63	39.9	26	23	88.5
Totals	1022	621	60.8	131	96	73.3

larger school districts. This may have been caused by the fact that the smaller school districts were closer to the end of their school year than were the larger districts at the time that the extra effort was being made. However, with the exception of the two school districts for which the superintendent declined, response was obtained from all sampled districts.

In order to determine whether there were significant differences between the responses to the items pertaining to the relative importance of student characteristics from the regular mailing and those elicited by the extra effort, multivariate and univariate analyses of variance were calculated. Table 3.9 presents the results of these tests, including the computed F ratios, the degrees of freedom, and the level of significance. As is shown, no significant multivariate or univariate F ratio was found at the .01 level. Therefore, the assumption was made that the respondents were an unbiased sample of the population of teachers in Michigan.

The two groups, respondents from the regular mailing and extra effort, were combined in Table 3.7 which indicates the total response rate as well as the response rate from each of the two sampling procedures.

A total response rate for teachers of 70.15 percent is reported. The response rate for Strata I is reported as 70.4 percent. The response rate for Strata II is 70.0 percent and for Strata III is 71.4 percent. The rate for

TABLE 3.9.--Multivariate and Univariate F-Ratios for Composition of Marks and Whether Response Received from Regular Mailing or from Extra Effort for Stratum I through V.

		F-Ratio	Significance	DF
Stratum I	-Overall	1.3115	.221	13/90
	Attendance	.3894	.534	1/103
	Appearance	.1565	.693	1/103
	Conduct	.0005	.983	1/103
	Effort	1.7952	.183	1/103
	Participation	.1337	.715	1/103
	Attitudes	.2191	.641	1/103
	Neatness	.4754	.492	1/103
	Homework	3.1034	.081	1/103
	Quizzes	.8958	.346	1/103
	Tests	.6450	.424	1/103
	Optional	.0244	.876	1/103
	Group Reports	4.1086	.045	1/103
	Individual Reports	1.1432	.288	1/103
Stratum II	-Overall	1.2860	.226	13/164
	Attendance	1.7908	.183	1/178
	Appearance	5.8821	.016	1/178
	Conduct	5.6242	.018	1/178
	Effort	.2944	.588	1/178
	Participation	.1105	.740	1/178
	Attitudes	.6081	.437	1/178
	Neatness	3.4906	.063	1/178
	Homework	.7146	.399	1/178
	Quizzes	2.4858	.117	1/178
	Tests	1.9959	.159	1/178
	Optional	5.3641	.022	1/178
	Group Reports	2.3343	.128	1/178
	Individual Reports	2.5419	.113	1/178
Stratum III	-Overall	.3981	.969	13/153
	Attendance	.0036	.952	1/166
	Appearance	.0033	.954	1/166
	Conduct	.0101	.920	1/166
	Effort	2.0113	.158	1/166
	Participation	1.7808	.184	1/166
	Attitudes	.2024	.653	1/166
	Neatness	.1321	.717	1/166
	Homework	1.3935	.240	1/166
	Quizzes	.3772	.540	1/166

TABLE 3.9.--Continued.

		F-Ratio	Significance	DF
	Tests	.3259	.569	1/166
	Optional	.0431	.836	1/166
	Group Reports	.0030	.956	1/166
	Individual Reports	.0812	.776	1/166
Stratum IV	-Overall	.8031	.656	13/167
	Attendance	.1476	.701	1/180
	Appearance	1.3820	.241	1/180
	Conduct	2.4136	.122	1/180
	Effort	1.2595	.263	1/180
	Participation	.0151	.902	1/180
	Attitudes	1.4282	.234	1/180
	Neatness	.0077	.930	1/180
	Homework	.5136	.475	1/180
	Quizzes	1.7210	.191	1/180
	Tests	.8812	.349	1/180
	Optional	.5704	.451	1/180
	Group Reports	1.0769	.301	1/180
	Individual Reports	.7437	.390	1/180
Stratum V	-Overall	1.1083	.366	13/72
	Attendance	.2310	.632	1/85
	Appearance	.7538	.388	1/85
	Conduct	.4426	.508	1/85
	Effort	3.4092	.068	1/85
	Participation	.1849	.668	1/85
	Attitudes	.7222	.398	1/85
	Neatness	.6482	.423	1/85
	Homework	.0021	.963	1/85
	Quizzes	2.5567	.114	1/85
	Tests	.0004	.983	1/85
	Optional	5.9661	.017	1/85
	Group Reports	1.9576	.165	1/85
	Individual Reports	1.2782	.261	1/85

Strata IV is reported as 80.0, while that for Strata V is 54.5 percent.

When the questionnaires were sent directly to the teachers, the response rate was higher for Strata I, II, and V, while for Strata III and IV the higher response rate was obtained from the teachers who received the questionnaires through the principals.

In order to determine whether significant differences were associated with to whom the questionnaire was sent, multivariate and univariate analyses of variance were calculated. Table 3.10 shows the F-ratios for these analyses, the degrees of freedom, and the significance level.

As is shown, only for Stratum II was a significant multivariate difference found at the .01 level. Because this resulted only from a single univariate significance, no further consideration was made of this difference.

Analysis Methodology

The range of all measures was restricted by eliminating those item options or values reported by fewer than two percent of the respondents, i.e., were not reported by at least two respondents in Strata I and V, by three in Stratum III, or by four in Strata II and IV. The respondents who reported these options were assigned to either the next higher or the next lower category. All of these changes assigned the respondents to the category closer to the mean for the particular measure.

TABLE 3.10.--Multivariate and Univariate F-Ratios for Composition of Marks and to Whom the Questionnaires were Sent for Stratum I through V.

		F-Ratio	Significance	DF
Stratum I	-Overall	.4916	.924	13/90
	Attendance	.7422	.391	1/103
	Appearance	.2132	.645	1/103
	Conduct	.0395	.843	1/103
	Effort	.0636	.802	1/103
	Participation	1.2899	.259	1/103
	Attitudes	.9444	.333	1/103
	Neatness	.7653	.384	1/103
	Homework	.0381	.846	1/103
	Quizzes	.0004	.984	1/103
	Tests	.2239	.637	1/103
	Optional	.0094	.933	1/103
	Group Reports	.8074	.371	1/103
	Individual Reports	2.5558	.113	1/103
Stratum II	-Overall	2.4093	.005*	13/165
	Attendance	1.3289	.251	1/178
	Appearance	.3187	.573	1/178
	Conduct	.5382	.464	1/178
	Effort	1.2335	.268	1/178
	Participation	.2942	.588	1/178
	Attitudes	.3303	.566	1/178
	Neatness	1.3059	.255	1/178
	Homework	7.5443	.007*	1/178
	Quizzes	2.9895	.086	1/178
	Tests	1.4250	.234	1/178
	Optional	.0161	.899	1/178
	Group Reports	.4748	.492	1/178
	Individual Reports	.1230	.726	1/178
Stratum III	-Overall	.5941	.856	13/153
	Attendance	.0444	.833	1/166
	Appearance	.5671	.452	1/166
	Conduct	.4494	.504	1/166
	Effort	.2025	.653	1/166
	Participation	.0084	.927	1/166
	Attitudes	.0783	.780	1/166
	Neatness	.3213	.572	1/166

*Indicates significance at $\alpha = .01$.

TABLE 3.10.--Continued.

		F-Ratio	Significance	DF
	Homework	2.3661	.126	1/166
	Quizzes	.0562	.813	1/166
	Tests	.3865	.535	1/166
	Optional	.3726	.543	1/166
	Group Reports	.1316	.717	1/166
	Individual Reports	.9766	.324	1/166
Stratum IV	-Overall	1.2421	.254	13/167
	Attendance	3.9622	.048	1/180
	Appearance	.0008	.978	1/180
	Conduct	1.2953	.257	1/180
	Effort	1.3974	.239	1/180
	Participation	.8588	.355	1/180
	Attitudes	2.8497	.093	1/180
	Neatness	4.2732	.040	1/180
	Homework	1.0128	.316	1/180
	Quizzes	.1011	.751	1/180
	Tests	.2335	.630	1/180
	Optional	.0822	.775	1/180
	Group Reports	1.0469	.308	1/180
	Individual Reports	.0121	.913	1/180
Stratum V	-Overall	.9785	.481	13/72
	Attendance	.1512	.698	1/85
	Appearance	1.0513	.308	1/85
	Conduct	.1259	.724	1/85
	Effort	5.5506	.021	1/85
	Participation	.8838	.350	1/85
	Attitudes	.0067	.935	1/85
	Neatness	.2338	.630	1/85
	Homework	.0926	.762	1/85
	Quizzes	.7784	.380	1/85
	Tests	.0431	.836	1/85
	Optional	.0000	.998	1/85
	Group Reports	.1299	.719	1/85
	Individual Reports	.0053	.942	1/85

In order to answer the questions which were defined as the objectives of the study, statistical analysis techniques were utilized.

Kendall's coefficient of concordance was used to answer the following questions:

Do teachers agree on the relative importance of student characteristics considered when assigning marks?

Chi square analyses were utilized to answer the following questions:

What personal factors are related to the basis on which marks are assigned?

What situational factors are related to the basis on which marks are assigned?

What attitudinal factors are related to the basis on which marks are assigned?

What other aspects of the marking procedures used are related to the basis on which marks are assigned?

Canonical correlational analyses were utilized to answer the following questions:

What personal factors of teachers are related to the weights applied to student characteristics considered when assigning marks?

What situational factors are related to the weights applied to student characteristics considered when assigning marks?

What attitudinal factors are related to the weights applied to student characteristics considered when assigning marks?

What other aspects of the marking procedure used are related to the weights applied to student characteristics considered when assigning marks?

A Chi square one sample test was employed in order to determine whether responses regarding the basis on which marks were assigned formed a uniform distribution.

Do teachers agree in regard to the basis on which marks are assigned?

The following remaining questions were answered by means of summary tables.

What types of objective measure of student performance are made by teachers?

How frequently are objective measurements of student performance made by teachers?

What changes do teachers suggest regarding marking procedures and policies?

How do school districts differ in their marking policies?

The first step in the analysis of the data was to determine whether differences existed between the strata according to the basis on which marks were assigned and the student characteristics that were considered when assigning marks.

A Chi square test for independence was utilized to determine whether there was a relationship between the stratum and the basis on which marks were assigned. The resulting information is presented in Table 3.11. The results were not significant at the .05 level of significance. Therefore, for this purpose of this study it may be said that the basis upon which marks are assigned is independent of the stratum from which the teacher was sampled.

TABLE 3.11.--Contingency Table for Stratum and the Basis on Which Marks Are Assigned.

	I	II	III	IV	V	Total
Growth	9	23	19	23	11	85
Mastery	13	25	15	17	11	81
Performance	5	2	3	5	4	19
Growth Mastery	40	76	90	83	38	327
Growth Performance	11	18	16	19	6	70
Mastery Performance	27	31	19	25	9	111
None	0	4	5	8	7	24
TOTAL	104	179	167	181	86	

$X^2 = 36.115$

df = 24 not significant at $\alpha = .05$.

In order to determine whether the stratum had a significant effect upon the student characteristics which were considered when marks were assigned multivariate and univariate F-ratios were calculated. These results are shown in Table 3.12. As shown, both the overall multivariate F-ratio and all univariate F-ratios indicate that the teachers' choice of student characteristics for consideration when assigning marks is not affected by the stratum, based upon school district size within which they teach.

TABLE 3.12.--Multivariate and Univariate F-Ratios for Composition of Marks and Stratum.

	F	Significance	DF
Overall	1.1560	.208	52/2720.945
Attendance	.8781	.477	4/714
Appearance	1.4776	.207	4/714
Conduct	.8137	.517	4/714
Effort	1.032	.390	4/714
Participation	1.240	.292	4/714
Attitudes	1.428	.223	4/714
Neatness	.7650	.548	4/714
Homework	1.0416	.385	4/714
Quizzes	1.4855	.205	4/714
Tests	.8210	.512	4/714
Optional	1.1200	.346	4/714
Group Reports	1.3359	.255	4/714
Individual Reports	.8398	.5002	4/714

Because the results of the Chi square test for independence for the basis for marks, the multivariate analysis of variance for the overall student characteristics considered, and the univariate analysis of variance for each of the student characteristics considered indicated that the basis for marks and the student characteristics represented by marks were independent of the stratum from which the teachers were sampled, the decision was made to eliminate the consideration of strata from the remaining portions of the study. Therefore, the analyses of the data, intended to derive answers to the questions which are the objectives of the study, utilized the entire sample of 717 respondents.

Overview

This chapter has presented a description of the population and sample, a description of the instrument used for data collection, and a description of the procedures employed for data collection and analysis. Chapter IV presents the data analyses and findings. Chapter V presents the conclusions and recommendations that were derived from the study.

CHAPTER IV

DATA ANALYSES AND FINDINGS OF THE STUDY

In this chapter is presented an explanation of the data analyses that were conducted and the findings that were derived. The first section includes a brief description of the statistical techniques which were utilized. The second section provides the statistical findings of each data analysis and the related interpretation of the findings. The third section provides an overall summary of the results of the study.

Statistical Methodology

Cannonical correlation analysis allowed the researcher to test the independence of two sets of variables in the multidimensional population. The test utilizes the covariance matrix for the two sets of standardized variables, from which are extracted the largest characteristic roots. These then are used to form a system of equations from which the coefficients are derived.

Canonical variates, weighted composites of variables are formed so as to maximize the proportion of variability in one set of variables that is explained by the other set. These are formed as unique, orthogonal sets of coefficients

with the number of variates that are formed dependent upon the smallest number of variables in either of the two sets. From these coefficients are calculated canonical R's, the maximum value that would be obtained by correlating the two weighted sets of values.

The interpretation of a canonical analysis is dependent upon the signs and relative magnitudes of the coefficients as well as upon the canonical R^2, which indicates the amount of variance in one set of weighted variables that is statistically accounted for by the other set of weighted variables. Canonical correlation analyses utilize data which must be measured at least on an interval scale.

On the other hand, nominal or categorical data is required for a Chi square test. The Chi square test of independence provided the researcher with a means of determining whether two variables were independent. The computed Chi square value is dependent upon the number of levels for the two variables as well as upon the sample size.

By computing the Chi square value for each level of the joint distribution, and using a single degree of freedom, an indication of the relative contribution of each variable level to the overall, significant Chi square value is obtained. Thereby, the specification of those levels of the joint distribution for which there is dependence between the two variables, is possible.

Kendall's coefficient of concordance provided a measure of the degree to which there was agreement among the teachers in regard to the relative importance applied to each student characteristic. The interpretation of Kendall's coefficient of concordance indicates whether the value is statistically different from complete disagreement. Data for this test included ranks from one to thirteen, the number of characteristics being ranked.

When the data consisted of the frequency of choice for the bases for marks, a Chi square one sample test enabled the researcher to determine whether a uniform distribution of responses was created. The interpretation of the obtained Chi square value depends upon the number of levels of the variable and the desired level of confidence. A significant value indicates that the distribution of the obtained data departs from a uniform distribution is at

All analyses were performed on either the CDC 6500 or the 3600 computer at Michigan State University utilizing programs developed by Jeremy Finn of the State University of New York at Buffalo and the Computer Institute for Social Science Research at Michigan State University.

Statistical Findings

Findings are presented below according to the type of analysis performed with the remaining organization dependent upon the questions that were to be answered by the study.

Whenever appropriate an α level of .05 was used as representing the desired level of confidence.

For the most part due to space limitations only those findings are presented which were statistically significant. However, in Appendices B and C the complete frequency distribution for all variables of interest as used in the analyses are presented.

Canonical Correlation Analyses

Question 1: What personal factors are significantly related to the weighting applied to student characteristics that are considered when assigning marks?

Findings: The correlation matrix presented in Table 4.1 reports the degree of relationship between the teachers' personal characteristics and the weighting applied to each student characteristic considered when assigning marks, as well as the intercorrelation for each of these measures.

A consideration of the interrelationships among the personal factors showed that, as would be expected, the older teachers reported more years of experience than did the younger teachers. The older teachers also reported that they had taken a greater number of measurement courses and held higher college degrees. Similar findings occurred between experience and both degree held and number of measurement courses. It was also found that male teachers were more often older than females, while female teachers were

less likely to hold advanced college degrees and more likely to report higher grade point averages.

A consideration of the interrelationship among the weightings applied to student characteristics showed either a marked or close relationship with the following exceptions. The importance which a teacher applies to attendance was found to be independent of the importance attached to conduct, neatness of work, homework, test marks, and individual projects or reports, while only slightly related to the importance associated with class participation and optional work. The importance applied to conduct was found to be only slightly related to the importance of effort and moderately related to attitudes, quiz marks, and group reports or projects. The importance of effort was found to be only slightly related to the importance of neatness of work, homework, test marks, and individual reports on projects. The importance of student attitudes was found to be moderately related to neatness of work, test marks, and individual projects or reports. The importance associated with group reports or projects was found to be moderately related to neatness of work, homework, and test marks, while individual reports or projects were found to be moderately related to quiz marks and group reports or projects. The importance of quiz and test marks was moderately related.

A consideration of the cross correlations between the teachers' personal factors and the weighting applied to

each student characteristic considered showed that while many of the correlations were statistically significant, none of them indicate more than a slight relationship.

Table 4.2 reports the first three canonical variates that were derived from the above correlation matrix. The first canonical variate indicated a canonical R of .421 between the weighted set of student characteristics and the weighted set of personal factors. The set of student characteristics demonstrated a high positive loading on neatness of work and a high negative loading on the students' personal appearance. Moderately high positive loadings were found for group reports or projects and attendance. Moderately high negative loadings were found for homework and test marks and for individual reports or projects.

The set of personal factors demonstrated all positive loadings with the highest loading for sex. Moderately high loadings were found for both experience and high school grade point average.

The second significant canonical variate indicated a canonical R of .253 between the two sets of measures. The set of student characteristics demonstrated a high positive loading on effort and a high negative loading on attendance. Moderately high positive loadings were found for neatness of work and optional work. Moderately high negative loadings were found for individual reports or projects, participation, and appearance.

TABLE 4.1.--Correlation matrices for personal factors and the weighting applied to student characteristics.

Sex	1.00																			
Age	.140	1.00																		
Experience	.042	.783	1.00																	
Degree	-.186	.196	.319	1.00																
Courses	-.058	.252	.316	.285	1.00															
GPA	-.239	-.011	-.010	.035	.064															
Attendance	.031	.083	.040	-.023	.004	-.026	1.00													
Appearance	-.048	.013	-.032	-.112	-.053	-.081	.799	1.00												
Conduct	-.107	-.076	-.088	-.152	-.080	-.086	.005	.590	1.00											
Effort	.028	.061	.015	-.057	-.016	-.043	.984	.878	.151	1.00										
Participation	-.093	-.046	-.064	-.158	-.074	-.101	.173	.721	.963	.317	1.00									
Attitudes	-.020	.031	-.016	-.102	-.038	-.072	.871	.987	.481	.934	.624	1.00								
Neatness	-.112	-.075	-.092	-.159	-.159	-.086	.004	.632	.983	.154	.976	.487	1.00							
Homework	-.140	-.095	-.113	-.158	-.088	-.095	.041	.625	.971	.186	.968	.513	.988	1.00						
Quizzes	-.034	.028	-.020	-.098	-.040	-.068	.871	.988	.425	.932	.621	.995	.486	.515	1.00					
Tests	-.177	-.092	-.100	-.147	-.076	-.080	.007	.587	.950	.146	.943	.473	.969	.964	.428	1.00				
Optional	-.103	-.068	-.091	-.161	-.084	-.095	.175	.727	.967	.320	.987	.628	.980	.977	.627	.952	1.00			
Group Reports	-.011	.041	-.004	-.080	-.027	-.057	.930	.961	.356	.973	.512	.988	.363	.394	.988	.356	.518	1.00		
Individual	-.119	-.079	-.095	-.156	-.083	-.091	.001	.600	.982	.151	.975	.484	.998	.987	.483	.968	.980	.362	1.00	
	Sex	Age	Experience	Degree	Courses	GPA	Attendance	Appearance	Conduct	Effort	Participation	Attitudes	Neatness	Homework	Quizzes	Tests	Optional	Group Reports	Individual	

The set of personal factors demonstrated negative factor loadings with the exception of sex, which had also the highest loading. A high negative loading was found for degree. Moderately high loadings were found for both age and experience.

The third significant canonical variate indicated a canonical R of .234 between the two sets of measures. The set of student characteristics demonstrated a high positive loading for group reports or projects and a high negative loading for neatness of work. A moderately high positive loading was found for individual reports or projects. Moderately high negative loadings were found for attitudes and attendance. An extremely low positive loading was found for test marks.

The set of personal factors demonstrated a high negative loading for age and high positive loadings for degree and high school grade point average. A moderately high loading was found for sex.

With the first canonical variate 17.8 percent of the variance in the weightings applied to student characteristics was statistically accounted for by the personal characteristics of the teacher. With the second canonical variate 6.4 percent of the variance was accounted for and with the third variate 5.5 percent was accounted for. It, therefore, seems that a knowledge of the personal characteristics of the teacher particularly in regard to sex would provide a slightly accurate

TABLE 4.2.--Significant Canonical Variates for Personal Factors and Student Characteristics

	First Canonical Variate	Second Canonical Variate	Third Canonical Variate
Student Characteristics			
Attendance	2.511	-7.438	-3.232
Appearance	-7.113	-2.125	-2.082
Conduct	- .425	- .114	1.719
Effort	- .309	7.229	1.524
Participation	1.043	-2.268	-2.867
Attitudes	.588	1.051	-3.563
Neatness	8.712	3.431	-4.608
Homework	-3.129	1.441	1.267
Quizzes	.442	- .651	-1.557
Tests	-2.040	-1.093	.118
Optional	.789	2.905	1.519
Group Reports	2.646	1.542	8.290
Individual	-2.610	-4.522	2.791
Personal Factors			
Sex	.915	.433	.297
Age	.119	- .347	- .629
Experience	.274	- .279	- .148
Degree	.125	- .400	.595
Courses	.041	- .143	.170
GPA	.295	- .141	.586
Canonical R	.421	.253	.234
Canonical R^2	.178	.064	.055
Redundancy	1.366	.491	.421
x^2	256.57	118.15	71.43
df	78	60	44
Significance	.0001	.0001	.0056

prediction of the comparative importance attached to the student characteristics considered when assigning marks.

> Question 3: What situational factors are significantly related to the weighting applied to student characteristics that are considered when assigning marks?

Findings: The correlation matrix presented in Table 4.3 reports the degree of relationship between the situational factors and the weighting applied to each student characteristic that is considered when assigning marks as well as the intercorrelation matrix for the situational factors. The intercorrelation matrix for the student characteristics is presented above in Table 4.1.

A consideration of the intercorrelations among the situational factors revealed a close relationship between grade taught and school level as well as between grade level and subject area taught. As would be expected, it was found that more minority students were enrolled in the metropolitan school districts than in rural districts.

A consideration of the cross correlations between the two measures showed several slight relationships. Teachers in urban school districts were found to attach less importance to the students' personal appearance than did teachers in rural districts. Teachers, in districts having a high percentage of minority students, attached more importance to attendance, conduct, and personal appearance than did other teachers. Teachers of nonacademic subject areas and elementary grade levels were found to attach more

TABLE 4.3.--Correlation Matrices for Situational Factors and Weighting Applied to Student Characteristics

Level	1.00						
Urban	.295	1.00					
Minorities	-.118	-.489	1.00				
Number District	-.127	-.308	.018	1.00			
Number School	-.338	-.159	-.173	-.174	1.00		
Subject	.467	.048	-.050	-.082	-.208	1.00	
Grade	-.818	-.025	.032	-.019	-.014	.837	1.00
Attendance	.016	-.040	.165	-.027	-.045	.092	-.038
Appearance	-.031	-.116	.203	.040	-.084	.079	-.039
Conduct	-.015	-.071	.115	.006	-.076	.088	-.023
Effort	.168	.000	.001	.017	-.081	.279	-.339
Participation	.082	.054	-.037	-.036	-.025	.151	-.102
Attitudes	.073	-.037	.037	.017	-.052	.189	-.141
Neatness	.132	-.037	.037	-.009	-.118	.207	-.212
Homework	-.080	.005	.074	-.069	-.006	-.210	.274
Quizzes	.011	-.056	.053	-.009	-.019	-.061	.096
Tests	-.092	-.051	.018	.003	.058	-.243	.339
Optional	.094	-.023	-.011	-.007	-.079	.123	-.040
Group Reports	.056	-.018	.023	.014	-.064	.066	-.025
Individual	-.058	-.017	.013	.019	-.031	.070	.206
	Level	Urban	Minorities	Number District	Number School	Subject	Grade

importance to effort, class participation, students' attitudes and neatness of work. Teachers of academic subject areas attach more importance to homework and test marks. Teachers of upper grade levels also attach more importance to homework, test marks, and individual reports on projects, while teachers of lower grade levels attach more importance to effort and neatness of work.

Table 4.4 reports the first four canonical variates that were derived from the correlation matrix in Table 4.3 and the intercorrelations for student characteristics, which appear in Table 4.1. The first canonical variate indicated a canonical R of .603 between the weighted set of student characteristics and the weighted set of situational factors. A high positive loading was found for test marks and a high negative loading for neatness of work. A moderately high positive loading was found for homework marks. Moderately high negative loadings were found for effort and optional work. Extremely low negative loadings were found for both appearance and attendance.

The set of situational factors showed a single positive loading of moderately high magnitude for the number of teachers in the school. The level of school and grade level taught both were found to have high negative loadings. A moderately high negative loading was found for subject area taught.

The second canonical variate indicated a canonical R of .392 between the two sets of measures. The set of

TABLE 4.4.--Significant Canonical Variates for Situational Factors and Student Characteristics.

	First Canonical Variate	Second Canonical Variate	Third Canonical Variate	Fourth Canonical Variate
Student Characteristics				
Attendance	-.011	.363	.632	.094
Appearance	-.028	.147	.629	-.209
Conduct	.059	.149	-.217	.444
Effort	-.314	.189	-.003	-.416
Participation	.072	.025	.092	-.767
Attitudes	-.090	-.350	.269	-.195
Neatness	-.496	-.191	-.229	.747
Homework	.349	-.244	-.171	-.379
Quizzes	.136	-.179	.155	.186
Tests	.720	.016	.204	.074
Optional	-.233	-.084	-.232	-.139
Group Reports	-.098	-.357	-.386	-.038
Individual Reports	.198	.987	-.028	-.033
Situational Factors				
Minorities	-.102	.080	.061	.075
Number in District	-.024	-.002	-.031	-.308
Urbanization	-.053	-.327	-.623	1.464
Grade Taught	-.646	-.184	.665	-.383
Subject Taught	-.251	.003	1.429	-1.361
Number in School	.499	.102	-.126	-.440
School Level	-.768	-.309	1.892	2.798
Canonical R	.603	.392	.372	.326
Canonical R^2	.364	.154	.139	.106
Redundancy	2.796	1.181	1.066	.816
x^2	813.13	498.67	382.64	278.84
df	338	300	264	230
Significance	.0001	.0001	.0001	.0153

student characteristics demonstrated a very high positive loading for individual reports or projects and a moderately high positive loading for attendance. Moderately high negative loadings were found for group reports or projects, attitudes, and appearance. Extremely low positive loadings were found for class participation and test marks.

The set of situational factors showed moderately high negative loadings for urbanization and school level.

The third canonical variate indicated a canonical R of .372 between the two sets of measures. The set of student characteristics demonstrated high positive loadings for attitudes on test marks. A high negative loading was found for group reports or projects and moderately high negative loadings for optional work, neatness of work, and conduct.

The set of situational factors demonstrated very high positive loadings for school level and subject area and a moderately high positive loading for grade level taught. A moderately high negative loading was found for urbanization.

The fourth canonical variate indicated a canonical R of .326 for the two measures. The set of student characteristics demonstrated a high positive loading for neatness of work and a high negative loading for class participation. A moderately high positive loading was found for

conduct. Moderatley high negative loadings were found for effort and homework marks. Extremely low negative loadings were found for both individual and group reports or projects.

The set of situational factors showed a very high positive loading for school level. A high positive loading was found for urbanization and a high negative loading for subject area taught.

With the first canonical variate 36.4 percent of the variation in the weighting applied to student characteristics was statistically accounted for by the situational factors. With the remaining canonical variates 15.4, 13.9 and 10.6 percent of the variation was accounted for by the respective variate. It, therefore, seems that a knowledge of the situational factors particularly in regard to the grade level of the students would provide a reasonably accurate prediction of the comparative importance attached to the student characteristics considered when assigning marks.

> Question 5: What attitudinal factors are significantly related to the weighting applied to student characteristics that are considered when assigning marks.

Findings: The correlation matrix presented in Table 4.5 reports the degree of relationship between the teachers' beliefs regarding the appropriate goals of education and the function of marks and the weighting applied to each student characteristic considered when assigning marks as well as the intercorrelation matrix for the attitudinal factors. The

TABLE 4.5.--Correlation Matrices for Attitudinal Factors and Weighting Applied to Student Characteristics.

	Goals	Functions Teacher	Functions Student	Functions Parents
Goals	1.00			
Functions Teacher	.999	1.00		
Functions Student	-.230	-.231	1.00	
Functions Parents	-.234	-.237	.366	1.00
Attendance	.998	.998	-.232	-.241
Appearance	.799	.799	-.188	-.154
Conduct	-.001	-.001	.001	.052
Effort	.985	.985	-.240	-.238
Participation	.169	.170	-.041	.016
Attitudes	.871	.871	-.207	-.187
Neatness	.002	.002	-.000	.058
Homework	.040	.040	-.007	.066
Quizzes	.872	.872	-.197	-.168
Tests	.006	.006	.040	.116
Optional	.173	.173	-.042	.013
Group Reports	.931	.931	-.216	-.204
Individual	-.001	-.001	.001	.060

intercorrelation matrix for the student characteristics is presented above in Table 4.1.

A consideration of the intercorrelations among the attitudinal factors showed that the teachers' beliefs regarding the appropriate goals of education and their opinions regarding the functions of marks for teachers were almost perfectly related. For this reason, the correlations between these two measures and the student characteristics were identical. A slight negative relationship was found between the functions of marks for teachers, for students, and for parents.

A consideration of the correlations between the attitudinal factors and the student characteristics showed a close relationship between the teachers' viewpoints regarding goals of education and the weighting applied to attendance, appearance, effort, attitudes, quiz marks and group reports or projects. A slight negative relationship was found between both functions of marks for students and parents and the weighting applied to attendance, appearance, effort, attitudes, quiz marks, and group reports.

Table 4.6 reports the first two canonical variates that were derived from the above correlation matrix. The first canonical variate indicated a canonical R of .9997 between the weighted set of student characteristics and the weighted set of attitudinal factors. A high positive loading was found for appearance and a high negative loading for

TABLE 4.6.-- Significant Canonical Variates for Attitudinal Factors and Student Characteristics.

	First Canonical Variate	Second Canonical Variate
Student Characteristics		
Attendance	.129	-.624
Appearance	.615	-.823
Conduct	-.044	-.347
Effort	.074	.927
Participation	-.025	-.679
Attitudes	.088	2.367
Neatness	-.254	2.999
Homework	-.000	-.706
Quizzes	.058	-6.291
Tests	-.001	-2.731
Optional	-.153	2.530
Group Reports	.208	3.763
Individual	-.151	-.397
Attitudinal Factors		
Goals	.126	-7.543
Functions Teacher	.872	7.270
Functions Student	-.004	-.272
Functions Parents	-.003	-.867
Canonical R	.999	.330
Canonical R^2	.999	.109
Redundancy	7.688	.839
X^2	5446.06	107.34
df	52	36
Significance	.0001	.0001

neatness of work. Moderately positive loadings were found for group reports or projects and attendance. Moderately high negative loadings were found for optional work and individual reports or projects.

The set of attitudinal factors showed a high positive loading for the functions of marks for teachers and a moderately high positive loading for the goals of education.

The second significant canonical variate indicates a canonical R of .330 between the two sets of measures. The set of student characteristics demonstrated a high positive loading for both group reports on projects and neatness of work. A high negative loading was found for quiz marks. Moderately high positive loadings were found for optional work and for attitudes. A moderately high negative loading was found for test marks.

The set of attitudinal factors showed a high positive loading for the functions of marks for teachers and a nearly equal negative loading for the goals of education.

With the first canonical variate 99.9 percent of the variation in the weighting applied to student characteristics was statistically accounted for by the teachers' attitudes toward the goals of education and the function of marks. With the second canonical variate 10.9 percent of the variation was explained by the attitudinal factors. It seems, therefore, that a knowledge of the teachers' attitudes regarding the appropriate goals of education and the

functions of marks would provide an accurate prediction of the comparative importance attached to the student characteristics considered when assigning marks.

> Question 7: What other aspects of the marking procedure used is significantly related to the weighting applied to the student characteristics that are considered when assigning marks?

Findings: The correlation matrix presented in Table 4.7 reports the degree of relationship between other aspects of the marking procedure and the weighting applied to each student characteristic that is considered when assigning marks, as well as the intracorrelation matrix for the other aspects. The intercorrelation matrix for the student characteristics is presented above in Table 4.1.

A consideration of the intercorrelations among the other aspects showed a slight relationship between the frequency of use of statistical techniques and the frequency of use of objective measurements.

A consideration of the cross correlations revealed a slight relationship between the frequency of use of statistical techniques and homework, quiz, and test marks. A slight relationship was also found between the frequency of objective measurements and homework, quiz, and test marks while this other aspect showed a slight negative relationship to effort.

Table 4.8 reports the first two canonical variates that were derived from the appropriate correlation matrices.

TABLE 4.7.--Correlation Matrices for Other Aspects of the Marking Procedure and Weightings Applied to Student Characteristics.

Sources	1.00		
Frequency Statistics	.044	1.00	
Frequency Objective	.085	.315	1.00
Attendance	-.036	.057	-.006
Appearance	-.043	.004	-.066
Conduct	-.077	.002	-.099
Effort	-.069	-.155	-.277
Participation	-.035	-.022	-.071
Attitudes	-.057	-.069	-.117
Neatness	-.085	-.020	-.059
Homework	.065	.167	.312
Quizzes	-.001	.152	.272
Tests	.094	.204	.317
Optional	-.032	.024	.019
Group Reports	-.036	.033	.065
Individual	-.028	.004	.042
	Sources	Frequency Statistics	Frequency Objective

TABLE 4.8.--Significant Canonical Variates for other Aspects of the Marking Procedure and Student Characteristics.

	First Canonical Variate	Second Canonical Variate
Student Characteristics		
Attendance	-.096	-.304
Appearance	.121	.920
Conduct	.025	.272
Effort	.485	.084
Participation	.018	.317
Attitudes	.048	-.260
Neatness	.166	-.402
Homework	-.367	.116
Quizzes	-.397	.157
Tests	-.520	.097
Optional	.205	-.006
Group Reports	-.120	-.134
Individual Reports	-.071	-.047
Other Aspects		
Sources Information	.335	.959
Use Statistics	-.263	.157
Use Objective Measures	-.850	.102
Canonical R	.583	.312
Canonical R^2	.340	.097
Redundancy	2.618	.747
X^2	407.06	113.36
df	78	66
Significance	.0001	.0001

The first canonical variate indicated a canonical R of .583 between the weighted set of student characteristics and the weighted set of other aspects. A high positive loading was found for effort and a high negative loading for test marks. Moderately high positive loadings were found for optional work and neatness of work. Moderately high negative loadings were found for quiz marks and homework marks.

The set of other aspects showed a high negative loading for the total frequency of use of objective measurements. Sources of useful information were found to have the only positive loading.

The second significant canonical variate indicated a canonical R of .312 between the two measures. The set of student characteristics demonstrated an extremely high positive loading for appearance and a high positive loading for class participation. A moderately high positive loading was found for conduct. A high negative loading was found for neatness of work and moderately high loadings for attendance and attitudes. An extremely low negative loading was found for optional work.

The set of other aspects showed all positive loading. Sources of useful information was found to have a high loading.

With the first canonical variate 34.0 percent of the variation in the weighting applied to each student characteristic was statistically accounted for by the other aspects

of the marking procedure. With the second canonical variate 9.7 percent of the variation was accounted for. It seems, therefore, that a knowledge of the other procedures that are related to the practice of marking students would provide a reasonably accurate prediction of the comparative importance attached to the student characteristics considered when assigning marks.

Chi Square Tests of Independence

Question 2: What personal factors are significantly related to the basis on which marks are assigned?

1. Findings: As indicated in Table 4.9, the basis on which marks are assigned was found to be significantly related to the sex of the teachers. Table 4.10 shows the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square value.

TABLE 4.9.--Chi Square Analyses for Personal Factors and Basis for Marks.

	X^2	df
Sex	43.935***	6
Age	43.487**	24
Experience	38.482	30
Degree	34.408	24
Number Courses	14.310	18
GPA	9.815	12

**Indicates X^2 significant at α equal to .01.
***Indicates X^2 significant at α equal to .001.

TABLE 4.10.--Chi Square Table for Sex and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Male (282)	Frequency	24	37	13	104	27	68	9
	Theoretical	33.82	31.86	7.47	128.61	27.53	43.26	9.44
	Chi Square	2.85	.83	4.09*	4.71*	.01	14.14**	.02
Female (435)	Frequency	62	44	6	223	43	42	15
	Theoretical	52.18	49.14	11.53	198.39	42.47	66.74	14.56
	Chi Square	1.85	.54	2.65	3.05	.01	9.17**	.01

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

Male teachers appear to base marks either on performance relative to that of others or on a combination of subject matter mastery and relative performance more often than was expected. Male teachers also appear to be less likely than female teachers to base marks on a combination of self growth and subject matter mastery.

Female teachers were less likely than male teachers to use a combination of subject matter mastery and relative performance than was expected.

2. Findings: As indicated in Table 4.9, the basis on which marks are assigned was found to be significantly related to the age of the teacher. Table 4.11 shows the joint frequency distribution for the two maasures as well as the theoretical expected frequency and the cell Chi square value.

Those teachers in the 46 to 55 year old group appear to less frequently use subject matter mastery alone as the basis for marks.

Those teachers in the oldest age group, 56 years and older, appear more likely to use the students' growth as the basis for assigning marks than was expected.

Question 4: What situational factors are significantly related to the basis on which marks are assigned?

1. Findings: As indicated in Table 4.12, the basis on which marks are assigned was found to be significantly related to the type of school within which the teachers

TABLE 4.11.--Chi Square Table for Age and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
20 - 25 (127)	Frequency	9	15	2	68	14	16	3
	Theoretical	15.21	14.33	3.36	58.02	12.38	19.46	4.25
	Chi Square	2.54	.03	.55	1.72	.21	.61	.37
26 - 35 (265)	Frequency	25	27	11	120	26	45	11
	Theoretical	31.74	29.90	7.01	121.06	25.84	40.60	8.86
	Chi Square	1.43	.28	2.27	.01	.00	.48	.52
36 - 45 (155)	Frequency	24	24	1	62	13	27	4
	Theoretical	18.57	17.49	4.10	70.81	15.11	23.75	5.18
	Chi Square	1.59	2.43	2.35	1.10	.30	.45	.27
46 - 55 (98)	Frequency	13	4	3	48	9	17	4
	Theoretical	11.74	11.06	2.59	44.77	9.55	15.01	3.28
	Chi Square	.14	4.50*	.06	.23	.03	.26	.16
56 & over (72)	Frequency	15	11	2	30	7	5	2
	Theoretical	8.74	8.23	1.94	33.35	7.12	11.18	2.44
	Chi Square	4.83*	1.13	.03	.71	.03	3.45	.10

*Indicates X_1^2 significant at α equal to .05.

were assigned, as determined by the grade levels in that school.

TABLE 4.12.--Chi Square Analyses for Situational Factors and Basis for Marks.

	X^2	df
Type of School (Grade Levels)	80.061***	24
Grade Level Taught	113.965***	30
Subject Taught	191.305***	54
Academic - Nonacademic	72.668***	12
Size of District (Within Stratum)	18.390	18
Size of School (Within Stratum)	33.643	24
Percent Minorities	27.506	24
Urbanization Level	35.487	24

***Indicates X^2 significant at α equal to .001.

Table 4.13 shows the joint frequency distribution for the two measures, as well as the theoretical expected frequency and the cell Chi square values.

Teachers in either a senior high school or junior high school were less likely to base marks on the students self growth than was expected. On the other hand, teachers in these schools are more likely to base marks on a combination of subject matter mastery and relative performance. This latter finding was true for both separated and combined schools.

TABLE 4.13.--Chi Square Table for Level of School and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Senior High 213	Frequency	8	20	7	69	20	33	6
	Theoretical	19.53	18.45	4.33	74.69	15.94	25.05	5.47
	Chi Square	12.55***	.26	2.18	.72	1.59	5.05*	.39
Junior High Middle School 214	Frequency	9	20	6	70	19	33	7
	Theoretical	19.53	18.45	4.33	74.69	15.94	25.05	6.47
	Chi Square	12.55***	.26	2.18	.72	1.59	5.05*	.39
Elementary (329)	Frequency	65	30	3	168	25	28	10
	Theoretical	39.41	37.12	8.71	150.30	32.08	50.40	11.00
	Chi Square	16.62***	1.36	3.74	2.09	1.56	9.96**	.09
Senior-Junior High (48)	Frequency	3	9	2	13	6	14	1
	Theoretical	5.75	5.42	1.27	21.93	4.68	7.35	1.60
	Chi Square	1.31	2.37	.42	3.63	.37	6.01*	.23
Junior High Elementary (9)	Frequency	0	1	1	5	0	2	0
	Theoretical	1.08	1.02	.24	4.11	.88	1.38	.30
	Chi Square	1.08	.00	2.44	.19	.88	.28	.30
All (5)	Frequency	1	1	0	3	0	0	0
	Theoretical	.60	.56	.13	2.28	.49	.77	.17
	Chi Square	.27	.34	.13	.22	.49	.77	.17

*Indicates X^2_1 significant at α equal to .05.

**Indicates X^2_1 significant at α equal to .01

***Indicates X^2_1 significant at α equal to .001.

Teachers in elementary schools were found to be disproportionately likely to base marks on the students' self growth. These teachers were also less likely to utilize a combination of subject matter mastery and relative performance.

2. Findings: As indicated in Table 4.12 the grade level taught by the teacher was found to be significantly related to the basis on which marks are assigned. Table 4.14 shows the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square value.

Teachers of the lower elementary grade levels, grades kindergarten through fourth, were found to be more likely to base marks on the students' self growth. Teachers of grades kindergarten through second were also more likely to utilize a combination of growth and subject matter mastery and were less likely to utilize a combination of subject matter mastery and relative performance as the basis for marks.

Teachers of grades five and six were less likely to use a combination of subject matter mastery and relative performance as the basis for marks.

On the other hand, teachers of grades seven through twelve are more likely to use a combination of subject matter mastery and relative performance while they are less likely to use growth as the basis for marks. Teachers of grades ten through twleve were also found to be more likely to use the

TABLE 4.14.--Chi Square Table for Grade Level Taught and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
K-1-2 (127)	Frequency	27	10	0	73	9	4	4
	Theoretical	15.21	14.33	3.36	58.02	12.38	19.46	4.25
	Chi Square	9.14**	1.31	3.36	3.87*	.92	12.28***	.01
3-4 (115)	Frequency	22	14	1	52	10	11	5
	Theoretical	13.77	12.97	3.04	52.53	11.21	17.62	3.84
	Chi Square	4.91*	.08	1.37	.01	.13	2.49	.35
5-6 (101)	Frequency	18	9	3	55	9	5	2
	Theoretical	12.10	11.39	2.67	46.14	9.85	15.47	3.38
	Chi Square	2.88	.50	.04	1.70	.07	7.09**	.56
7-8-9 (176)	Frequency	11	22	4	71	22	40	6
	Theoretical	21.08	19.86	4.66	80.40	17.16	26.96	5.88
	Chi Square	4.82*	.23	.09	1.10	1.37	6.30*	.00
10-11-12 (187)	Frequency	6	26	11	71	17	50	6
	Theoretical	22.40	21.10	4.95	85.43	18.23	28.65	6.25
	Chi Square	12.01***	1.14	7.40**	2.44	.08	15.91***	.01
Nongraded Elementary (11)	Frequency	2	0	0	5	3	0	1
	Theoretical	1.32	1.24	.29	5.03	1.07	1.69	.37
	Chi Square	.35	1.24	.29	.00	3.46	1.69	1.09

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

students' performance when compared to others as the basis for marks.

Teachers of non-graded elementary classes were not found to be significantly different from other teachers in regard to the basis for marks. This finding may have resulted from the small number of respondents in this classification.

3. Findings: As indicated in Table 4.12, the basis on which marks are assigned was found to be significantly related to whether the teachers taught elementary students, academic, or nonacademic subject areas. Table 4.15 shows the joint frequency distribution of the two measures as well as the theoretical expected frequency and the cell Chi square values.

Secondary teachers in academic subject areas more often than was expected reported that they used a combination of subject matter mastery and relative performance as the basis for marks. On the other hand, secondary teachers in academic subject areas are less likely to use growth as the basis.

As reported above, elementary teachers were found to be less likely to use a combination of subject matter and relative performance and more likely to use growth.

4. Findings: As indicated in Table 4.12, similar findings were made regarding the particular subject area taught, as related to the basis for marks. Table 4.16 shows

TABLE 4.15.--Chi Square Table for Academic versus Nonacademic and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Academic (346)	Frequency	24	45	12	149	34	72	10
	Theoretical	41.02	39.09	9.17	158.28	33.78	53.08	11.58
	Chi Square	7.06**	.89	.87	.54	.00	6.74**	.22
Nonacademic (86)	Frequency	4	10	5	29	12	20	6.00
	Theoretical	10.20	9.72	2.28	39.34	8.40	13.19	2.88
	Chi Square	3.76	.01	3.25	2.72	1.55	3.51	3.38
Elementary (281)	Frequency	55	26	2	149	23	18	8.00
	Theoretical	33.31	31.74	7.45	128.55	27.43	43.11	9.41
	Chi Square	14.12***	1.04	3.98*	3.25	.72	14.63***	.21

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

TABLE 4.16.--Chi Square Table for Subject Area Taught and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Science (58)	Frequency	4	6	3	18	7	17	3
	Theoretical	6.88	6.55	1.54	26.45	5.66	8.98	1.94
	Chi Square	1.20	.05	1.39	2.70	.32	7.17**	.58
Mathematics (69)	Frequency	4	14	4	22	2	22	1
	Theoretical	8.18	7.79	1.83	31.47	6.74	10.68	2.31
	Chi Square	2.14	4.94*	2.58	2.85	3.33	11.99***	.74
Social Studies (84)	Frequency	6	16	2	29	10	19	2
	Theoretical	9.96	9.49	2.23	38.31	8.20	13.00	2.81
	Chi Square	1.57	4.47*	.02	2.26	.39	2/76	.23
Language Arts (122)	Frequency	9	8	3	74	15	9	4
	Theoretical	14.46	13.78	3.23	55.64	11.91	18.89	4.08
	Chi Square	2.06	2.43	.02	6.06*	.80	5.18*	.00
Foreign Language (14)	Frequency	1	1	0	7	0	5	0
	Theoretical	1.66	1.58	.37	6.38	1.37	2.17	.47
	Chi Square	.26	.21	.37	.06	1.37	3.70	.47
Business Commercial (21)	Frequency	0	6	0	5	1	9	0
	Theoretical	2.49	2.37	.56	9.58	2.05	3.25	.70
	Chi Square	2.49	5.55*	.56	2.19	.54	10.17**	.70
Music Art (19)	Frequency	2	1	3	8	3	1	1
	Theoretical	2.25	2.15	.50	8.67	1.85	2.94	.64
	Chi Square	.03	.61	12.38***	.05	.71	1.28	.21
Industrial Arts (25)	Frequency	1	0	2	9	7	5	1
	Theoretical	2.96	2.82	.66	11.40	2.44	3.87	.84
	Chi Square	1.30	2.82	2.70	.51	8.52**	.33	.03
Physical Education Health (21)	Frequency	0	1	0	7	2	5	4
	Theoretical	2.49	2.37	.56	9.58	2.05	3.25	.70
	Chi Square	2.49	.17	.56	.69	.00	.94	15.46***
Elementary All Subjects (284)	Frequency	58	26	2	148	23	19	8
	Theoretical	33.67	32.08	7.53	129.52	27.73	43.97	9.51
	Chi Square	17.58***	1.15	4.06	2.64	.91	14.18***	.24

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

The use of a combination of subject matter mastery and relative performance is reported by more science, mathematics and business or commercial teachers.

The use of a combination of growth and relative performance is reported by more home economics or industrial arts teachers.

The use of a combination of growth and subject matter mastery is reported by more language arts teachers. Language arts teachers were also less likely to use a combination of mastery and relative performance.

The use of subject matter mastery is reported by more mathematics and social studies teachers.

The use of relative performance is reported by more music or art teachers.

Teachers of physical education or health class appear to be less likely to utilize any of the suggested bases for marks.

The reponses from foreign language teachers were not found to be significantly different from chance expectations based on the total sample of responding teachers; however, few foreign language teachers were included in the sample.

Question 6: What attitudinal factors are significantly related to the basis on which marks are assigned?

1. Findings: As indicated in Table 4.17, the basis on which marks are assigned was found to be significantly related to the teachers' reported viewpoint regarding the appropriate goals of education. Table 4.18 shows the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square value.

TABLE 4.17.--Chi Square Analysis for Attitudinal Factors and Basis for Marks.

	X^2	df
Goals of Education	65.107***	30
Function of Marks for Teacher	11.395	12
Function of Marks for Student	53.992***	18
Function of Marks for Parents	84.646***	18

***Indicates X^2 significant at α equal to .001.

Although Table 4.18 demonstrated clearly that the majority of teachers indicated a belief in the traditional goals of education, findings, regarding the relationship between this belief and the basis on which marks are assigned, are somehwat less clear.

Teachers, expressing a moderately progressive view, regarding the appropriate goals of education, were found to be more likely to base marks on growth and less likely to utilize subject matter mastery. This group of teachers were also more likely to report that they used none of the suggested bases.

TABLE 4.18.--Chi Square Table for Goals of Education and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Very	Frequency	1	2	1	2	1	7	0
Progressive	Theoretical	1.66	1.58	.37	6.40	1.37	2.15	.47
(14)	Chi Square	.26	.11	1.07	3.03	.10	10.96***	.47
Moderate	Frequency	10	0	0	19	5	2	4
Progressive	Theoretical	4.74	4.54	1.06	18.30	3.91	6.14	1.34
(40)	Chi Square	5.83*	4.54*	1.06	.03	.31	2.79	5.29*
Low	Frequency	21	15	0	82	15	14	1
Progressive	Theoretical	17.55	16.72	3.92	67.70	14.45	22.71	4.95
(148)	Chi Square	.68	.18	3.92*	3.02	.02	3.34	3.16
Low	Frequency	36	41	11	164	30	54	13
Traditional	Theoretical	41.37	39.43	9.25	159.65	34.07	53.54	11.68
(349)	Chi Square	.70	.06	.33	.12	.49	.00	.15
Moderate	Frequency	15	16	6	50	15	29	5
Traditional	Theoretical	16.12	15.36	3.60	62.21	13.28	20.86	4.55
(136)	Chi Square	.08	.03	1.59	2.40	.22	3.17	.04
Very	Frequency	2	7	1	11	4	4	1
Traditional	Theoretical	3.56	7.39	.79	13.72	2.93	4.60	1.00
(30)	Chi Square	.68	3.85*	.05	.54	.39	.08	.00

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

Teachers, expressing a low progressive view, were found to be less likely to use relative performance as the basis for marks.

On the other hand, teachers, expressing a highly progressive view, were found to be more likely to utilize a combination of subject matter mastery and relative performance, while those, expressing a highly traditional view, were less likely to utilize subject matter mastery.

2. Findings: As indicated in Table 4.17, the basis on which marks are assigned was found to be significantly related to the viewpoint of the teacher regarding the degree to which marks serve a useful function for students. Table 4.19 presents the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

The majority of teachers reported that marks are at least moderately useful for students.

Those teachers, who indicated that marks did not serve a useful function for students, were more likely to either base marks on growth or to report that they did not utilize any of the suggested bases.

Teachers, who indicated that marks were very useful for students, were more likely to base marks on relative performance.

3. Findings: As indicated in Table 4.17, the basis on which marks are assigned was found to be significantly

TABLE 4.19.--Chi Square Table for Functions of Marks for Students and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Useless (89)	Frequency	21	13	0	29	11	7	8
	Theoretical	10.55	10.05	2.36	40.71	8.69	13.65	2.98
	Chi Square	10.35**	.86	2.36	3.37	.61	3.24	8.46
Little Use (16)	Frequency	3	2	0	7	1	1	2
	Theoretical	1.90	1.81	.42	7.32	1.56	2.45	.54
	Chi Square	.64	.02	.42	.01	.20	.86	4.00*
Moderately Useful (432)	Frequency	42	41	9	219	37	75	9
	Theoretical	51.21	48.80	11.45	197.62	42.18	66.28	14.46
	Chi Square	1.66	1.25	.52	2.31	.64	1.15	2.06
Very Useful (180)	Frequency	19	25	10	73	21	27	5
	Theoretical	21.34	20.33	4.77	82.34	17.57	27.62	6.03
	Chi Square	.26	1.07	5.73*	1.06	.67	.01	.17

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

related to the viewpoint of the teacher regarding the degree to which marks serve a useful function for parent. Table 4.20 presents the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

The majority of teachers reported that marks are at least moderately useful to parents.

Those teachers, who indicated that marks are useless to parents, were more likely either to base marks on growth or on none of the suggested bases.

Those, indicating that marks are moderately useful to parents, appear to be more likely to use a combination of growth and subject matter mastery and are less likely to use a combination of subject matter mastery and relative performance.

To the contrary, those teachers, who expressed a belief that marks are very useful to parents, showed a lower likelihood of utilizing a combination of growth and mastery and a higher likelihood of utilizing a combination of subject matter mastery and relative performance. This group was also less likely to use growth as the basis for marks.

Question 8: What other aspects of the marking procedure used are significantly related to the basis on which marks are assigned?

1. Findings: As indicated in Table 4.21, the basis on which marks are assigned was found to be significantly

TABLE 4.20.--Chi Square Table for Functions of Marks for Parents and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Useless (53)	Frequency	15	5	0	18	6	4	5
	Theoretical	6.28	5.99	1.40	24.25	5.17	8.13	1.77
	Chi Square	12.09***	.16	1.40	1.61	.13	2.10	5.87*
Little Use (20)	Frequency	5	3	0	9	1	1	1
	Theoretical	2.37	2.26	.53	9.15	1.95	3.07	.67
	Chi Square	2.92	.24	.53	.00	.46	1.39	.16
Moderately Useful (448)	Frequency	54	43	10	235	44	48	14
	Theoretical	53.11	50.61	11.87	204.94	43.74	68.73	15
	Chi Square	.01	1.14	.30	4.41*	.00	6.25*	.07
Very Useful (196)	Frequency	11	30	9	66	19	57	4
	Theoretical	23.24	22.14	5.19	89.66	19.14	30.07	6.56
	Chi Square	6.44*	2.79	2.79	6.24*	.00	24.12***	1.00

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

related to the sources of information which were reported by teachers as being useful. Table 4.22 presents the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

TABLE 4.21.--Chi Square Analyses for Other Aspects of the Marking Procedure and Basis for Marks.

	X^2	df
Sources of useful information	108.018***	24
Total frequency objective measures	149.689***	60
Frequency quizzes	76.902***	18
Frequency tests	59.817***	18
Total frequency use of statistics		
Frequency use of standard scores	6.911	6
Frequency use of percentile ranks	19.454**	6
Frequency use of normal curve	43.519***	6

Teachers, who reported that parents were a source of useful information, were more likely to base marks on growth than was predicted.

Teachers, who reported that the students themselves were a source of useful information, were less likely to utilize none of the suggested bases.

Teachers who reported that they found none of the listed sources of information useful when assigning marks,

TABLE 4.22.--Chi Square Table for Sources of Useful Information and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Parents (14)	Frequency	6	4	4	2	0	1	1
	Theoretical	1.68	1.58	.37	6.40	1.36	2.14	.47
	Chi Square	11.15***	3.71	.37	3.02	1.36	.61	.60
Student (567)	Frequency	69	56	11	283	58	81	9
	Theoretical	67.91	63.97	15.00	259.02	55.28	86.87	18.95
	Chi Square	.02	.99	1.07	2.22	.13	.40	5.23*
Teachers (3)	Frequency	1	0	0	0	0	0	2
	Theoretical	.36	.34	.08	1.37	.29	.46	10
	Chi Square	1.14	.34	.08	1.37	.29	.46	35.99***
Records (19)	Frequency	2	3	1	8	2	3	0
	Theoretical	2.28	2.14	.50	8.68	1.85	2.91	.64
	Chi Square	.03	.34	.49	.05	.01	.00	.64
None (115)	Frequency	8	18	7	35	10	25	12
	Theoretical	13.77	12.97	3.04	52.53	11.21	17.62	3.84
	Chi Square	2.42	1.95	5.14*	5.85*	.13	3.09	17.30***

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

appear to base marks on either relative performance or to utilize none of the suggested bases. On the other hand, this group of teachers were less likely to use a combination of subject matter mastery and growth.

2. Findings: As indicated in Table 4.21, the basis for marks was found to be significantly related to the overall total frequency of use of objective measurements of student performance. Table 4.23 presents the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

Those teachers, who reported that they very seldom employed objective measurements of student performance, were found to be more likely to base marks on growth or to use none of the suggested bases. On the other hand, this group were less likely to use either subject matter mastery or a combination of mastery and relative performance.

Those teachers at the frequent end of the continuum were found to be more likely to base marks on either subject matter mastery or a combination of mastery and relative performance.

3. Findings: As indicated in Table 4.21, the basis on which marks are assigned was found to be significantly related to the frequency of use of each individual objective measurement, homework, quizzes, and tests. Table 4.24 presents the joint frequency distribution for the basis for

TABLE 4.23.--Chi Square Table for Frequency of Objective Measurements and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Very Seldom (57)	Frequency	20	1	0	21	7	0	8
	Theoretical	6.90	6.31	1.41	26.17	5.48	8.72	1.99
	Chi Square	24.90***	4.47*	1.41	1.02	.42	8.72**	18.09***
(30)	Frequency	7	1	1	15	1	4	1
	Theoretical	3.63	3.32	.74	13.78	2.89	4.59	1.05
	Chi Square	3.13	1.62	.09	.11	1.23	.08	.00
(27)	Frequency	7	1	1	13	2	2	1
	Theoretical	3.27	2.99	.67	12.40	2.60	4.13	.94
	Chi Square	4.27*	1.33	.16	.03	.14	1.10	.00
(23)	Frequency	2	5	0	9	5	0	2
	Theoretical	2.78	2.55	.57	10.56	2.21	3.52	.80
	Chi Square	.22	2.36	.57	.23	3.51	3.52	1.78
(53)	Frequency	3	6	2	25	9	6	2
	Theoretical	6.41	5.87	1.31	24.34	5.10	8.11	1.85
	Chi Square	1.82	.00	.36	.02	2.93	.55	.01
Average (66)	Frequency	9	7	0	36	7	5	2
	Theoretical	7.99	7.31	1.64	30.31	6.35	10.10	2.31
	Chi Square	.13	.01	1.64	1.07	.07	2.58	.04
(56)	Frequency	11	5	0	30	3	5	2
	Theoretical	6.78	6.20	1.39	25.71	5.39	8.57	1.96
	Chi Square	2.63	.23	1.39	.71	1.06	1.59	.00
(79)	Frequency	6	11	1	39	8	13	1
	Theoretical	9.56	8.75	1.96	36.28	7.60	12.09	2.76
	Chi Square	1.32	.58	.47	.20	.02	.07	1.13
(98)	Frequency	6	8	4	42	12	22	4
	Theoretical	11.86	10.86	2.43	45.00	9.43	15.00	3.43
	Chi Square	2.89	.75	1.02	.20	.70	3.27	.10
(114)	Frequency	7	21	5	46	5	29	1
	Theoretical	13.79	12.63	2.83	52.35	10.97	17.45	3.99
	Chi Square	3.35	5.55*	1.67	.77	3.25	7.65**	2.24
Very Often (83)	Frequency	5	10	3	39	7	19	0
	Theoretical	10.04	9.20	2.06	38.11	7.99	12.70	0
	Chi Square	2.53	.07	.43	.02	.12	3.12	2.90

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01

***Indicates X_1^2 significant at α equal to .001.

marks and the frequency of homework as well as the theoretical expected frequency and cell Chi square values.

Teachers who reported that they never assign homework are more likely to base marks on growth and are less likely to use either relative performance or a combination of subject matter mastery and relative performance.

On the other hand, teachers who reported that they assign homework are more likely to base marks on a combination of mastery and relative performance and less likely to use growth as the basis for marks.

Table 4.25 presents the joint frequency distribution for the basis on which marks are assigned and the frequency of quizzes as well as the theoretical expected frequency and cell Chi square values.

Those teachers, who reported that they did not give quizzes, were more likely to base marks on growth or to utilize none of the suggested bases. This group of teachers were also less likely to base marks on either subject matter mastery or on a combination of mastery and relative performance.

Table 4.26 presents the joint frequency distribution for the basis on which marks are assigned and the frequency of tests as well as the theoretical expected frequency and the cell Chi square value.

Those teachers, who reported that they did not give tests, were found to be more likely to base marks on growth

TABLE 4.24.--Chi Square Table for Use of Homework and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Never (282)	Frequency	56	26	2	148	23	19	8
	Theoretical	33.43	31.86	7.47	128.61	27.53	43.65	9.44
	Chi Square	15.24***	1.08	4.01*	2.92	.75	13.93***	.22
Use (431)	Frequency	28	55	17	176	47	92	16
	Theoretical	51.09	48.69	11.42	196.56	42.08	69.32	14.43
	Chi Square	10.44**	.82	2.73	2.15	.58	9.58**	.17

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

TABLE 4.25.--Chi Square Table for Frequency of Quizzes and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Not	Frequency	29	4	4	37	16	8	13
Used	Theoretical	13.50	12.40	2.67	50.71	10.88	17.11	3.72
(111)	Chi Square	17.79***	5.69*	.66	3.71	2.46	4.85*	22.62***
Seldom	Frequency	11	9	0	49	9	14	2
Used	Theoretical	11.43	10.50	2.26	42.94	9.17	14.49	3.19
(94)	Chi Square	.02	.22	2.26	.85	.00	.02	.44
Often	Frequency	26	34	7	144	22	50	7
Used	Theoretical	35.28	32.40	6.97	132.49	28.30	44.71	9.84
(290)	Chi Square	2.44	.08	.00	1.00	1.40	.63	.82
Very	Frequency	20	32	6	93	22	37	2
Often Used	Theoretical	25.79	23.69	5.10	98.85	20.69	32.68	7.20
(212)	Chi Square	1.30	2.92	.16	.15	.08	.57	3.75

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X^2 significant at α equal to .001.

TABLE 4.26.--Chi Square Table for Frequency of Tests and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Not	Frequency	51	18	1	124	23	18	10
Used	Theoretical	30.58	27.70	6.84	111.53	23.74	35.98	8.63
(245)	Chi Square	13.64***	3.40	4.98*	1.39	.02	8.98**	.22
Seldom	Frequency	1	0	0	0	0	0	0
Used	Theoretical	.12	.11	.03	.46	.10	.15	.04
(1)	Chi Square	6.14*	.11	.03	.46	.10	.15	.04
Often	Frequency	23	36	11	117	28	52	10
Used	Theoretical	34.57	31.32	7.73	126.09	26.85	40.68	9.76
(277)	Chi Square	3.87*	.70	1.39	.66	.05	3.15	.01
Very	Frequency	10	23	7	69	15	30	4
Often Used	Theoretical	19.72	17.86	4.41	71.92	15.31	23.20	5.57
(158)	Chi Square	4.79*	1.48	1.52	.12	.01	1.99	.44

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

and less likely to use either relative performance or a combination of relative performance and mastery of subject matter.

On the other hand, those teachers who reported that they often gave tests, were found to be less likely to base marks on growth.

4. Findings: As indicated in Table 4.21, the basis on which marks are assigned was found to be significantly related to the overall total frequency of use of specified statistical techniques. Table 4.27 presents the joint frequency distribution for the two measures as well as the theoretical expected frequency and the cell Chi square values.

Those teachers who reported that they did not use the suggested statistical techniques were more likely to either use growth as the basis for marks or to use none of the bases listed. This group were less likely to use a combination of mastery and relative performance.

Those teachers who seldom used statistical techniques were more likely to base marks on subject matter mastery.

Those who often used these methods were less likely to use growth as the basis for marks and more likely to use a combination of subject matter mastery and relative performance.

The teachers, reporting that they always used statistical procedures in the determination of marks were more

TABLE 4.27.--Chi Square Table for Frequency of use of Statistical Techniques and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Never (260)	Frequency	47	21	5	135	19	18	15
	Theoretical	30.82	29.40	6.89	118.58	25.38	40.25	8.70
	Chi Square	8.49**	2.39	.52	2.27	1.61	12.40***	4.56*
Seldom (236)	Frequency	25	39	6	93	24	46	3
	Theoretical	27.98	26.66	6.26	107.63	23.04	36.39	7.90
	Chi Square	.32	5.71*	.01	1.99	.04	2.45	3.04
Often (157)	Frequency	5	16	6	75	15	37	3
	Theoretical	18.61	17.74	4.16	71.60	15.33	24.31	5.26
	Chi Square	9.96**	.17	.81	.16	.01	6.63**	.97
Always (64)	Frequency	8	5	2	24	12	10	3
	Theoretical	7.59	7.23	1.70	29.23	6.25	9.94	2.14
	Chi Square	.02	.69	.05	.92	5.29*	.00	.34

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

***Indicates X_1^2 significant at α equal to .001.

likely to base marks on a combination of growth and relative performance.

5. Findings: As indicated in Table 4.21, the basis on which marks are assigned was found to be significantly related to the use of percentile ranks and to the use of a normal curve. Table 4.28 presents the joint frequency distribution for the basis for marks and the use of percentile ranks as well as the theoretical expected frequency and cell Chi square values.

Those teachers, who reported the use of percentile ranks in the determination of marks, appear to be more likely to use a combination of subject matter mastery and relative performance as the basis for marks and are also less likely to base marks on growth.

Table 4.29 presents the joint frequency distribution for the basis for marks and the use of the normal curve as well as the theoretical expected frequency and the cell Chi square values.

Those teachers, who reported the use of the normal curve in the determination of marks, were found to be more likely to utilize relative performance independently or in combination with either growth or subject matter mastery as the basis for marks.

Those teachers, who reported that they do not use the normal curve, were found to be less likely to use a combination of mastery and relative performance.

TABLE 4.28.--Chi Square Table for Use of Percentile Ranks and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Not Used (493)	Frequency	68	49	12	219	44	56	18
	Theoretical	55.82	52.57	12.33	212.88	45.43	71.39	15.58
	Chi Square	2.66	.24	.01	.18	.05	3.32	.38
Used (225)	Frequency	18	32	7	109	26	54	6
	Theoretical	30.18	28.43	6.67	115.12	24.57	38.61	8.42
	Chi Square	4.92*	.45	.02	.33	.08	6.14*	.70

*Indicates X_1^2 significant at α equal to .05.

TABLE 4.29.--Chi Square Table for Use of Normal Curve and Basis for Marks.

		Growth	Mastery	Performance	Growth Mastery	Growth Performance	Mastery Performance	None
Not Used (453)	Frequency	65	58	6	223	32	51	18
	Theoretical	54.26	51.10	11.99	206.94	44.16	69.40	15.14
	Chi Square	2.13	.93	2.99	1.25	3.35	4.88*	.54
Used (265)	Frequency	21	23	13	105	38	59	6
	Theoretical	31.74	29.90	7.01	121.06	25.84	40.60	8.86
	Chi Square	3.63	1.59	5.11*	2.13	5.73*	8.34**	.92

*Indicates X_1^2 significant at α equal to .05.

**Indicates X_1^2 significant at α equal to .01.

Tests for Agreement

Question 9: Do teachers agree in regard to the relative importance of student characteristics considered when assigning marks?

Table 4.30 presents the frequency with which each weighting for each student characteristic was reported. Four student characteristics were considered by the majority of teachers. These characteristics were effort, class participation, quiz marks, and test marks. Approximately one half of the teachers reported the use of homework marks and individual reports or projects when assigning marks. Slightly over one third of the teachers reported the consideration of conduct, optional work, group reports or projects, and neatness of work.

Table 4.31 presents the frequency and percent of teachers according to the relative importance of each student characteristic. Approximately one half of the teachers reported that they considered test marks of greatest importance when assigning marks. Effort was considered of greatest importance by 40.7 percent of the teachers. Equal proportions, 28.6 percent, reported that class participation and quiz marks was second in importance.

In order to determine whether there was agreement among the teachers in regard to the relative importance of each student characteristic, a Kendall's coefficient of concordance was computed. The value for this statistic was

TABLE 4.30.--Relative Weightings for Student Characteristics.

	1-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-99	100%	Totals*
Attendance	96/45.5	91/43.1	24/11.4	10/ 4.7		10/ 4.7				2/ .9		211/28.9
Appearance	26/52.0	19/38.0		2/ 4.0		3/ 6.0						50/ 6.9
Conduct	105/39.0	106/39.4	29/10.8	15/ 5.6	6/2.2	8/ 3.0						269/37.5
Effort	74/12.5	138/23.3	125/21.1	103/17.4	27/4.5	69/11.6	11/1.9	2/ .3	27/4.5	11/1.9	6/1.0	593/82.7
Participation	126/24.0	187/35.6	99/18.8	80/15.2	6/1.1	22/ 4.2			6/1.1			526/73.3
Attitudes	111/31.2	128/35.9	54/15.2	38/10.7		8/ 2.2	5/1.4		4/1.1	5/1.4	3/ .8	356/49.6
Neatness	152/53/0	92/32.1	19/ 6.6	13/ 4.5	8/2.8	3/ 1.0						287/40.0
Homework	89/24.7	117/32.4	70/19.4	61/16.9	9/2.5	10/ 2.7	5/1.4			2/ .6		361/50.4
Quizzes	76/15.4	150/30.4	129/26.2	97/19.5	16/3.2	21/ 4.3	4/ .8					493/68.8
Tests	28/ 5.0	81/14.3	109/19.3	148/26.2	61/10.8	92/16.3	14/2.5	12/2.1	15/2.7	3/ .5	2/ .4	563/78.5
Optional	128/47.9	94/35.2	19/ 7.1	2/ .7	10/3.7	14/ 5.2						267/37.2
Group Reports	104/36.7	98/34.6	39/13.8	27/ 9.5	3/1.1	12/ 4.2						283/39.4
Individual	94/26.0	130/36.0	65/17.7	42/11.6	6/1.7	24/ 6.6						361/50.3

*Frequency/percentage

TABLE 4.31.--Rankings of Importance of Student Characteristics.*

	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth	Ninth
Attendance	45/ 6.3	81/11.3	163/22.7	208/29.0	152/21.2	48/617	12/1.7	7/1.0	1/.1
Appearance	26/ 3.6	32/ 4.5	126/17.6	250/34.9	172/24.0	81/ 11.3	20/2.8	8/1.1	2/.3
Conduct	43/ 6.0	86/12.1	185/25.8	215/30.0	133/18.5	44/ 6.1	8/1.1	3/ .4	
Effort	293/40.7	172/24.0	139/19.4	68/ 9.5	36/ 5.0	7/ 1.0	2/ .3		
Participation	102/14.2	205/28.6	201/28.0	139/19.4	52/ 7.2	18/ 2.5			
Attitudes	64/ 8.9	129/18.0	204/28.4	177/24.7	106/14.8	33/ 4.6	4/ .6		
Neatness	32/ 4.5	69/ 9.6	199/27.7	236/39.9	131/18.3	42/ 5.9	7/1.0	1/ .1	
Homework	83/11.6	141/19.7	170/23.8	181/25.2	105/14.6	21/ .1	12/1.7	3/ .4	1/.1
Quizzes	123/17.1	205/28.6	195/27.2	133/18.5	44/ 6.1	11/ 1.5	5/ .7	1/ .1	
Tests	363/50.6	132/18.4	105/14.6	77/10.7	30/ 4.2	8/ 1.1	2/ .3		
Optional	37/ 5.2	75/10.5	188/26.2	232/32.4	131/18.3	46/ 6.4	6/ .8	2/.3	
Group Reports	48/ 6.7	101/14.1	180/25.1	218/30.4	120/16.7	43/ 6.0	5/ .7	2/.3	
Individual	84/11.7	140/19.5	180/25.1	193/26.9	86/12.0	29/ 4.1	4/ .5	1/.1	

*Frequency/percentage

.3584, indicating at the .001 level of confidence that there was slight, but statistically significant agreement among the teachers regarding this measure.

Question 10: Do teachers agree in regard to the basis on which marks are assigned?

In order to determine whether teachers agree in regard to the basis on which marks are assigned, a Chi square one sample test was utilized. The Chi square one sample test determined the probability that the distribution of responses formed a uniform distribution; that is, approximately equal numbers of teachers reported that they utilized each of the bases. Table 4.32 presents the frequency distribution for the basis on which marks are assigned.

Table 4.32.--Frequency Distribution for Basis for Marks.

	Number	Percent
Growth compared to self	85	11.9
Subject matter mastery	81	11.3
Performance compared to others	19	2.6
Growth and mastery	327	45.6
Growth and performance	70	9.8
Mastery and performance	111	15.5
None of these	24	3.3
TOTAL	717	

The obtained Chi square value of 25.33 with 6 degrees of freedom at the .05 level of significance indicated that the different bases for marks were not selected with equal likelihood.

Apparently, there was a statistically significant level of agreement among the teachers in regard to the basis on which marks are assigned.

Summary Tables

Question 11: What types of objective measurements of student performance are made by teachers?

Table 4.33 presents the information from these items dealing with the formats used for test and quiz items.

As indicated more than one-third (39.3 percent) of the teachers reported that they used short answer formatted items for quiz items, while slightly less than one-third (32.4 percent) of the teachers reported that they used this format for test items. The second most frequently used type of items for both tests and quizzes was reported as multiple choice with 17.7 percent of the teachers using this form for quiz items and 28.6 percent for test items.

Between 15 and 20 percent of the teachers reported problems as the third most frequently used, and fill-in items as the fourth most frequently used form for both quiz and test items.

While essay type items was the least frequently used form of quiz items, reported by 7.1 percent, this format was

used by 13.4 percent of the teachers as test items. The use of true-false items for quizzes and tests was reported by 8.8 and 7.6 percent of the teachers respectively.

The percent of teachers reporting that they did not give quizzes was equal to that who reported that they did not give tests.

TABLE 4.33.--Formats of Items Used for Quizzes and Tests.

	Number	Percent*
Quiz Items		
Oral	96	13.4
Problems	111	15.5
Essay	51	7.1
Short Answer	282	39.3
Fill In	109	15.2
Multiple Choice	127	17.7
True False	63	8.8
Do Not Give	90	12.6
Test Items		
Oral	44	6.1
Problems	134	18.7
Essay	96	13.4
Short Answer	232	32.4
Fill In	128	17.9
Multiple Choice	205	28.6
True False	55	7.6
Do Not Give	91	12.6

*Total percents exceed 100 due to multiple responses.

Question 12: How frequently are objective measurements of student performance made by teachers?

As indicated in Table 4.34, 16.3 percent of the teachers reported that they assigned homework daily, while 23.3 percent reported that they never assigned homework.

Approximately 30.0 percent reported that they assigned homework once or twice weekly with approximately the same proportion reporting that while they did assign homework it was less often.

While 29.7 percent of the teachers reported that they gave quizzes once or twice weekly, only 1.5 percent reported that this was a daily occurrence.

Quizzes and tests are not universal phenomena in schools as shown by the 14.9 percent of the teachers who reported that they did not give quizzes and the 16.0 percent who did not give tests. Quizzes are given at the end of each unit by 12.8 percent, while 39.1 percent reported the

TABLE 4.34.--Frequency of Objective Measurements.

	Number	Per cent
Homework		
Daily	117	16.3
Once or twice weekly	211	29.4
Less often	220	30.7
Never	167	23.3
Quizzes		
Daily	11	1.5
Once or twice weekly	213	29.7
Less often	291	40.6
End of unit	92	12.8
Never	107	14.9
Tests		
Once or twice weekly	36	5.0
Every two or three weeks	158	22.0
End of unit	280	39.1
Once or twice in marking period	122	17.1
Never	115	16.0

use of tests for this purpose. While 22.0 percent gave a test every two or three weeks, 17.1 percent reported that they gave only one or two tests per marking period.

Question 13: What changes do teachers suggest regarding marking procedures and policies?

Findings: Table 4.35 presents information pertaining to the number and proportion of respondents who suggested each category of proposed changes. Because only 72.3 percent of the teachers who returned the questionnaire responded to the item pertaining to suggested changes, the proportions were computed on the basis of those teachers responding to the item.

The largest proportion, 19.8 percent, of the responding teachers indicated satisfaction with present marking procedures. The next most frequently selected change over the traditional report card was the use of the parent-teacher conferences. An equal number of teachers proposed the use of a pass/fail procedure as suggested the complete elimination of marks. Apparently, many teachers would like to provide more information than is possible on the traditional report card as indicated by their suggestions of a check list of objectives or of narrative reports.

More teachers indicated that increased emphasis should be placed on academic characteristics than on the affective or behavioral aspects as shown by the proportions of 6.0 and 4.9 respectively.

A number of teachers, 6.4 percent, indicated the need for a system wide or school wide marking procedure.

TABLE 4.35.--Suggested Changes in Marking Procedures and Policies.

	Number	Percent
Eliminate	65	11.9
Use Narrative	48	8.8
Use Pass/Fail	65	11.9
Use Checklist of Objectives	51	9.3
Use Parent Conferences	78	14.2
More Uniformity	35	6.4
More Emphasis on Affective	27	4.9
More Emphasis on Academic	33	6.0
Other Change	38	6.9
No Change	108	19.8
TOTAL	548	

Question 14: How do school districts differ in their marking policies?

Table 4.36 presents the summarized information obtained from the report forms. Because not all schools returned the requested forms and because many elementary schools used different report forms for different grade levels, the numbers presented represent a total of 278 received report forms, although all responding teachers were included in the other analyses.

Two thirds of the forms made provision for a separate mark to indicate conduct. For the most part this included the 60.4 percent which makes use of a dual system. wherein separate marks were provided for academic and affective characteristics.

Over fifty percent (52.4) of the forms utilized some type of check list to indicate the students' mastery of particular objectives.

Only 37.1 percent of the forms required that the parents' signatures be obtained prior to returning the form to the teacher or school.

TABLE 4.36.--Frequency of Characteristics Regarding the Report Forms.

	Number	Per cent
Dual Report System	166	60.4
Conduct Mark	184	66.9
Check List	144	52.4
Narrative	7	2.5
Return to School	102	37.1
Times per Year		
2 times	60	21.8
3 times	16	5.8
4 times	138	50.2
6 times	58	21.1

Approximately one half of the school districts provided marks to parents four times during the school year. Approximately one half of the remaining districts provide only two marking periods per year. These were frequently supplemented by parent-teacher conferences. The remaining 21.1 percent of the districts provided six marking periods per year.

Summary of Findings

The findings of the study were summarized into the following thirty-one statements.

1. The combined personal characteristics of the teacher appear to explain approximately one-fifth of the variation in the relative importance allotted student characteristics.

2. The sex of the teacher appears to be related closely to the importance allotted to neatness of work and the students' personal appearance.

3. The number of measurement courses taken by the teacher appears to contribute little to the importance allotted to student characteristics.

4. The combined situational factors appear to explain approximately one third of the variation in the relative importance allotted student characteristics.

5. The grade level of the students being marked appears to be positively related to the relative importance allotted to test marks.

6. The percent of minority students in the school district or the organization level of the community appears to contribute little to the relative importance allotted to student characteristics.

7. The subject area taught appears to be only slightly related to the relative importance assigned to student characteristics.

8. The combined teachers' attitudes regarding the goals of education and the functions of marks appear to explain almost completely the variations in the relative importance allotted to student characteristics.

9. The teachers' viewpoints regarding the appropriate goals of education and the degree to which marks serve a useful function for teachers appear to be closely related.

10. The relative importance assigned to the students' personal appearance appears to be related to the degree to which teachers believe marks are useful to teachers.

11. The degree to which teachers reported that marks serve a useful function to either parents or students appears to be unrelated to the relative importance allotted to student characteristics.

12. Teachers' attitudes appear to be unrelated to the importance allotted to test and homework marks.

13. The combined other aspects of the marking procedure appear to explain approximately one third of the

variations in the relative importance allotted to student characteristics.

14. The frequency with which teachers utilize objective measurements of student performance appears to be negatively related to the relative importance assigned to the students' effort.

15. The frequency with which teachers utilize objective measurements of student performance appears to be positively related to the relative importance assigned to test marks.

16. The majority of teachers appear to base marks on a combination of growth and subject matter mastery.

17. The likelihood that teachers' select either growth or a combination of subject matter mastery and relative performance as the basis for marks appears to be affected by the sex and age of the teacher.

18. Teachers of the lower elementary grades appear to be more likely to base marks on growth while secondary teachers are more likely to use relative performance either singly or in combination with subject matter mastery.

19. Teachers of academic subject areas appear to be more likely to base marks either on subject matter mastery or on a combination of subject matter mastery and relative performance than are teachers of nonacademic subject areas.

20. Teachers, reporting a belief in progressive goals for education, appear to be less likely to base marks on either subject matter mastery or relative performance.

21. Teachers, who report that marks are useful to students, appear to base marks on relative performance.

22. Teachers, who report that marks are useful to parents, appear to be more likely to base marks on a combination of subject matter mastery and relative performance.

23. Teachers, who base marks on growth, appear not to utilize objective measurements of student performance.

24. Teachers who base marks on a combination of subject matter mastery and relative performance are more likely to utilize the statistical technique of either percentil ranks or the normal curve.

25. Teachers most often use short answer or multiple choice items for both quizzes and tests.

26. Teachers more often use oral items for quizzes than essay items, while the reverse is true for tests.

27. The majority of teachers appear to favor frequent use of objective measurements.

28. Many teachers appear to be satisfied with the presently used marking procedures and policies.

29. Dual marking systems are employed by the majority of school districts.

30. Many school districts make use of a check list of objectives rather than single marks for reporting student progress.

31. The majority of school districts report student progress four times during the school year.

Overview

This chapter has presented a description of the data analyses and a summary of the findings. Chapter V presents the conclusion and recommendations that were derived from the study.

CHAPTER V

RATIONALE FOR THE STUDY, METHODOLOGY, CONCLUSIONS, AND RECOMMENDATIONS FOR FURTHER STUDY

The intent of this chapter is to provide a summary of the research study. The rationale for the study will be defined through a description of the background of the problem. A summary of the methodology utilized for the study of the topic of interest will be presented. The next section will deal with the conclusions that were derived from the statistical analyses performed. In conclusion recommendations will be presented regarding those aspects of the topic which require further examination before a complete understanding of the status quo is obtained.

Rationale for the Study

Although the assigning of marks to students, as indicators of their in-school performance, has become a universal practice by teachers at all grade levels from kindergarten through graduate school, dissatisfaction has continued to be expressed regarding the reliability of marks and their effect on learning.

Evidence presented during the second decade of this century reported that marks assigned by different teachers, for identical work are not comparable. Two major explanations for these differences have been offered.

Teachers may differ in regard to the basis on which marks are assigned. Surveys have indicated that teachers choose to utilize either singly or in combination the following bases for marks: (1) performance relative to that of others, (2) mastery of subject matter, or (3) the students' self growth. However, for the most part, a single "best" choice has not been unquestionably defined.

Another point of difference among teachers in regard to the assignment of marks concerns those characteristics of the students that are included in the specifications of marks. While it has been claimed that marks should represent only the academic performance of the students, studies have indicated that personality characteristics are also often considered when assigning marks. It was for this reason that dual marking systems have frequently been adopted. However, this has not usually been successful in removing the subjective elements, representing the students' personality, from the objective elements, representing the students' academic performance.

Until the goals of education have been agreed upon, agreement will not be reached regarding the question of

those student characteristics that should be represented by marks. If the traditional viewpoint, that the intent of the school is the accumulation of organized useful knowledge, is accepted then only objective elements are relevant. On the other hand, humanistic advocates believe that subjective elements are most important to the overall growth of the child.

Other complaints concern the effect of marking practice on the students who are being marked. Concerns indicate that the self concept of those students, who receive poor marks may be damaged as the child begins to feel inferior. On the other hand, the advocates of marking claim that marks enable the child to establish a sense of relative worth that will be useful as an adult.

This effect of marks may be the result not of the system itself, but of the attitudes of the teachers, parents, and peer groups, as well as of the unreliability or unfairness of the marks.

Not all teachers or schools utilize traditional marking procedures as many modifications have been offered during the past decades. However, evidence is not available that by changing the external format of the reporting system the deficiencies associated with marking can be overcome.

A knowledge of the status quo in regard to the assigning of marks to students by teachers was seen as a necessary first step in the process of improving the system.

Methodology

Therefore, the intent of this study was to investigate, by means of a survey, the question of how teachers are currently assigning marks to students. It was believed that significant relationships could be established between specific personal characteristics of the teacher, the teachers' attitudes, situational factors, other aspects of the marking procedure and the basis on which marks are assigned as well as the student characteristics that are considered when marks are assigned.

The basis or framework upon which marks are assigned was defined as follows: (1) the students' growth when compared to self, (2) the students' mastery of the subject matter, (3) the students' performance relative to others, (4) combinations of these or (5) none of the mentioned possibilities. The other measure concerned the relative importance attached to each of the following student characteristics, considered when assigning marks: (1) attendance, (2) personal appearance, (3) behavior or conduct (4) amount of effort expended by the students, (5) in class participation, (6) students' attitudes, (7) neatness of work, (8) homework, (9) quiz marks, (10) test marks, (11) optional work, (12) group reports or projects, and (13) individual reports or projects.

Objectives

The primary objective of the study was to answer the following question as it pertained to the State of Michigan: How are teachers assigning marks to students?

In order to answer this question, responses to the following set of questions were first obtained.

What personal factors are significantly related to the weighting applied to student characteristics that are considered when assigning marks?

What personal factors are significantly related to the basis on which marks are assigned?

What situational factors are significantly related to the weighting applied to student characteristics considered when assigning marks?

What situational factors are significantly related to the basis on which marks are assigned?

What attitudinal factors are significantly related to the weighting applied to student characteristics considered when assigning marks?

What attitudinal factors are significantly related to the basis on which marks are assigned?

What other aspects of the marking procedure used are significantly related to the weighting applied to student characteristics considered when assigning marks?

What other aspects of the marking procedure used are significantly related to the basis on which marks are assigned?

Do teachers agree in regard to the relative importance of student characteristics considered when assigning marks?

Do teachers agree in regard to the basis on which marks are assigned?

What types of objective measures of student performance are made by teachers?

How frequently are objective measurement of student performance made by teachers?

What changes do teachers suggest regarding marking procedures and policies?

How do school districts differ in their marking policies?

Sample

In order to accomplish the objectives, a sample of 140 school districts in the State of Michigan were randomly selected according to a stratified sampling plan, utilizing proportional allocation. From each of these districts, two elementary schools, two junior high or middle schools, and two senior high schools, when available, were randomly selected. From each sampled school two teachers were randomly selected. A total sample of 1022 teachers was thereby selected to receive copies of the mailed questionnaires.

Because significant differences were not found in the responses of the teachers between the strata in regard to the variables of interest, the statistical analyses included the total sample of responses rather than individual stratum. A response rate of 70.15 percent or 717 responses were obtained, including both those obtained from the regular mailing and those obtained as a result of an intensified effort. Because a significant difference was not obtained according to the mailing that resulted in response it was

assumed that the sample of responders was unbiased and, therefore, was representative of the population of teachers in Michigan.

Data Collection

The primary source of information was a mailed questionnaire which was designed specifically for the study. Approximately 15 minutes of the teachers' time was required to respond. Twenty two items were included, two of which asked for an open-ended response, while the others provided from two to six options.

Information was available from these questionnaires regarding the teachers' personal characteristics, attitudes, situational factors, and the marking procedures utilized.

Additional information concerning situational factors was obtained from the Michigan Education Directory. Copies of the report forms used within the sampled schools provided information regarding the marking policies of the school and district.

Data Analysis

Canonical correlation analyses were employed to determine the maximum correlations which could be obtained from the set of weighted values, representing the relative importance associated with each of the student characteristics, considered when assigning marks and each other set

of weighted values, representing personal characteristics, situational factors, attitudinal factors and other aspects of the marking procedure.

Chi square tests of independence were utilized to test for the existence of a relationship between the bases on which marks are assigned and each element within the sets of personal characteristics, situational factors, attitudinal factors, and other aspects of the marking procedure. The associated cell Chi square values provided an indication of specific relationships.

A Kendall's coefficient of concordance was calculated in order to measure the level of agreement among the teachers in regard to the relative importance associated with each student characteristic considered.

In order to determine whether there was agreement among the teachers in regard to the bases on which marks are assigned, a Chi square one sample test was performed.

Limitations

The primary factor that limited the derivation of conclusions from this study was the small number of teachers sampled from any single grade level or subject area. This precluded the interpretation of results at any particular level. Other studies have found significant differences between grade levels and subject areas. Therefore, further

study is needed, utilizing narowly defined groups of teachers.

Another limitation pertained to the instrument that was utilized. The item which dealt with the relative importance associated with the student characteristics considered may have been misinterpreted. Because many school policies provide for multi-bases or at certain grade levels only for parent teacher conferences, misunderstandings may have occurred on the part of the teachers as to how to respond.

Conclusions

Several conclusions were formulated from the results of the statistical analyses performed.

The evidence obtained from this study indicates that teachers belong to either one of two categories in regard to the procedures which they utilize when assigning marks to students. The first group includes those teachers who rely primarily upon their subjective estimations concerning the students' performance. The second group includes those teachers who primarily depend upon objective measurements in relation to marking.

The former group demonstrates a definite tendency to report that the marks which they assign are based upon the students' growth when compared to himself. In the same connection, the latter group are more likely to report that

they assign marks either solely upon the students' relative performance when compared to others or on a combination of relative performance and subject matter mastery.

Marks are viewed, by those teachers who utilize primarily objective measurements, as serving a useful function for both parents and students. On the other hand, those teachers who determine marks according to their own subjective estimation of the students' growth do not believe that marks serve a useful function particularly for teachers. This latter finding may be also due to this group's opinions regarding the appropriate goals of education.

The progressive viewpoint regarding the goals of education which this group holds claims that education should seek primarily to socialize the child rather than to instill knowledge. An important element within this framework claims that marks are dehumanizing and are harmful to the child. This group claims that objective measurements do not reveal the students' performance relative to the true goals of education, to socialize the student. To the contrary, test marks for the other group of teachers are said to provide a measure of the degree to which the student has achieved the intent of the school, to provide a sound framework and substantive knowledge. Therefore, the amount of effort expended by the student is less relevant than is level of subject matter mastery. Although it would appear

that the marks, assigned by those teachers who consider test and quiz marks to be of greatest importance, would be highly reliable. However, because the majority of teachers reported a preference for short answer items and because of the frequent ambiguity associated with these items, it can be assumed that even these marks lack reliability.

Percentiles and percentile ranks are useful indicators to this group of teachers of the level of mastery, while a normal curve can serve as an indicator of the students' relative performance. Such techniques, however, do not provide a useful means for the teacher who uses subjective estimation procedures to indicate performance.

An overall profile of the two groups would demonstrate that male teachers are more likely to utilize objective measurements. Within the same group are more likely to be the younger teachers, who teach academic subject areas to students at the upper grade levels. Those teachers, who prefer subjective estimation methods, are more likely to be older females who teach either lower grade levels or secondary nonacademic subject areas.

Differences were not observed between the two groups according to the following measures which had been expected to affect marking procedures: (1) teacher experience, (2) college degree held, (3) number of measurement courses taken, (4) teacher's high school grade point average, (5) community urbanization level, (6) percent of minority students, (7)

size of the school or school district, (8) the grade levels included within the school, or (9) the specific subject area taught.

For the most part, teachers are satisfied with existing marking procedures, although many would like to provide more information than can be done by traditional methods. This is particularly true for those teachers who utilize subjective estimation methods. Conferences between the teacher and parents are, therefore, preferred by this group.

The primary concern based upon the finding of this study is whether there should be two distinct categories of teachers in regard to how marks are assigned. Because categorical identifications were not clearly defined even according to grade level or subject area taught, the misinterpretation of marks seems to be very likely.

In fact, evidence indicates that teachers themselves are unable to define precisely those steps taken to determine the specific mark or marks to assign a particular student. The usual response to the question concerning the precise steps taken was as follows: (1) average grades assigned during period, (2) consider effort, attitude, etc., and (3) adjust mark accordingly.

Parents, students, teachers, and all other individuals who receive, assign, or in any way have contact with marks, which are intended as indicators of student

performance, should be aware that differences in meaning do exist between marks. The teachers particularly must recognize and acknowledge, in a way that is known to all others, the procedures which they individually use in order to decrease confusion and misunderstandings. Teachers, also, should strive to develop and utilize the most reliable methods possible, and those best suited to their own unique situation.

Recommendations for Further Study

This study was unable to provide information pertaining to several aspects of the questions concerning marking practices. If a complete understanding of the status quo is to be gained, further studies should be conducted in order to derive answers to the following questions.

How do teachers of different grade levels differ in regard to the marking procedures utilized?

How do teachers at the elementary level differ in connection with the assignment of marks representing the varied subject areas?

How do secondary teachers of different subject areas differ in regard to the marking procedures utilized?

What are the expectations of parents and students in regard to marks and their associated interpretations?

What is the actual role of the school administration in regard to marking?

How does the amount of time required by teachers differ in regard to the varied marking procedures used?

What are the characteristics of the tests, quizzes, and homework utilized by teachers?

What techniques do teachers use to appraise the affective characteristics of students?

How effective are the various technical changes proposed as means of improving marking practices?

What is the content of college courses which pertain to measurement?

Does the content of college measurement courses serve teachers' needs?

How should in-service training be structured in order to meet the needs of teachers?

Although the present study has provided some important information regarding the status quo of marking in the State of Michigan, a great deal of time, effort, and funding should be devoted to this study in the future. The marks a student received while attending school determine to a large extent the entire direction of his future life. Therefore, this aspect as a part of the entire educational process should be considered carefully both by future researchers and teachers, when assigning the marks that may control the students' entire future.

APPENDICES

APPENDIX A

SURVEY QUESTIONNAIRE

Please respond to the following items by circling the correct response. If none of the options are exactly correct, you are to choose the one that is closest. Please leave no items blank. Limit your responses to one response per item.

1. Sex
   1) male
   2) female

2. Age at your nearest birthday
   1) 20-25
   2) 26-35
   3) 36-45
   4) 46-55
   5) 56-65
   6) over 65

3. Number of years of teaching experience, include the current year as 1 year.
   1) 1
   2) 2-5
   3) 6-10
   4) 11-20
   5) 21-40
   6) over 40

4. What is the highest college degree that you hold?
   1) none
   2) B.A. or B.S.
   3) M.A. or M.S.
   4) Ed. Spec.
   5) Ed.D. or Ph.D.

5. How many college courses have you taken in the area of grading or measurement?
   1) none
   2) 1
   3) 2 or 3
   4) 4 or more

6. What kind of grades did you receive in high school?
   1) high
   2) average
   3) low

7. In which one of the following grade categories do you teach mostly this year?
   1) K-1-2
   2) 3-4
   3) 5-6
   4) 7-8-9
   5) 10-11-12
   6) nongraded elementary
   7) special education

8. In which one of the following subject areas do you teach mostly this year?
   1) science
   2) mathematics
   3) social science
   4) language arts including reading
   5) foreign language
   6) business and/or commercial
   7) music or art
   8) home economics or industrial arts
   9) health and/or physical education
   10) elementary, all subject areas

9. How often do you assign homework?
   1) never
   2) daily
   3) once or twice weekly
   4) less often than twice weekly

10. How often do you give quizzes? (requiring 15 or less minutes)
   1) never
   2) daily
   3) once or twice weekly
   4) less often than twice weekly
   5) only at the end of each unit

11. How often do you give longer tests?
   1) never
   2) once or twice weekly
   3) once every two or three weeks
   4) at the end of each unit
   5) once or twice each marking period

12. What format do you usually use for quiz items?
   1) oral
   2) problems
   3) essay
   4) short answer
   5) fill in
   6) multiple choice
   7) true false
   8) do not give quizzes

13. What format do you usually use for test items?
   1) oral
   2) problems
   3) essay
   4) short answer
   5) fill in
   6) multiple choice
   7) true false
   8) do not give tests

14. On which of the following bases do you assign marks?
   1) the student's growth compared to himself
   2) the student's mastery of subject matter
   3) the student's performance compared to others
   4) 1 and 2 above
   5) 1 and 3 above
   6) 2 and 3 above
   7) none of the above

15. Which of the following sources of information do you find most helpful when assigning marks?
   1) parents 2) the student himself 3) other teachers 4) student past records 5) none of these

16. Teachers sometimes use certain techniques in their efforts to measure student performance. One such technique is the use of standard scores. Do you use Standard Scores when marking?
   1) yes, always 2) yes, sometimes 3) no, they take too much time 4) no, other reasons 5) do not know

17.. Another such technique is percentile ranks or percentiles. Do you use Percentile Ranks or Percentiles when assigning marks?
1) yes, always 2) yes, sometimes 3) no, they take too much time 4) no, other reasons
5) do not know how

18. Another technique used is the normal distribution or normal curve. Do you use a Normal Distribtuion or Normal Curve when assigning marks?
1) yes, always 2) yes, sometimes 3) no, they take too much time 4) no, other reasons
5) do not know how

19. Please indicate on a scale from 0, none, to 5, high, how much emphasis you think school should give to each of the following.

1) to provide the opportunity for students to learn the value of academic knowledge	0	1	2	3	4	5
2) to teach useful skills	0	1	2	3	4	5
3) to teach social awareness	0	1	2	3	4	5
4) to provide students with academic competency	0	1	2	3	4	5
5) to help students become socialized individuals	0	1	2	3	4	5
6) to assist students in their preparation for a future vocation	0	1	2	3	4	5
7) to assist students in the development of a positive self concept	0	1	2	3	4	5
8) to provide the opportunity for students to become aware of societal problems	0	1	2	3	4	5

20. Please indicate on a scale from 0, not at all successful, to 5, very successful, how successful you think marks are at present for each of the following pruposes.

1) to indicate to the teacher the degree to which specific objectives are mastered	0	1	2	3	4	5
2) to punish student for misbehaving	0	1	2	3	4	5
3) to inform parents of the student's relative standing in the class	0	1	2	3	4	5
4) to guide the student in planning for college or a career	0	1	2	3	4	5
5) to indicate to the teacher the academic strengths of the students	0	1	2	3	4	5
6) to inform parents of the student's academic progress	0	1	2	3	4	5
7) to direct the learning of the students	0	1	2	3	4	5
8) to indicate to the parents the social needs of the student	0	1	2	3	4	5
9) to indicate to the teacher whether methods and materials were successful	0	1	2	3	4	5

21. Below are listed several characteristics of students that teachers sometimes consider when assigning marks. Please indicate which of the following you consider by giving the percent of the mark that is dependent upon each characteristic.
For example, if half of a mark is dependent upon test marks, etc., you might indicate as follows:

attendance	10 %
class participation	____ %
test marks	50 %
group reports or projects	25 %

---

1) attendance	____%	8) homework marks	____%
2) personal appearance	____%	9) quiz marks	____%
3) behavior or conduct	____%	10) test marks	____%
4) effort put forth	____%	11) optional work	____%
5) class participation	____%	12) group reports or projects	____%
6) attitudes	____%	13) individual reports or projects	____%
7) neatness of work	____%		

Please respond to the next item either on the back of these pages or on your own paper.

22. What changes would you suggest in regard to the current marking procedures you use or in the system set by the school administration? What effect might these changes have on education?

THANK YOU AGAIN FOR YOUR HELP. BE SURE TO LET US KNOW IF YOU WANT A COPY OF THE FINDINGS.

No. 1761

APPENDIX B

TABLE Appendix B.--Frequency Distribution for Responses to Questionnaire Items.

	I	II	III	IV	V	Total	Total Percent
Sex							
Male	48	59	68	76	31	282	39
Female	56	120	99	105	55	435	61
Age							
20-25	20	34	37	28	7	126	18
26-35	35	58	67	70	35	265	37
36-45	24	39	32	41	20	156	22
46-55	13	23	23	25	15	99	14
56-65	2	25	8	17	9	71	10
Experience							
1	10	17	16	10	3	46	6
2- 5	29	60	53	51	21	214	30
6-10	17	32	44	47	25	165	23
11-20	40	39	37	50	25	191	27
20-40	8	31	17	23	12	11	13
College Degree							
None	0	0	0	0	0	0	
B.A.-B.S.	77	133	112	103	41	466	65
M.A.-M.S.	26	45	54	73	37	235	33
Ed. Spec.	0	0	0	4	6	10	1
Ed. D.-Ph.D.	1	0	1	1	2	5	1
Measurement Courses							
None	35	61	62	48	25	231	32
1	34	52	55	72	27	240	33
2-3	31	63	43	57	29	223	31
4 or more	4	3	7	4	5	23	3
Self Grades							
High	63	105	104	114	53	439	61
Average	40	70	58	63	30	261	36
Low	1	4	5	4	3	17	3
Grade-Teach							
K-1-2	19	39	35	24	10	127	17
3-4	19	36	29	23	8	115	16
5-6	12	23	20	35	12	102	14
7-8-9	21	36	41	55	23	176	25
10-11-12	32	42	42	41	30	187	26
nongraded	1	3	1	3	3	11	2
Subject teach							
Science	10	16	13	14	5	58	8
Mathematics	10	17	11	22	9	69	10
Social Science	14	21	19	17	13	84	12
Language Arts	19	25	25	30	23	122	17
Foreign Language	2	5	0	6	1	14	2

Table Appendix B Continued.

	I	II	III	IV	V	Total	Total Percent
Business	1	4	5	6	5	21	3
Music-Art	2	4	7	6	0	19	2
Home Ec-Industrial Arts	6	3	6	6	4	25	3
Health-P.E.	2	4	6	6	3	21	3
All Elementary	38	80	75	68	23	284	40
Frequency Homework							
Never	18	45	49	43	14	169	24
Daily	21	32	20	24	20	117	16
Once or Twice Weekly	41	45	40	58	27	211	29
Less Often	24	57	58	56	25	220	31
Frequency Quizzes							
Never	9	27	28	34	12	110	15
Daily	1	3	2	2	3	11	2
Once or Twice Weekly	45	52	44	49	23	213	30
Less Often	37	73	72	72	37	291	41
Unit End	12	24	21	24	11	92	13
Frequency Tests							
Never	14	29	34	33	11	121	17
Once or Twice Weekly	9	10	4	6	7	36	5
Two or Three Weeks	22	39	37	44	16	158	22
Unit End	45	76	56	72	31	280	39
Once or Twice Marking Period	14	25	36	26	21	122	17
Format Quiz Items							
Oral	16	22	21	28	9	96	13
Problems	14	25	23	37	12	111	15
Essay	9	4	8	16	12	49	7
Short Answer	53	65	71	65	30	284	40
Fill In	17	30	26	25	10	108	15
Multiple Choice	18	31	25	32	22	128	18
True-False	7	16	16	15	9	63	9
None	8	28	28	27	12	103	14
Format Test Items							
Oral	4	8	15	12	5	44	6
Problems	14	34	26	43	17	134	19
Essay	16	13	22	29	16	96	13
Short Answer	41	58	49	57	27	232	32
Fill In	20	36	27	33	13	129	18
Multiple Choice	31	55	40	51	27	204	28
True-False	12	21	15	20	8	76	11
None	8	22	23	28	9	90	13
Useful Information							
Parents	3	4	2	5	2	16	2
Student	63	137	141	144	79	564	79
Teachers	1	0	1	1	0	3	1
Records	2	9	3	3	2	19	3
None	17	29	20	28	21	75	10

Table Appendix B Continued.

	I	II	III	IV	V	Total	Total Percent
Use Standard Score							
	4	9	6	10	8	37	5
	22	56	30	51	30	189	26
	1	1	2	4	5	13	2
	53	99	115	99	56	422	59
	6	14	14	17	5	56	8
Use Percentiles							
	3	15	12	8	9	47	7
	24	53	46	51	32	206	29
	4	4	2	4	5	19	3
	54	99	102	112	55	422	59
	1	8	5	6	3	23	3
Use Normal Curve							
	4	7	4	8	2	25	3
	29	65	43	66	37	240	33
	1	3	2	3	2	11	2
	50	100	115	102	59	426	59
	2	4	3	2	4	15	2
Goals of Education							
1. To provide the opportunity for students to learn the value of academic knowledge							
0	3	4	6	4	5	22	3
1	1	4	3	4	1	13	2
2	7	11	11	13	7	49	7
3	26	43	42	38	17	166	22
4	26	48	40	42	26	182	29
5	41	69	65	80	30	285	40
2. To teach useful skills							
0	2	4	2	1	6	15	2
1	5	1	0	1	0	0	7
2	0	3	5	4	3	15	2
3	7	12	14	8	7	48	7
4	21	31	34	30	13	129	18
5	69	128	112	137	57	503	70
3. To teach social awareness							
0	3	4	3	6	3	19	3
1	3	1	3	0	0	7	1
2	4	6	3	4	3	20	3
3	23	30	24	22	14	113	16
4	27	51	52	49	21	200	28
5	44	87	82	100	45	358	50
4. To provide students with academic competency							
0	3	4	4	4	5	20	3
1	2	1	1	5	2	11	2

Table Appendix B Continued.

	I	II	III	IV	V	Total	Total Percent
2	6	3	5	3	3	20	3
3	20	52	39	37	19	167	23
4	32	63	59	65	27	246	34
5	41	56	59	67	30	253	39
5. To help students become socialized individuals							
0	6	5	3	5	4	23	3
1	4	2	1	4	0	11	2
2	5	12	8	11	4	40	8
3	18	40	36	19	18	131	18
4	29	48	54	57	24	212	30
5	42	72	65	85	36	300	42
6. To assist students in their preparation for a future vocation							
0	3	3	4	3	4	17	2
1	2	2	1	2	2	9	1
2	3	7	10	6	2	28	4
3	15	22	18	29	10	93	13
4	21	41	42	46	17	167	23
5	60	104	92	95	51	402	56
7. To assist students in the development of a positive self concept							
0	4	3	9	3	3	13	2
1	2	2	1	0	1	6	1
2	3	1	0	3	1	8	1
3	12	12	11	11	5	51	7
4	15	34	35	31	19	134	19
5	68	127	120	133	57	505	70
8. To provide the opportunity for students to become aware of societal problems							
0	4	3	4	6	5	22	3
1	3	2	2	2	0	9	1
2	6	12	9	6	8	41	6
3	22	40	37	49	13	167	22
4	31	63	65	51	25	235	33
5	38	59	50	67	35	249	35
Functions of Marks for Parents							
8. To indicate to the parents the social needs of the student							
0	39	58	44	62	29	232	32
1	7	26	43	38	18	132	18

Table Appendix B Continued.

	I	II	III	IV	V	Total	Total Percent
2	18	36	32	29	11	126	18
3	26	36	38	38	19	157	22
4	10	17	10	10	7	54	8
5	4	6	0	4	2	16	2
6. To inform parents of the student's academic progress							
0	5	9	9	10	8	41	6
1	5	3	5	6	4	23	3
2	13	23	17	19	12	84	12
3	28	64	59	63	23	237	33
4	35	50	57	63	24	229	32
5	18	30	20	20	15	103	14
3. To inform parents of the student's relative standing in the class							
0	6	13	16	21	12	68	9
1	9	13	16	10	8	56	8
2	19	30	15	20	12	96	13
3	37	68	57	58	26	246	34
4	20	39	47	51	21	178	25
5	13	16	16	21	7	73	10
Functions of Marks for Students							
4. To guide the student in planning for college or a career							
0	12	19	17	20	16	84	12
1	5	8	14	5	9	41	6
2	13	28	30	35	10	116	16
3	34	65	61	61	24	245	34
4	31	43	31	43	23	171	24
5	9	16	14	17	4	60	8
7. To direct the learning of the students							
0	14	25	27	27	20	113	16
1	10	15	18	21	7	71	10
2	19	32	30	33	13	127	18
3	30	69	54	54	29	236	33
4	24	28	28	32	10	122	17
5	7	10	10	14	7	48	7
2. To punish student for misbehaving							
0	62	110	106	102	52	432	60
1	15	28	26	35	17	121	17
2	9	22	17	25	9	82	11

Table Appendix B Continued.

		I	II	III	IV	V	Total	Total Percent
	3	12	11	10	9	5	47	7
	4	4	4	6	6	3	23	3
	5	2	4	2	4	0	12	2
For Teachers								
1.	To indicate to the teacher the degree to which specific objectives are mastered							
	0	7	12	11	13	11	54	8
	1	9	8	10	8	5	40	6
	2	10	20	16	21	10	77	11
	3	42	63	65	64	24	258	36
	4	23	52	46	53	27	201	28
	5	13	24	19	22	9	87	12
5.	To indicate to the teacher the academic strengths of the students							
	0	6	9	12	12	8	47	7
	1	5	7	5	9	7	33	5
	2	14	22	18	22	10	86	12
	3	33	51	62	62	29	237	33
	4	37	67	57	60	22	243	34
	5	9	23	13	16	10	71	10
9.	To indicate to the teacher whether methods and materials were successful							
	0	13	21	14	22	10	80	11
	1	9	13	11	17	6	56	8
	2	6	25	22	21	10	84	12
	3	38	48	52	50	20	208	24
	4	24	50	42	46	29	191	27
	5	14	22	26	25	11	98	14

APPENDIX C

Computer Institute for Social Science Research
Michigan State University
501 Computer Center
East Lansing, Michigan 48823

Dear

During recent years education has been the recipient of much criticism, a part of which has centered around the grading or marking of students. Claims have been made that marks are harmful to students and that they are not accurate indicators of the students' abilities. However, little is actually known about the processes by which teachers assign marks or the relationship between personal, situational, and attitudinal factors and the marking procedures used by teachers.

We have designed a survey through which we hope to derive some answers to the question of how teachers in Michigan are marking students. Your school district has been selected to participate in this study. Your cooperation will be greatly appreciated.

You are being asked to support this study in two ways. First by supplying the names and addresses of the teachers assigned on a regular basis to each of the following schools within your district.

Two teachers from each of these schools will be selected to receive mailed questionnaires pertaining to the procedures by which they assign marks. This questionnaire will require approximately 15 minutes for its completion. No further contact will be made to any teacher, unless a request is made by a teacher for a copy of the findings.

Secondly you are being asked to supply us with an unused copy of the student progress report form used in each of the above mentioned schools.

Because the entire study is dependent upon the cooperation of school superintendents, your support is requested.

We would appreciate very much if the requested information could be sent immediately. Unless all information is received before April 20, the study can not be completed during the current school year.

If you would like to receive a copy of the complete findings, please so indicate, when you send the list of teachers and the progress report forms.

Respectfully,

Bonnie J. Steller

Dear Superintendent:

Recently you should have received a letter requesting your support of a study we have designed, regarding the marking procedures used by teachers in Michigan. If you have not as yet replied, your immediate attention would be appreciated. If your failure to respond is due to policies regarding the release of teachers' addresses, please send us the remaining information. Additional procedures have been designed for districts which prefer not to even release teacher names.

Please confirm your cooperation at your earliest convenience, as without complete cooperation from school superintendents, the entire study will be worthless. Thank you for your help.

BONNIE J. STELLER
Computer Institute for Social Science Research
501 Computer Center
Michigan State University
East Lansing, Michigan 48823

APPENDIX D

COMPUTER INSTITUTE FOR SOCIAL SCIENCE RESEARCH
501 Computer Center
Michigan State University
East Lansing, Michigan 48823

Dear Sir:

During recent years education has been the recipient of much criticism, a part of which has centered around the marking or grading of students. Claims have been made that marks are harmful to students and that they are not accurate indicators of students' abilities, yet little is known about how teachers actually assign marks to students.

We have designed a survey through which we hope to derive some answers to the question of how teachers in Michigan are assigning marks. Two teachers from your school have been selected to participate in this study.

Enclosed with this letter are two sets of materials, a cover letter explaining the purpose of the survey, a two-page questionnaire, and a return addressed, stamped envelope, each set being addressed by a number.

You are being asked to distribute these two sets of materials according to the following instructions. It is very important that these instructions be followed, as only in this way can we later make statements of our findings as they apply to all teachers within the state. In other words, we have applied random selection procedures.

First you are to number consecutively the names on a roster of teachers employed in your school, omitting all special education teachers, librarians, counselors, etc. The teachers, who in this way are assigned the numbers on the envelopes, are to be given the set of materials. If, by chance, an envelope is addressed to a number greater than any that you have assigned, the materials are to be given the last teacher on the roster.

Because we plan to have no direct contact with any teacher, it would be very greatly appreciated if you would be responsible for reminding the teachers on or about May 11 that the questionnaires are to be returned before May 15, as we want to complete the collection of information prior to the close of schools for summer vacation.

It would also be helpful if you would send us an unused copy of the form used for reporting student progress by your teachers.

If you are unable to participate in this study, please drop us a note, indicating your reasons for refusal to comply with our needs.

If you would like a copy of our findings, please let us know. They should be available during the Fall of 1973.

Thank you for your participation.

Respectfully,

THIRD MAILING

COMPUTER INSTITUTE FOR SOCIAL SCIENCE RESEARCH

501 Computer Center
Michigan State University
East Lansing, Michigan 48823

Dear Principal:

Recently you were requested to deliver copies of the enclosed questionnaire to two teachers in your building. As yet we have not received both completed responses. Therefore, we are asking you to again distribute copies of the questionnaire to the selected teachers from whom we have received no response.

We are aware of the fact that there were unplanned delays in the mailing and delivery of these materials; therefore, we are extending the deadline to June 7. Also we realize that the first copy of the questionnaire may have been mislaid or for some other legitimate reason been forgotten. However, as you realize it is crucial that we obtain complete responses, if the findings of the study are to be generalized to the entire population of teachers within the State of Michigan.

If you have not sent us a copy of the student progress report form used in your school, please put it in the mail at your earliest convenience.

Again thank you for your support and cooperation.

Please remind the teachers to whom you distribute the questionnaires that we must have them returned before June 7, as we realize that after that date teachers often are not available through the schools. We certainly hope that you are able to secure the cooperation of the teachers in your building.

Sincerely,

Bonnie J. Steller

Bonnie J. Steller

Enclosures

APPENDIX E

Computer Institute for Social Science Research
501 Computer Center
Michigan State University
East Lansing, Michigan 48823

Dear Teacher:

During recent years education has been the recipient of much criticism, a part of which has centered around the marking or grading of students. Claims have been made that marks are harmful to students and that they are not accurate indicators of the students' abilities. However, little is actually known about the processes used by teachers when assigning marks.

We have designed a survey through which we hope to derive some answers to the question of how teachers in Michigan are marking students. Your school district was selected to participate in this study. Your school superintendent has been contacted and has indicated his willingness for you to participate in this study. Your cooperation will be very greatly appreciated.

You are being asked to support this study by responding to the enclosed questionnaire. It should require approximately 15 minutes for completion. Please respond to those items directly pertaining to your marking methods according to what you actually do, rather than what you would like to do or think might be better. Other items ask for your own opinions rather than for facts. Only one response should be made to each item.

Because the entire study is dependent upon the cooperation of teachers such as you, your support is requested.

We would appreciate very much if you could complete the questionnaire at your earliest possible convenience and return it to us in the enclosed, addressed envelope. Those questionnaires received after May will not be included in the study.

If you would like to receive a copy of the findings of the study, please either so indicate on the questionnaire by including your name and address or send us a postcard. The quesionnaires do not need to be signed unless you want to request a copy of the findings. The findings should be available early in September.

Thank you for your time and assistance in helping us answer an important question for educators today.

Sincerely,

Bonnie J. Steller

COMPUTER INSTITUTE FOR SOCIAL SCIENCE RESEARCH

501 Computer Center
Michigan State University
East Lansing, Michigan 48823

Dear Teacher:

Recently you should have received a questionnaire dealing with the procedures by which you assign marks to students. However, as yet we have not received your completed response.

We are aware of the fact that there were delays in the mailing and delivery of this earlier letter and questionnaire. Also knowing that teachers are very busy people particularly at this time of year, and that, for this reason, it is easy to mislay and forget something such as this, we have enclosed another copy of the questionnaire with this letter. Please complete and return it before June 1 so that we can analyze the responses.

Remember that it is crucial that all of the teachers, whom we selected, do respond regardless of their opinions regarding marking. Unless we receive all of the responses, the findings of the study will tell us nothing about how teachers in Michigan are assigning marks but will only tell us how those who took the 15 minutes required to respond are assigning marks to students.

If solutions to problems confronting educations are to be found, it must be through the cooperation of all teachers. Only a few can accomplish little.

Thanks again for your support and cooperation. We appreciate your help.

Don't forget to let us know if you would like a copy of the findings.

Sincerely,

Bonnie J. Steller

Bonnie J. Steller

Enclosure

Dear Teacher,

This is just a reminder that we have not as yet received your response to the questionnaire concerning your marking procedures. June 1 is not far off so please remember to take care of it at your earliest convenience.

Thank You Again.

Bonnie J. Steller
Computer Institute for Social Science Research

BIBLIOGRAPHY

BIBLIOGRAPHY

Primary Sources

Anderson, Robert H. "The Importance and Purposes of Reporting." National Elementary School Principals, XLV (1966), 6-11.

Arnold, Richard Dean. "The Relationship of Teachers' Sex to Assignment Marks and Tested Achievement among Upper Elementary Grade Boys and Girls." (Unpublished Ph.D. dissertation, University of Minnesota, 1966).

Bates, Bean. "Can we stop the merry-go-round of grades?" Arizona Teacher, LV (January, 1967), 19 and 26.

Bicknell, George L. "Sand V Grading System." Texas Outlook, XXV (July, 1941), 18.

Bolmeier, E. C. "Principles pertaining to marking and reporting of pupil progress." School Review, LIX (January, 1951), 15-24.

Bramlette, Merle. "Is the Sand V Grading System Satisfactory or Unsatisfactory?" Texas Outlook, XXVI (April, 1942) 29-30.

Butterworth, Thomas W. "A Guide to the Evaluation of Pupil Progress with Particular Emphasis on Pupil Marks." ERIC No. 036 814 (August, 1966).

Chadwick, Ruth E., Rose Durham, and Marion Morse. "The Report Card in a Nongraded School." National Elementary Principals, XLVII (January, 1968), 22-28.

Crooks, A. D. "Marks and Marking Systems: A Digest." Journal of Educational Research, XXVII (1933), 259-273.

Delgalvis, N. "Suggesting a Different Grading System." Pennsylvania School Journal (January, 1965).

De.Zoeche, Dorothy. "The Wound is Mortal." The Clearing House XIX (February, 1945), 339-344.

Dionne, Joseph. "Future of Testing: A Look at the Trends." ERIC No. 043 669 (March, 1970).

Doerhoff, Roy. "Grading as Easy as 1, 2, 3, 4." School and Community, LV (April, 1969), 19.

Doherty, William V. "A Survey of the Evaluation of Pupil Progress in Selected Elementary Schools in Ohio." (Unpublished Ph.D. dissertation, Ohio State University, 1954).

Dunivan, Lindell P. "Grading as a system of guidance: a rebuttal." School and Community, LVI (May, 1970), 31.

Ebel, Robert L. "The relation of scale fineness to grade accuracy." Journal of Educational Measurement, VI (Winter, 1969), 217-221.

Editorial, "Report Cards on Parent-Teacher Conferences: Which is Better?" Kentucky School Journal, XIXL (November, 1970), 23.

Benjamin Fine, "A.B.C. of Grading Puzzles Parents." New York Times Magazine, November 18, 1957, cited by Ann Z. Smith and John E. Dobbin, "Marks and Marking Systems," Encyclopedia of Educational Research, 3rd ed., edited by Chester W. Harris (New York: The Macmillan Company 1960), pp. 783-791.

Finn, Jeremy D. "Expectations and the Educational Environment." Review of Educational Research, XLII (Summer, 1972), 387-410.

Fish, Enrica. "The Relationship of Teachers' Assigned Marks among Elementary Grades." (Unpublished Ph.D. dissertation, University of Minnesota, 1969).

Gage, N. L. "Desirable Behaviors of Teachers." Urban Education (November, 1965), 85-95.

Gould, George. "Practices in Marking and Examinations." School Review XL (1932), 142-146.

Halliwell, Joseph W and Joseph P. Robitaille. "The relationship between theory and practice in a dual reporting program." Journal of Educational Research, LVII (November, 1963), 137-141.

Harris, Wilbur S. "Agreement among N.C.M.E. Members on Issues in Educational Measurement." Journal of Educational Measurement, X (Spring, 1973), 63-70.

Hinely, R. T. "An Equal Chance--A Fantasy." Phi Delta Kappan, XLVIII (June, 1967), 523-524.

Housdorff, Henry and S. David Farr. "The effect of grading practices on the marks of gifted sixth grade children." Journal of Educational Research, LIX (December, 1965), 169-172.

Johnson, Icie F. "Injustice of Grades." School and Community, LIV (October, 1967), 24-25.

Johnston, Samuel R. "Ar there better ways of evaluating, recording, and reporting pupil progress in the junior and senior high schools?" National Association of Secondary School Principals, XXXIV (March, 1950), 79-89.

Keck, Winston B. "Testing and Evaluation Practices in New England Public Schools." (Unpublished Ph.D. dissertation, Boston University, 1961).

Kelley, Allan C. and Paul Zarembka. "Normalization of Student Test Scores: An Experimental Justification." Journal of Educational Research, LXII (December, 1968), 160-164.

Krause, Dorothy. "A study of present practices and attitudes concerning the letter grading system of secondary language arts teachers." (Unpublished Ph.D. dissertation, University of South Dakota, 1969).

Lawson, Douglas E. "Scoring of subjective tests with several variables controlled." Elementary School Journal, XXVII (February, 1938), 450-457.

Liggett, William A. "Are there better ways of evaluating, recording, and reporting pupil progress in the junior and senior high schools?" National Association of Secondary School Principals, XXXIV (March, 1950), 79-89.

Linder, Ivan. "Is there a substitute for teachers' grades?" American School Board Journal (July, 1940), 25-26.

Little, Ruth Coyner. "Whither Grading?" National Education Association Journal (May, 1967), 95.

Lund, John. "More Truth about Marks." Journal of Education (1929), 609-639.

Marshall, John C. "Evaluation: Does it mean anything?" The Clearing House, XLII (May, 1968), 535-538.

Mason, Geoffrey. "Studies and Reports: An Empirical Analysis of a System of Achievement Grading." ERIC No. 014 131-132 (November, 1965).

Maxson, Wilbur G. "Grading--a serious matter." National Education Association Journal, LIII (October, 1964), 56-57.

McGee, H. M. "Measurement of Authoritarianism and its Relation to Teachers' Classroom Behavior." Genetic Psychology Monographs (1955), 89-196.

Melby, Ernest O. "It's time for schools to abolish the marking system." Nations Schools (May, 1966), 104.

Metzner, Seymour. "Teacher Bias in Pupil Evaluation: A Critical Analysis." Journal of Teacher Education, XXII (Spring, 1971).

Millman, Jason. "Reporting Student Progress: A case for a criterion referenced marking system." Phi Delta Kappan, LII (December, 1970), 226-230.

Miner, Betty Crowther. "Three factors of school achievement." Journal of Educational Research, LX (April, 1967), 370-376.

Morris, Lucile, "Evaluating and Reporting Pupil Progress." Elementary School Journal, LIII (November, 1952), 144-149.

Mosely, Aubrey Howard. "A Study of Teachers' Perceptions of Factors Related to Educational Research." (Unpublished Ph.D. dissertation, University of Alabama, 1966).

Newsome, Ralph P. "Assignment of School Marks." Texas Outlook, XXV (August, 1941), 39-40.

Niedermeyer, Fred C. and Howard J. Sullivan. "Differential effects of individual and group testing strategies in an objective based instructional program." Journal of Educational Measurement, IX (Fall, 1972). 199-204.

Palmer, Orville. "Seven Classic Ways of Grading Dishonesty." The English Journal, LI (1962), 464-467.

Pewitt, Joe. "Grading as a system of coercion." School and Community, LVI (February, 1970), 20.

Philbrick, Joseph L. and Patrick I. O'Donnell. "Precision in Grading Practices--Panacea or Problem?" Journal of Educational Research, LXII (December, 1968), 173-176.

Pope, Billy Newton, "Factorial Study of Teacher Marks and Standardized Tests. (Unpublished Ph.D. dissertation, University of Houston, 1965).

Purkey, Ernest. "An Analysis of Test Data and Teachers' Marks over Twelve Years of School Attendance." (Unpublished Ph.D. dissertation, University of Missouri, Columbia, 1966).

Ribon, David M. "A Grade-Weighting System Investigated." National Association of Secondary School Principals, L (November, 1966), 94-98.

Roberts, Bernice R. "A Comparison of Teacher Judgements and Multi-factor Test Scores as Predictors of Pupil Capacity." (Unpublished Ph.D. dissertation, Wayne State University, 1958).

Roeder, Harold H. "Teacher Education Curricula--Your final grade is F." Journal of Educational Measurement, X (Summer, 1973), 141-143.

Roelfs, R. M. "Trends in Junior High School Progress Reporting." Journal of Educational Research, XLIX (December, 1955), 241-249.

Russell, Ivan L. and W. A. Thalman. "Personality: Does it influence teachers' marks?" Journal of Educational Research, XLVIII (April, 1955), 561-564.

Sparks, Merla. "Alternative to the Traditional Grading System." English Journal (October, 1967), 119-121.

Starch, Daniel and E. C. Elliott, "Reliability and Grading Work in History." School Review, XXI (1913), 676-681.

Starch, Daniel and E. C. Elliott. "Reliability of Grading Work in Mathematics." School Review, XXI (1913), 254-259.

Starch, Daniel and E. C. Elliott. "Reliability of Grading of High School Work in English." School Review, XX (1912), 442-457.

Tees, Arthur Thomas. "In Defense of Grades." School and Community, LV (October, 1968), 12-13.

Terwillinger, James S. "A Survey of Secondary School Marking Practices and Policies." National Association of Secondary School Principals, L (March, 1966), 1-37.

Theimer, Jr. William Charles. "The Relationships between Certain Personality and Educational Factors and the Grading Patterns of Secondary School Teachers." (Unpublished Ph.D. dissertation, University of Denver, 1967).

Thorndike, E. L. and E. O. Bergman. "On the form of distribution of intellect in the ninth grade." Journal of Educational Research, X (November, 1974), 271-278.

Toliem, William Joseph. "Factors which Affect the Decision Making of Classroom Teachers as they Evaluate Students." (Unpublished Ph.D. dissertation, State University of New York at Buffalo, 1972).

Trow, William Clark. "On Marks, Norms, and Proficiency Scores." Phi Delta Kappan, XLVIII (December, 1966), 171-173.

Wetzel, William A. "The use of the normal curve of distribution in estimating students' marks." School Review, XXIX (May, 1921), 373-378.

Wrinkle, William, "The Story of a Secondary School Experiment in Marking and Reporting." Educational Administration and Supervision, XXIII (October, 1937), 481-500.

Secondary Sources

Bangs C. W. and H. A. Greene. Teachers' Marks and Marking Systems (University of Iowa, 1930).

Ebel, Robert L. Essentials of Educational Measurement (Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1972).

Ebel, Robert L. Measuring Educational Achievement (Englewood Cliffs, New Jersey: Prentice-Hall, 1965).

Fleming, Cecile W. A Detailed Analysis of Achievement in High School (New York: Division of Publications, Teachers College, 1925).

Goodland, John and Robert H. Anderson. The Nongraded Elementary School (New York: Harcourt, Brace and Co., 1959).

Kirschenbaum, Howard, Sidney B. Siman, and Rodney W. Napier. Wad-ja-get? (New York: Hart Publishing Co., Inc., 1971).

Mayo, Samuel T. Pre-service Preparation of Teachers in Educational Measurement. (Toyola University, 1967).

Odell, C. W. "Marks and Marking Systems." Encyclopedia of Educational Research (New York: Macmillan, 1950).

Rothney, John W. M. "Evaluating and Reporting Pupil Progress." What Research Says to the Teacher, (March, 1955).

Ryans, D. G. Characteristics of Teachers (Washington, D. C.: American Council on Education, 1960).

Schwartz, Alfred and Steward C. Tiedmann. Evaluating Student Progress in a Secondary School (New York: David McKay Co., Inc., 1957).

Smith, Ann Z. and John E. Dobbin. "Marks and Marking Systems." Encyclopedia of Educational Research (New York: Macmillan, 1960), 783-791.

Sobel, Frances S. Teachers' Marks and Objective Tests as Indices of School Adjustment (New York: Division of Publications, Teachers College, 1936).

Starch, Daniel. Educational Measurement (New York: Macmillan, 1916).

Starch, Daniel. Educational Psychology (New York: Macmillan, 1927).

Taylor, J. Corey. The reliability of quarterly marks in the seventh grade of junior high school. (Baltimore: John Hopkins Press, 1931).

Thorndike, Robert L. "Marks and Marking Systems." Encyclopedia of Educational Research (New York: Macmillan, 1970), 759-766.

Traxler, Arthur. Techniques of Guidance (New York: Harper and Brothers, 1945).

Wrinkle, William. Improving Marking and Reporting Practices (New York: Rinehard and Co., Inc., 1947).