

REPORT CARD FORMS AND ACHIEVEMENT

by

SISTER MARY LOU PFANNENSTIEL
B. A., Marymount College, 1967

3735

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1970

Approved by:



Major Professor

ACKNOWLEDGMENTS

Sincere appreciation and gratitude is expressed to Dr. Charles Peccolo, Professor of Education, Kansas State University, by the writer. His goodness, patience, interest and invaluable assistance was a source of great encouragement.

The writer is also indebted to Dr. J. Harvey Littrell and Dr. A. Bower Sageser for serving on the Committee, to Sister Mary Augustine and Sister Francis de Sales for their patient assistance, to my pupils, their parents, and the principal of Seven Dolors Grade School, Manhattan, Kansas for their cooperation, and to my parents and friends for their encouragement and enthusiasm.

LD
2668
T4
1970
P48
C.2

III

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	3
Hypothesis	4
Limitations and Delimitations	4
Definition of Terms	5
II. REVIEW OF LITERATURE	6
Scope and Functions of Report Cards	7
Historical Review of Report Cards	12
III. DESIGN AND PROCEDURE	55
Sampling	55
Measuring Devices	64
Procedure	66
IV. REPORT OF FINDINGS	69
Analysis Techniques	69
Description of Findings	69
V. SUMMARY AND CONCLUSIONS	75
Purpose of the Study	75
Procedures	75
Summary of the Results	76
Conclusions, Recommendations and Implications	78

CHAPTER	PAGE
BIBLIOGRAPHY	80
APPENDICES	88
Appendix A	89
Appendix B	91
Appendix C	94

LIST OF TABLES

TABLE		PAGE
I.	Control Group Pre-Test Grade Scores	57
II.	Experimental Group Pre-Test Grade Scores	58
III.	Description of the Control Group	59
IV.	Description of the Experimental Group	61
V.	Pre-Test Grade Score Means of Control and Experimental Group	62
VI.	Post-Test Grade Score Means of Control and Experimental Group	70
VII.	Mean Gains of Control and Experimental Group	72
VIII.	Control Group Post-Test Grade Scores	73
IX.	Experimental Group Post-Test Grade Scores	74
X.	Summary of Test Results for Control and Experimental Group	77

CHAPTER I

INTRODUCTION

"Turmoil," according to Caudle, "is the word to describe the concern of teachers and administrators in the elementary schools, as well as the patrons, as to the relative merits of various grading systems now in use."¹ Both lay and professional magazines seem to intensify rather than subdue the controversy. In some school systems even the election of school board members was determined on the basis of reporting form preferences advocated by the candidates.²

The inhabitants of Anatevka in Joseph Stein's Fiddler on the Roof were governed by tradition. Americans are no exception for they, too, are bound by tradition in their grading system. Caudle believed that in view of the social and technological advances, which have been evidenced in recent years, it would seem that the method of reporting to parents concerning the scholastic abilities of their children were in a state of stagnation.³

Hammel, however, stated that with the complexity and compromise involved in the composition of a reporting form, educators were

¹James B. Caudle, "Educators Look at Grading," Part III. Elementary Level - "A Linger Problem," The Texas Outlook, XLV (July, 1961), 14.

²Ibid.

³Ibid.

anxiously seeking a more efficient and representative picture of pupil-teacher efforts.⁴

Report cards and grading were as frequently discussed as the weather, but rarely was there any positive, constructive action as a follow-up according to Cummins.⁵

Link envisioned an advanced system that would make grades obsolete. He believed that grades did not aid students in discovering themselves, but only served to increase tension and anxiety.⁶

"Pupil progress reports, regardless of form," stated Rolf, "are not going to solve the problems associated with achievement or lack of achievement. However, sound progress reports can aid achievement if properly designed and administered."⁷

Discouraging as is our present system, Sherry optimistically noted that whatever improvements were made, they could not be any more objectionable or inadequate than the present practices.⁸

⁴John A. Hammel, "Report Cards: A Rationale," The National Elementary Principal, XLIII (May, 1964), 50.

⁵Robert E. Cummins, "Evaluating and Grading," Education LXXXII (March, 1962), 403.

⁶Francis R. Link, "To Grade or Not to Grade," The PTA Magazine, LXII (November, 1967), 12.

⁷Fred J. Rolf, "Reporting to Parents," Ohio Schools, XXXIX (May, 1961), 10.

⁸Joseph E. Wherry, "What Are Current Trends In Reporting Student Growth and Achievement To Parents?" The Bulletin of the National Association of Secondary-School Principals, XLIII (April, 1959), 155.

Misner stated that although much experimentation with various substitute reports had been conducted, few schools had found it feasible to disregard formal reports altogether. Today's reports may differ in content and format from previous ones, but they play just as prominent a role in contemporary learning as before.⁹

Perhaps of all the controversies concerning the elementary schools, Stockard believed that the most prominent and most pertinent to the home-school relationship was the matter of grades and marking.¹⁰

"One simple solution to our problem," Brimm declared, "is to eliminate the use of the report card, but that cannot be done unless we replace it with something better. . . . The report card is a time-honored institution and it is obvious that any attempt to eliminate it would meet with strong resistance."¹¹

In view of this fact the writer conducted a comparative study in achievement of pupils receiving traditional report cards and pupils receiving a more recently developed diagnostic report card.

STATEMENT OF THE PROBLEM

Education needs to be examined from all aspects--social, physical, psycho-physical, psychological, social-philosophical,

⁹Paul J. Misner, "The Restoration of Report Cards - A Report on Reporting," The PTA Magazine, LVIII (February, 1964), 10.

¹⁰Jimmy Stockard, "Case for Dual Grading. . .Ability Versus Achievement," The Texas Outlook, XLVI (May, 1962), 30.

¹¹R. P. Brimm, "Report Cards - Yesterday and Today," The Clearing House, XXXIII (September, 1958), 18.

curricular, literary, and organizational-methodology. Will traditions regarding marks and report cards change or lessen? Do the present report cards view the "whole child," do they take into consideration the intelligence and rational capacity of the child, the dignity and worth of the individual human personality, and do they respect his freedom with responsibility? Are educators caught up in a spiral of greater efficiency at the expense of human values? How can the human spirit survive if it is killed in the schools?

The purpose of this study was to determine whether pupil achievement differed because of the type of report card received.

HYPOTHESIS

This study will concern itself with the following hypothesis:

H.: There is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards.

LIMITATIONS AND DELIMITATIONS

The writer limited the study to pupil achievement as affected by report cards.

The reference materials for this study were obtained from the Kansas State University and Marymount College Libraries at Manhattan and Salina, Kansas, respectively.

The pupils were selected from the sixth grade class of Seven Dolors Grade School, Manhattan, Kansas. The sampling was limited to one classroom.

The necessity of parent-teacher conferences, because of school policies added an unwanted dependent variable to the study.

DEFINITION OF TERMS

The following terms as used in this study are defined for the purpose of clarification.

Report Card. (as an instrument of evaluation). "A bundle of compromises tied together by the sincere efforts of a study committee which sought to produce a card that would reflect the local educational enterprise."¹²

Traditional Report Card. This term includes percentage grades, letter grades (the five letter system A-B-C-D-E or -F), (the four letter system E-G-F-P), (the three letter system S-I-U), (the two letter system P-F), check lists and the dual system of achievement grades and ability grades.

Diagnostic Report Card. (concerned with prescription for the pupil). "It is analytical in that it seeks to identify strong and weak aspects of a pupil's performance so appropriate corrective measures may be taken."¹³

¹²Hammel, op. cit., p. 50.

¹³Mauritz Johnson, Jr., "Solving the Mess in Marks," New York State Education, XLIX (November, 1961), 12.

CHAPTER II

REVIEW OF LITERATURE

Report cards are an important as well as a useful and integral part of education for they ascertain a child's educational and vocational future. They record the achievement of all pupils at all levels, and school records are considered incomplete if not substantiated by marks.

Wrinkle stated that it was unnecessary to delve into the history of education to find "that tradition, whatever good may be attributed to it, is constantly operating as a retarding influence tending to delay progressive adjustments long after the justification for such adjustments is apparent."¹ He maintained that the social and economic development of our civilization surpassed the curriculum development of our schools. However, the present century has subjected the traditional practices in school marking to study and revision. This is more apparent in elementary education for "it is the greatest distance from academic domination by higher institutions and the academically inclined teacher."² Often entrance into institutions of higher learning and scholarship grants are determined in large part by marks made in high school.

¹William L. Wrinkle, "School Marks - Why, What, and How?" Educational Administration and Supervision, XXI (March, 1935), 218.

²Ibid.

SCOPE AND FUNCTIONS OF REPORT CARDS

School marks may be classified under various headings. Wrinkle in his book, Improving Marking and Reporting Practices In Elementary and Secondary Schools, classified them under headings representing their basic functions as administrative, guidance, information, and motivation and discipline functions,³ while Smith and Wright stated them as administrative, educational, and social.⁴

The administrative functions of promotion, placement, graduation, and transfer prove to be the most obvious uses of the report card. Because the report card is the all-important predictor of future success, it is vitally important that it be as accurate as possible for the promotion or failure of promotion aspect has a definite influence on a pupil's whole educational career.⁵

Wrinkle believed that relative to the functions of administration, the teachers, the school and its practices, and the administration exist for the education of the pupil.⁶ Report cards, serving as an evaluating scale, are important to parents, pupils, and teachers. (It was in this

³William L. Wrinkle, Improving Marking and Reporting Practices in Elementary and Secondary Schools (New York: Rinehart and Company, 1947), p. 31.

⁴Henry L. Smith and Wendell W. Wright, Tests and Measurements (New York: Silver, Burdett and Company, 1928), p. 11.

⁵Pervival M. Symonds, Measurements in Secondary Education (New York: The Macmillan Company, 1927), p. 499.

⁶Wrinkle, "School Marks - Why, What, and How?" Educational Administration and Supervision. p. 218.

manner that the writer viewed the scope and functions of report cards.)

All ratings have essentially the same function--to report a child's learning progress to his parents. The purpose of reporting to parents may be summarized as twofold: (1) to impart information to the parents concerning their child's progress and standing in school, and (2) to promote understanding and cooperation between the home and the school. Because parents tend to rely solely on report cards for information regarding their child's progress, they want sufficient factual information in order to evaluate their child's progress.⁷

Misner quoted a parent (Mrs. Sternberg) as saying, "We like to use the reports as a means of helping our children understand their individual strong points and weak points. We can show them where they're making progress and where greater effort is needed. In this way children learn to evaluate themselves."⁸

Detter stated that the method of evaluation is non-essential since "the main purpose in reporting to parents is to enable them to cooperate more intelligently with the school in guiding the child."⁹ The parent is, therefore, entitled to an accurate evaluation of his child's progress because the child is his most prized possession and responsibility.

⁷Misner, op. cit., p. 11.

⁸Ibid.

⁹Shirley Detter, "Reporting Progress to Parents," School Activities, XXX (November, 1958), 82.

Williams believed that the total reporting system can serve not only as a powerful force in protecting the pupil's self-respect and self-confidence but also motivate the pupil toward greater growth and development.¹⁰ Realizing that his progress is important, he must understand what the report reflects and be able to interpret the report according to his needs or progress and thereby encouraged to improve.¹¹

Chadwick aptly summarized the value of the report card when he said that it must

present to the child a realistic and acceptable picture of himself, a picture that gradually leads him to a realization of the person he is and may become. His strengths and his weaknesses should be made clear. At the same time, he must see himself as a worthwhile member of society with whom he can live happily and comfortably. Unless we can do this for each child all else we attempt to do is largely a waste of time.¹²

Williams noted that often pupils interpreted report cards "as approval or disapproval of them as persons rather than evaluations of school work."¹³

Cagle considered the purpose of marking and evaluating as helping students to maximum growth¹⁴ while Burton regarded marking

¹⁰Lois Williams, "Teachers and Parents: Did You Know That Your Children Feel This Way?" Childhood Education, XXXV (October, 1958), 64.

¹¹Marilyn H. Cutler, "Does Your Report Card Format Rate an A?" The Nation's Schools, LXXII (September, 1963), 60.

¹²Ruth E. Chadwick and others, "The Report Card in a Nongraded School," The National Elementary Principal, XLII (January, 1968), 28.

¹³Williams, op. cit., p. 61.

¹⁴Dan F. Cagle, "How May We Make the Evaluation and Reporting of Student Achievement More Meaningful?" National Association of Secondary School Principals Bulletin, (April, 1955), 27.

as a means of encouraging students through an increased home-school cooperation.¹⁵

Morris saw the purpose of marking aimed at increasing the effectiveness of student learning rather than just recording results as had been the case prior to new developments in educational philosophy.¹⁶

The underlying principle of marking, according to Bolmeier, was that marks had to be designed and utilized for the student's benefit and not for the teacher's convenience.¹⁷

Report card grades tend to be misconstrued. Students who have little self-confidence are discouraged with low grades while a superior student feels proud about his high grades--grades often earned with little or no effort.¹⁸

Maxson wrote that "the grade goes far beyond the students' first reaction of elation, dismay, indifference, or resentment."¹⁹

¹⁵William H. Burton, The Guidance of Learning Activities (3rd ed.), (New York: Appleton-Century-Crofts, Inc., 1962), p. 601.

¹⁶Lucile Morris, "Evaluating and Reporting Pupil Progress," Elementary School Journal, LIII (November, 1952), 144-147.

¹⁷Edward C. Bolmeier, "Principles Pertaining to Marking and Reporting Pupil Progress," School Review, LIX (January, 1951), 16.

¹⁸Brimm, op. cit., p. 17.

¹⁹Wilbur B. Maxson, "Grading, A Serious Matter," NEA Journal, LIII (October, 1964), 56.

True grades may evoke enthusiasm for improvement, but they may also provoke a "What's the use?" attitude for they serve as a status symbol. Maxson realized the power of grades when he stated: "When I take my grade book in hand, I have great power over the lives of my students."²⁰

The report card is not only important to the parents and pupils but also to teachers. It is perhaps the latter who incurs the greatest responsibility toward both parents and pupils. The teacher is confronted with the problem of reporting accurately to parents on the progress or failure of their child. Prudence demands tact so that the results will be constructive. On the other hand a pupil's growth and performance are determined by the teacher's decision based on the knowledge of data which he collected, analyzed, and evaluated.²¹

White and Boehm voiced the cry of many teachers. Elementary pupils have little or no concern for grades, and if they do, the concern is induced by the parents.²²

Teachers have become more aware of pupils as individuals. The concept of person became real and meaningful while the positive rather than the negative was stressed for a pupil thrives on acceptance and wilts under anxiety.²³

²⁰ ibid., pp. 56-57.

²¹ Faith Smither, "Report Cards - Problems and Possibilities," The National Elementary Principal, XL (September, 1960), 168.

²² Mary Alice White and Ann E. Boehm, "The Child's World of Marks," NEA Journal, LVII (January, 1968), 13.

²³ William, op. cit., p. 60.

Teachers in conjunction with parents have striven to develop some forms of reporting conferences and evaluations which would serve as an aid to the pupil's growth--growth "in terms of the child's abilities rather than in comparison to those of others or to the average of his group."²⁴

Smither's advice to both parents and teachers, the educators of youth, should also be heeded today:

If teachers and parents want children to fulfill their dreams, if schools are to prepare young people for the world ahead, we should take a different path. We should be observing, feeling, and trying out the potentials of children. We should be seeking to know their motivation to learn, their eagerness for certain experiences, their growing sensitivity to people, and their awareness of life, its problems and promises.²⁵

HISTORICAL REVIEW OF REPORT CARDS

The interest in systematic reporting of learning progress may be roughly divided into two periods: (1) 1910-1940 in which the semantic and mechanical problems of marking were the focal interest of research; and (2) 1940-present in which the improvement of marks in the area of communication and in comprehensiveness had demanded vital interest.²⁶ For convenience and clarity, the writer has further divided these periods into decades.

²⁴ibid.

²⁵Smither, op. cit., p. 171.

²⁶John E. Dobbin and Ann Z. Smith, "Marks And Marking Systems," Encyclopedia of Educational Research. (3rd Ed.), (New York: The Macmillan Company, 1960), 783.

1910-1920

Concern and controversy regarding marking can be traced to the period 1910 to 1920 when studies indicated the unreliability and variability of marks. All levels of education during this period favored the percentage system which was in general use, with a trend toward the use of a three to seven point system. Arranging scores in order and then changing these scores into marks according to an adopted distribution found supporters as well as antagonists.²⁷

In a 1910 study of fifteen thousand or more grades given by approximately two hundred and fifty teachers in elementary and high schools, and in the College of Letters and Science (University of Wisconsin), Dearborn concluded that the chief causes of inequality were believed to be due to a lack of uniformity in standard.²⁸

Johnson declared that an in-depth study was not essential to notice the lack of uniformity of standard in grading between secondary schools and colleges, between different secondary schools, and within the same school between different departments and even between different teachers in the same department. The value of the results after spending enormous amounts of time and energy by teachers, principals and clerks in grading, making monthly, quarterly or half-yearly reports,

²⁷ Ibid.

²⁸ Walter Fenno Dearborn, School And University Grades (Wisconsin: University of Wisconsin Press, 1910), p. 6.

and transferring them to permanent school records did not commensurate with the effort involved.²⁹

The Kelly study, published in 1914, showed the variability of marks given to pupils by the teachers. Kelly concluded that marks meant a variety of things to different teachers, even to the point where in some cases it meant a difference between an F- (fair minus) and a G (good) in elementary school systems using the E-G-F-P (excellent, good, fair, and poor, respectively). High schools faced the same problem. An example illustrated the point. Two schools had the same point system for its passing standard, yet seventy points in one school meant more than eighty-one points in the other.³⁰

Teachers varied in their reliability of marking and standards of marking under individualistic marking systems in use. Rugg stated:

There are large individual differences in teachers' marks of the same students in the same subjects, on the same examination papers and the same drawings and lettering samples. The mean variations in many of the instances tested run as high as 15 percent. They practically never are less than 5 percent.³¹

Starch described the wide discrepancies in standards of marking in English, mathematics, and history as a result of an investigation

²⁹Franklin W. Johnson, "A Study of High-School Grades," School Review, XIX (January, 1911), 13.

³⁰Frederick James Kelly, Ph.D., Teachers' Marks Their Variability And Standardization (New York: Teachers College, Columbia University Press, 1914), p. 133.

³¹Harold Ordway Rugg, "Teachers' Marks and Marking Systems," Educational Administration and Supervision, I (January, 1915), 137.

which was made to determine the range of variation and the reliability of the marks assigned by different teachers to the same papers. A startling and almost shocking fact of this investigation was the tremendously wide range of variation. Marks assigned by different teachers to the same paper varied enormously. The unreliability and variability of marks was as great in one subject as in another. Even mathematics grades were as unreliable as language or history grades. Yet despite these results, Starch saw marks as indispensable for their real personal value and administrative value.³²

An experiment in the Lawrence, Kansas schools conducted by Jaggard evidenced that a group of teachers could extensively correct their faulty distribution of marks when they were properly informed. The systematic effort to educate the teachers on what was a proper distribution of marks seemed to be the chief reason for improvement in the grading system.³³

Cajori, in 1914, proposed that school grades A, B, C, D, E, be distributed symmetrically in the proportions 7-24-38-24-7 percent.³⁴

Resulting from the previous fifteen years of discussion led Rugg to state that all agreed that the methods for measuring the

³²Daniel Starch, Ph.D., Educational Measurements (New York: The Macmillan Company, 1916), p. 8.

³³Guy H. Jaggard, "Improving the Marking System," Educational Administration and Supervision, V (January, 1919), 34.

³⁴F. Cajori, "A New Marking System and Means of Measuring Mathematical Abilities," Science, XXXIX, (1914), 874.

outcome of instruction should be reconstructed because of variability, unreliability, and inconsistency in pupil marks given by teachers. The marks were on a subjective basis with no uniform standards. A recommendation of a five-division marking scale (A-B-C-D-E or excellent, superior, medium, inferior, and poor) was suggested. Rugg confirmed the recommendation by stating that teachers could accurately handle five divisions and also suggested the desirability of "measurement by ranking with subsequent transmutation to absolute marks by means of a distribution-curve" as a means of helping to rebuild the marking system.³⁵

R. H. Johnson called the direct marking system (100 or 10 is the highest mark in which marks are given as tenths or hundredths of this grade) the "naive system." Although it was in common use, all versed in statistical method rejected it. The use of the coefficient method in which the grade was determined by dividing the student's marks by the class median for the same work as advocated by Johnson. The median was graded as one (1). Marks above and below were determined by the following formula: $\frac{\text{student's mark}}{\text{median mark}} = \text{grade}.$ ³⁶

In 1915 Rugg reported that teachers' marks on report cards greatly influenced various phases of educational administration, yet school and college administrators did not recognize the importance of

³⁵Harold O. Rugg, "Teachers' Marks and the Reconstruction of the Marking System," Elementary School Journal, XVIII (May, 1918), 702-18.

³⁶Roswell H. Johnson. "The Coefficient Marking System," School and Society, VII (June, 1918), 714-715.

the problem. These report card marks largely controlled promotions of instructional levels, acceptance or rejection to schools, bestowal of honors and degrees which were viewed as an index to teaching efficiency, honorary societies membership election basis, scholarship and fellowship appointment basis, and even college teaching positions. While complex or blanket abilities, development, attainment, and accomplishment were measured directly, educational specialists agreed that native ability or capacity were measured indirectly by these report card marks. Specialists preferred letter marks to percentile marks and accepted a five division marking system as the best for many administrative and logical reasons.³⁷

1920-1930

In the 1920's, ability grouping and individualization of instruction received impetus from the increased use of standardized intelligence tests and added to the problem concerning marks, especially to the range of marks to be used for each ability group.³⁸

The period 1920 to 1930 stimulated many facets of life including education. The wealth of ideas had a marked influence on marking. Individualization of instruction had become an important topic and it highlighted interest in ability groupings and accelerated changes in

³⁷Harold Ordway Rugg, "Teachers' Marks and Marking Systems," Educational Administration and Supervision, 1 (January, 1915), 117-135.

³⁸Dobbin and Smith, op. cit., p. 785.

promotion practices. Improvements took place in the quality and quantity of educational materials as well as in the area of measurement techniques, increasing the use of standardized tests. Rapid replacement of the percentage system occurred with the use of four to seven symbols.³⁹

Thorndike and Bregman reported a research study of ninth grade intelligence. Their conclusion supported the use of the normal distribution curve in assigning marks.⁴⁰

The use of the normal curve was accepted by Monroe, but he emphasized that a standard distribution was only a device used to reduce errors in grading and not as an end in itself. Monroe stated:

Whenever common sense indicates that the distribution of grades for a particular class should depart from the standard distribution, no instructor should hesitate to award the grades which he believes his students deserve. A standard distribution will be closely approximated only in large unselected groups of students. Relatively few classes in high school include more than 35 students and not infrequently the group is selected. Hence significant departures from a standard distribution may be expected. On the other hand, the distribution of grades will frequently approach the standard shape.⁴¹

To insure uniformity of standards, the normal probability curve had become the most widely accepted plan. Oftentimes classes were not

³⁹ibid., p. 783.

⁴⁰E. L. Thorndike and E. O. Bregman, "On the Form of Distribution of Intellect in the Ninth Grade," Journal of Educational Research, X (November, 1924), 278.

⁴¹Walter S. Monroe, Directing Learning in High School (New York: Doubleday, Page and Company, 1927), pp. 527-528.

normally distributed and thus the normal probability curve received criticism. Yet, Symonds felt that more injustice was done by teachers having free rein in marking than by the utilization of the normal probability curve.⁴²

Freyd pointed out the advantages of the popular graphic rating method which originated in the Scott Company Laboratory in 1920. The methods of rating on a line and the checking descriptive terms were not original. However, the combination of both gave the graphic rating scale a novel feature. The study proved that ratings of the same subjects made by various raters were similar.⁴³

In 1925, Symonds observed that rating scales were being more commonly used in education as well as in personnel work in industry. Unmeasurable qualities and traits were quantitatively determined by rating scales. Ranking, though generally considered more reliable, was not preferred over rating scales. Rating was considered more pleasant than ranking. The graphic rating scale presented fewer problems in its usage and was, therefore, more acceptable. Experimentation by Symonds produced evidence that ranking and graphic rating scales yielded similar results. However, confusion arose when ranking involved large groups. Symonds explained it as follows:

⁴²Percival M. Symonds, Measurement in Secondary Education (New York: The Macmillan Company, 1927), pp. 510-517.

⁴³Max Freyd, "The Graphic Rating Scale," Journal of Educational Psychology, XIV (February, 1923), 92-94.

In rating, comparison is made with roughly defined steps or classes in the scale and the individual is placed in the nearest fitting class according to the opinion of the rater. But in ranking, one individual is compared directly with another.⁴⁴

The competence and moral character of teachers were not questioned because of a disagreement on school marks, but it was a major criticism against the system of assigning marks according to the Trabue research studies published in 1924. In the training of teachers, colleges gave no attention to methods and rules in marking. Studies showed that even teachers in the same departments of the same schools disagreed in the distribution of their marks, and that different teachers varied up to forty percentage points in marking the same pupil examination papers. Trabue proclaimed that marks would not have highly significant meanings unless teachers were taught to use more objective evidence as the bases for marks as well as utilizing better examinations with uniform methods of scoring.⁴⁵

A clearer definition of the marking base was needed during the 1920's. Ruch stated that marking was arbitrary, and that without definition, marks had no meaning. Two things were essential if a defensible plan for evaluating pupil's accomplishment was developed by a school system. Ruch stated these as

⁴⁴Percival M. Symonds, "Notes on Rating," Journal of Applied Psychology, IX (1925), 188-190.

⁴⁵Marion Rex Trabue, Ph.D., Measuring Results in Education (New York: American Book Company, 1924), pp. 49-56.

1. The pupils must be placed in correct, relative positions or ranks with respect to each other; and
2. The adopted marking scheme must be defined. Its sole meaning and value rest upon its definition to pupils, teachers, and parents alike.⁴⁶

Ruch believed that nearly any properly defined marking system would be good if the teacher had an adequate basis for evaluating each pupil in rank-order. The basic data had to be valid and reliable.

Almost any scheme of recording marks, provided it be adhered to by all teachers in the same school or school system, and provided further that it be understood by all concerned, will prove adequate if there is a valid and reliable provision for the measurement of the relative abilities of the pupils to be graded. At the same time, the definition of local practices is essential in order that there be meaning to the final marks given.⁴⁷

Karrer advocated a new method of grading in which he proposed the speed of learning as the basis of the grading system. Teachers were judged by Karrer according to the acceleration of the speed of learning or to the maintenance of a high level of speed of learning regarding their pupils. Several major requirements that reporting systems should meet were as follows:

- (1) The mark should show or measure an individual's positive quality or his characteristics,
- (2) the measure should be a simple numerical quantity,
- (3) the measurement should involve an objective method,
- (4) the method should be applicable to the applier as well as to the applied,

⁴⁶Giles Murrel Ruch, The Objective or New-Type Examination (New York: Scott, Foresman and Company, 1929), p. 376.

⁴⁷Ibid., pp. 377-378.

- (5) the method of measurement (reporting system) should be statistically sound, and
- (6) the numbers or marks must be easily interpreted.⁴⁸

To determine the general ability level of any group of pupils, Symonds advocated the use of intelligence tests because these tests represented a measure of ability which was applied to the entire group and had a tolerable degree of accuracy. He emphasized that these intelligence tests used were not to be used as achievement measures.⁴⁹

Abell recommended the use of the standard deviation as the basis for marking and pointed out that the real advantage of using the deviations was the fact that they were relative measures. This standard deviation method was tested extensively by Abell in the late 1920's.⁵⁰

Some educators felt that pupils should be marked on an intelligence test basis because studies showed a perfect positive correlation between ability and achievement with maximum motivation present. Symonds strongly disagreed because maximum motivation did not occur in real life classroom situations, and thus, marks should be based strictly on achievement and independently of intelligence test scores. Achievement should be measured by objective evidence (test results) and used for the marking system. Symonds stated:

⁴⁸Enack Karrer, "Reflections on a New Method of Grading," School and Society, XXIV (November, 1926), 582-583.

⁴⁹Percival M. Symonds, Measurement in Secondary Education, (New York: The Macmillan Company, 1927), p. 517.

⁵⁰E. L. Abell, "Grading From the Median," Journal of Educational Research, XVIII (December, 1928), 426.

Let it be stated here with emphasis that school marks should be thoroughbred and not hybrid. School marks should represent only one thing - achievement. School marks must not be a composite of several different qualities, otherwise they mean nothing. Above all, school marks should not be based merely on conscientious work. Quite unconsciously teachers use marks not only as a measure of achievement but as a disciplinary weapon.⁵¹

Research studies were undertaken in the 1920's to determine the correlation between school marks and intelligence-test scores. Ohlson reported the correlation between the average school marks and the Terman group test of mental ability for all five hundred six graduates of the Everett, Washington High School as .38 with a $\pm .03$ probable error. The correlation for the boys and girls was also taken separately and proved surprising. The boys had a .32 ($\pm .04$) correlation and a .47 ($\pm .02$) correlation was reported for the girls. Ohlson concluded that girls were more conscientious. The boys, equally intelligent or possibly more so (the median intelligence score for the boys was ten points higher) were more inclined to be happy-go-lucky. Possibly because girls were more quiet and thus less irritating to the teachers could account, in part, for the higher correlation.⁵²

Bolton did a re-study of some of Starch's data. (Starch showed evidence through research data of the variability of marks different teachers gave to the same examination papers.) His new data seemed to

⁵¹Symonds, Measurement in Secondary Education, pp. 505-506.

⁵²David Ohlson, "School Marks vs. Intelligence Rating," Educational Administration and Supervision, XIII (February, 1927), 93-102.

nullify Starch's claim to little or no uniformity in marks given by teachers. Bolton's opposing evidence indicated that under usual conditions (examinations given the teachers conducting the course and marking the examinations themselves) teachers' marks were reliable. He stated:

In all probability there is sufficient reliability in teachers' marks to justify their continued use as a means of determining promotions in the grades or of graduation from the high school or college and for purposes of determining college entrance. When all the grades assigned to a pupil at all times in a given subject, and all the grades given by the different teachers in the pupil's school career are massed and a composite rating is secured undoubtedly it represents quite fairly the pupil's past performance. Incidentally other studies go to show that such a composite is fairly prophetic of future accomplishment.⁵³

Concerning tests and markings, Bolton asked,

Has there not been altogether too much attention given to tests and markings during the last few years? The rank and file of teachers have been required to spend so much time on types of examinations and methods of marking that their entire attention has been badly diverted from gathering materials for teaching and studying best methods of instruction. Many seem obsessed with the idea that materials and methods of instruction are correct and that all energy must be put upon some methods of marking that will discover abilities and make possible the segregation of the "dumbbells" from the geniuses.⁵⁴

1930-1940

A standard distribution of A, B, C, D, E grades of 6-22-44-22-6 percents was proposed by Eells in 1930. On the assumption of normality

⁵³Frederick E. Bolton, "Do Teachers' Marks Vary as Much as Supposed?" Education, XLVIII (September, 1927), 38.

⁵⁴ibid., p. 39.

of distribution, the unit distance between true means of successive grade groups would be uniformly one sigma, with less than one percent error. During the past twenty-five years, various writers proposed different distributions of grades for the five-point grading system which contained discrepancies in the relative length of units between successive grade means ranging from four to sixteen percent.⁵⁵

Davis severely criticized Eells' proposal of the 6-22-44-22-6 normal curve system and concluded in a research study that comparing grades beyond a semester had little meaning and that the grade value changed in succeeding semesters.⁵⁶

In a study undertaken to determine some general trends, Middleton stated:

At first, the task of investigating the literature seemed to be a rather hopeless one. What a mass and a mess it all was! Could order be brought out of such chaos? Could points of agreement among American educators concerning the perplexing grading problem actually be discovered?⁵⁷

Middleton and his committee on grading studies carefully a well-selected bibliography of over eighty books and journal articles which revealed a strong preference for standardization in grading, the use of the normal probability curve and grading areas, the publication

⁵⁵Walter Crosby Eells, "An Improvement In the Theoretical Basis of Five Point Grading Systems Based on the Normal Probability Curve," Journal of Educational Psychology, XXI (February, 1930), 135.

⁵⁶J. DeWitt Davis, "The Effect of the 6-22-44-22-6 Normal Curve System on Failure and Grade Levels," Journal of Educational Psychology, XXII (November, 1931), 639.

⁵⁷Warren C. Middleton, "Some General Trends in Grading Procedure," Education, LIV (September, 1933), 5.

of each teacher's grades at the end of each term, and a favoritism of a five-point grading scale (usually A, B, C, D, and F) along with a normal distribution of grades.⁵⁸

Gould's questionnaire study method revealed minor uniform standards for measuring pupil progress. Regarding the distribution of school marks, the study revealed that the normal curve was conservatively used by the teachers, thus refuting the accusation that teachers slavishly use the normal curve.⁵⁹

Norsted deplored the fact that

Though marks may be defined and understood, there still remains a great lack of agreement due to variations in the interpretation of marks, relative merit of work, weighting of course requirements, use of the distribution curve, sex of teachers and students, standards of judgment, influence of extraneous elements, and lack of plain common sense.⁶⁰

According to Forman, education tended toward a markless or "gradeless" age. Grades and marks were vices perpetuated by tradition. Since progress required the removal of vices, grades and marks had to be removed from our educational system. They were viewed as arbitrary standards and subjective measures conforming to the laws of chance and as artificial stimulators furnishing fear motivation resulting only in artificial education. Colleges and universities were responsible for the pressures and demands for grades and marks. Forman advocated the

⁵⁸ibid., pp. 5-10.

⁵⁹George Gould, "Practices In Marking and Examination," The School Review, XL (February, 1932), 146.

⁶⁰Roy A. Norsted, "To Mark or Not to Mark?" The Journal of Education, CXXI (March, 1938), 82.

"creative credit" plan which provided credit to a pupil who created anything showing original effort, use of knowledge, and work completed.⁶¹

Tiegs claimed that school systems used the fear of failure as a motivational factor when in reality it was a definite obstacle to learning and affected the morale of pupils and the general public. Pupils should be promoted unless extraordinary circumstances prevent it because of very limited definite knowledge concerning the nature of failure as well as the uncertainty in the ability to predict success.⁶²

The use of the school mark as a motivating factor was detrimental maintained Wrinkle. The use of the hickory-stick for motivational purposes was viewed as inhumane and thus was discontinued in our evolutionary education, yet mental flogging of pupils for non-conformity by the marking system on the report card still existed. The mark was inadequate as a device used to inform parents and pupils. Wrinkle stated:

Probably the best report form would be a blank sheet of paper on which the teacher would make pertinent statements regarding the progress and achievement and the weaknesses and strengths of the student.⁶³

Campbell believed that pupil effectiveness and efficiency were greatly reduced because marks caused constant uneasiness and it was impossible for a great number of pupils to fully realize their maximum potential. Health could be impaired by the overconscientious pupil,

⁶¹W. O. Forman, "The Gradeless Era in High School," Journal of Education, CXI (May, 1930), 501-502.

⁶²Ernest W. Tiegs, Ph.D., Tests and Measurements For Teachers (New York: Houghton Mifflin Company, 1931), 193.

⁶³Wrinkle, "Social Marks - Why, What and How?" pp. 219-224.

while the rest became discouraged. Every teacher should share the information regarding marks by explaining and defining them with the pupil so that each pupil "may know, and knowing, may succeed."⁶⁴

Marks give pupils a perverted attitude toward his education and learning pointed out Hillbrand. He condemned the "rank-in-class" method and viewed marking systems based on it as a form of social snobbery. In the days of feudalism, rank-in-class as a method of social appraisal could have been accepted perhaps, but not in twentieth-century America. Hillbrand hoped to see the day when grades would be abolished and students would realize that education is a growth process.⁶⁵

Odell supported the retention of school marks, although some argued for their abolishment. The chief fault, according to him, was their unreliability and subjectivity because different teachers based marks on many different factors and had a variety of standards in mind in giving them. The significance and basis of the marks should be decided. A system of marking that employed only five or six letters or other symbols, of which two were failing, was viewed as being more satisfactory than percentile marks, despite their popularity. Regarding the normal curve, Odell suggested that in giving marks it was better

⁶⁴Laurence R. Campbell, "So Pupils May Know," School and Society, XXXII (December, 1930), 762-763.

⁶⁵Earl K. Hillbrand, "College Marks - and Remarks," School and Society, XXXIII (May, 1931), 727-729.

to adopt limits for the percents of pupils receiving each mark rather than giving single exact percents.⁶⁶

Numerous pupils credit--or discredit--their teachers with a kind of "I'll-give-you-any-grade-I-damn-please-and-you-can't-do-a-damn-thing-about-it" complex, according to Wakeham. Most progressive teachers hate inflicting grades, but as long as grades are part of the educational system, effort must be made to inject helpful elements into the system. Grades should be impartial, impersonal, uniform, fair, and comprehensible to the pupil.⁶⁷

Despite the fact that tests may be made by technicians, Williamson argued that local administrators and teachers determined the standards and objectives. The danger was not that test-makers will dominate educational practices, but rather that educators failed to define clear objectives and failed to evaluate their efforts for achieving these goals. If education were properly defined, all valid measuring instruments would be usable regardless of the source of construction.⁶⁸

Fay, despite the fact that he openly admitted the obvious unreliability of the results of his research study, offered some

⁶⁶Charles W. Odell, Ph.D., Educational Measurement in High School (New York: The Century Company, 1930), 469.

⁶⁷Glen Wakeham, "Humanizing Grades," School and Society, XXXIV (October, 1931), 596.

⁶⁸E. G. Williamson, "The Cooperative Guidance Movement," The School Review, XLIII (April, 1935), 279-280.

tentative conclusions in which he stated that A students do better when told their marks; B students somewhat better; C students only slightly better while lower intelligent students needed a knowledge of marks as an incentive to increase their achievement.⁶⁹

Convincing evidence in a study on teachers was presented by Lawson. Even when teachers have similar educational preparation, actual teaching experience, cultural background as well as present environment, they still cannot estimate reliably the marks to be assigned for essay-type examinations.⁷⁰

1940-1950

De Pencier, in giving a resume of the pupil progress reporting trends of the period 1938-49, stated that the same outlook and concepts found in other phases of education were present in the reporting trends. The greatest emphasis was on the "whole child" and on techniques which help the child's all-round development. "Reporting has been called the most retarded phase of American education."⁷¹

Smith and others on the evaluation staff of the Commission on the Relation of School and College of the Progressive Education

⁶⁹Paul J. Fay, "The Effect of the Knowledge of Marks on the Subsequent Achievement of College Students," The Journal of Educational Psychology, XXVIII, (October, 1937), 554.

⁷⁰Douglas E. Lawson, "Scoring of Subjective Tests with Several Variable, Controlled," Elementary School Journal, XXXVIII (February, 1938), 457.

⁷¹Ida B. De Pencier, "Trends in Reporting Pupil Progress in the Elementary Grades, 1938-1949," Elementary School Journal, LI (May, 1951), 519-522.

Association published in 1942 their findings regarding evaluation, records and reports of student progress in thirty schools. The committee obtained report cards used by various kinds of schools for the purpose of careful study. An analysis was made of the suggestions and criticism sent in by the schools. The most popular demand was for a report card that would provide usable information of the pupil's strengths and weaknesses, and would replace the letter or numerical marks. The single mark only hid the facts about pupil progress. A report card form showing pupil strengths and weaknesses would provide an analysis of pupil achievement and serve as a safe guidance basis. The consensus of opinion was that marks had become much too competitive, were harmful to pupils, and drew the attention of pupils, parents, and teachers toward the marking symbols per se rather than toward the real purposes of education. Most schools, that had replaced their marking system, used the writing paragraphs form of report card in which pupil growth was analyzed by each teacher. This method proved to be very time-consuming. Smith warned that parents and pupils must receive some explanation about the information provided on the report card to avoid antagonism and confusion.⁷²

Good observed that many educators believed marks should have been abolished because they were more harmful than helpful. In spite of marking system improvements and greater teacher effort to use the

⁷²Eugene Randolph Smith and others, Appraising and Recording Student Progress (New York: Harper & Brothers, 1942), pp. 488-493.

mark conscientiously, the simple truth established the fact that the marking system proved valueless. Marks were not reliable, usually discouraged pupils, and hindered learning.⁷³

Sanders identified the value of a report card as dependent upon the value attached to it. If it must be given, then it must be given in wisdom and received in understanding. Sanders stated:

School reports have led to lying, cheating, stealing and suicide. Anguish and heartbreak, persecution and punishment are too common for sympathy. False values are inevitable. Wrong attitudes are developed. Surely it is time that reports should be recognized for what they are and be used as an aid, instead of a hindrance to education.⁷⁴

In a research study, Michaelis and Howard concluded that the greatest needs for improvement in evaluation concerned the personal-social development of the pupil. Little was done in the California schools to appraise emotional and social adjustments as well as the related needs of the learner.⁷⁵

After ten years of continuous study and experimentation for the purpose of improving marking and reporting practices, the research-laboratory school of Colorado State College of Education in Greeley, Colorado reported that the important discovery was that "educational

⁷³Warren R. Good, "Should School Marks Be Abolished?" The Education Digest, XI (December, 1945), 11-12.

⁷⁴Eugene Sanders, ". . . Behind the Report Cards," Nation's Schools, XXXI (February, 1943), 31-32.

⁷⁵John U. Michaelis and Charles Howard, "Current Practices in Evaluation in City School Systems in California," Journal of Educational Research, XLIII (December, 1949), 260.

objectives should be set up in terms of behavior." Wrinkle recommended that pupils serve as working partners in the development of new practices in school systems. The same consideration should be given to these pupil suggestions and recommendations as were given to the faculty. Six fallacies as seen by Wrinkle supported the use of the single letter marking system:

- (1) The mark is an effective conveyor of information;
- (2) anyone can achieve any mark he wishes if he is willing to make the necessary effort;
- (3) people succeed in out-of-school life about the same as they do in school;
- (4) the mark is rightly comparable to a pay check;
- (5) marking practices provide a justifiable introduction to competitive adult life; and
- (6) the mark can be used as a means without its eventually being recognized as an end in itself.⁷⁶

Berman pointed out that little or no valid research had been done regarding the junior high school report card. In fact, the junior high school report cards incorporated the worst features of senior high school reporting systems despite the fact that junior high school education had its own avowed objectives. Berman concluded that the tradition of a competitive rating system was being continued as shown in the survey began in 1940 of one hundred forty-nine report cards of junior high schools from across the country. Realistic recognition of differences in the ability to achieve and in the rate of learning must take place if social significance is to be achieved as presented by the objectives of the junior high school program of education.⁷⁷

⁷⁶Wrinkle, Improving Marking and Reporting Practices in Elementary and Secondary Schools. pp. 49-112.

⁷⁷Samuel Berman, "Revising the Junior High-School Report Card," National Association of Secondary-School Principals Bulletin, XXVII (May, 1943), 49-62.

According to Drake, the first tests emphasized acquiring knowledge and skills. Tests concerned with "intangibles," such as understandings, attitudes, and appreciations of the learner emerged. These intangibles received great attention in the past five years, and will more so in the next decade. He proclaimed, "We must become more interested in children themselves than in devices and subject-matter. Evaluation in the future should have human welfare as its chief concern."⁷⁸

1950-1960

Vredevoe, in reporting a research study based on personal interviews extended over four years with teachers and administrators, revealed that teachers differed in interpreting achievement, the value of school marks differed from school to school and from teacher to teacher within the same school, the use of "satisfactory" and "unsatisfactory" did not offer a solution to the reporting problem, some teachers added a plus or minus to the "S" and the "U" in order to widen the grading range, secondary schools used the A, B, C, D, E, or their equivalent while elementary schools tended to provide other methods.⁷⁹

⁷⁸C. Elwood Drake, "Trends in the Field of Evaluating Secondary Education," Educational Administration and Supervision, XXVI (April, 1940), 242-243.

⁷⁹Lawrence E. Vredevoe, "How May We Make the Recording and Reporting of Pupil Achievement More Meaningful?" National Association of Secondary-School Principals Bulletin, XXXVII (April, 1953), 180-181.

A research study aimed at identifying recent significant trends in pupil progress reporting of the junior high schools was conducted by Roelfs. It was reported that nearly all of the schools studied used report forms of one kind or another. The report was supplemented with parent-teacher-pupil conferences, telephone calls, letters to parents, and other informal methods by thirty-eight percent of the schools studied. Not a single junior high school in the research study relied solely on informal methods. There was a trend for subdividing each academic subject into goals, skills, or habits on the report. Many junior high schools who had adopted a two, three, or four step marking scale returned to the A, B, C, D, F marking. Reporting to parents was viewed by teachers and administrators as becoming too complicated and laborious. There was no evidence that simplicity in reporting procedures was present.⁸⁰

Traxler stated that, in general, although many current articles of educational literature lead one to believe that newer practices in reporting procedures were accepted by many school systems, many schools were still very conservative and used traditional forms. Many schools that have tried the pass-fail, satisfactory-unsatisfactory, and the mastery-failure to master type of plan have not found it completely satisfactory. Three main arguments against using marks have their origin in research, logical inference supported by human experience, and mental hygiene. Some people have become convinced of an inherent badness in all marking systems. In the 1950's more attention was placed

⁸⁰R. M. Roelfs, "Trends in Junior High School Progress Reporting," Journal of Educational Research, XLIX (December, 1955), 243-248.

on the educational objectives which were to serve as a basis for the reporting system. A flexible reporting plan developed by Traxler for schools not wishing to abandon report cards was to have either each teacher or each department construct the report forms to be used. The teacher or departmental objectives could be included on the report forms.⁸¹

Parents and teachers preferred the traditional five-letter system of marking with an included carefully planned checklist of individual characteristics showing social, mental, emotional, and physical growth, according to Cagle.⁸²

Morris stated that little progress had been made in reporting to parents since the days of 1840, when colored slips of paper were given out by the teacher to show approval or blame. All report card forms were made to fit past educational practices because they complied with authoritarian systems, overstressed the adherence to a textbook, and required the ability to memorize. The parent-teacher conference was advocated as a replacement to report cards.⁸³

The traditional report card (using percentages) was viewed by Bolmeier as a convenient, antiquated device used for "separating the sheep from the goats" and not favored by many intelligent persons.

⁸¹Arthur E. Traxler, Techniques of Guidance (New York: Harper and Brothers, Publishers, 1957), pp. 233-247.

⁸²Cagle, op. cit., p. 27.

⁸³Morris, op. cit., p. 149.

Relative strengths, weaknesses, aptitudes, interests, and study habits of pupils should be indicated on the report card in order to serve as a basis for counsel. The single-mark system was unreliable and inadequate in revealing relative accomplishments of the pupil. It did not provide the essential information needed for effective guidance. Alphabetic marks were interpreted differently by parents, pupils, and teachers. Occasional modifications of a reporting system were advocated.⁸⁴

In a research study conducted to study the relation of the personality traits to school success, Russell and Thalman pointed out that from continued failure in school serious and permanent damage to a pupil's personality can result. They stated:

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 on the 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.⁸⁵

In an investigation to determine whether teachers tended to favor one sex in giving marks and whether the sex-favored tended to be determined by the teacher's sex, Carter concluded that regardless of male or female teachers, boys were penalized in the marks assigned, although not as much by a male teacher. The data also indicated that

⁸⁴Bolmeier, op. cit., pp. 16-24.

⁸⁵Ivan L. Russell and Wellington A. Thalman, "Personality: Does It Influence Teachers' Marks?" Journal of Educational Research, XLVIII (April, 1955), 564.

if marks were to reflect true achievement, then a refining of marks was necessary. Carter also concluded that intelligence was a factor in marks assigned by teachers. Thus, marks reflected not only achievement but intelligence as well.⁸⁶

Hadley, in a research study, reported that the results showed evidence of a tendency that the most-liked pupils were marked higher than their measured achievement would justify. Fifty percent of the most-liked pupils in the sample received higher marks while fifty percent of the least-liked received lower marks than actual attainment. The pupils who were neither most-liked nor least-liked had an even chance of being marked too high or too low. However, Hadley believed that most teachers mark as objectively as possible.⁸⁷

1960-Present

Austin noted that the earliest report cards which were devised in the 1800's reflected the educational philosophy of that period. A philosophy which stressed the subject-matter rather than the learner was emphasized. The purpose of these report cards was to inform the pupil and his parents of the learner's progress. Numbers or letters were utilized to rate pupil achievement. The traditional report cards contained several deficiencies. No description of the basis for evaluating the pupil's work was included on the traditional card. Grades

⁸⁶Robert S. Carter, "How Invalid Are Marks Assigned by Teachers?" Journal of Educational Psychology, XLIII (April, 1952), 227.

⁸⁷S. Trevor Hadley, "A School Mark - Fact or Fancy," Educational Administration and Supervision, XL (May, 1954), 308-312.

were subjected to many misinterpretations because they were not objectively determined. Thus, parents, having no real basis for understanding the grading system frequently became antagonistic toward the child and the school. Parents and children developed poor attitudes. The value of learning and education was ignored because the focus was upon marks. Pupils worked either to "get by" or to obtain good grades. For some well-meaning parents, the report card became a lethal weapon. Cramming and cheating resulted. Many parents withdrew their reassurance and love from their children when their grades were low. Confusions arose and frequently centered around these questions:

Did the grade relate to the student's potential or to the norm for the class? Did it indicate the pupil's present standing or his progress since a previous report? Did a high mark in a low ability group mean the same grade in a high ability group?⁸⁸

Burton explained that the traditional report card, developed during the 1800's, also reflected the educational philosophy and practices of that period. Emphasis was not upon the learner, but entirely upon the academic subject. The new-type report card was a natural development because of changes in educational thinking. Great progress had been made in the 1950's. Thousands of new-type report cards were used, yet many school systems had not heard of these improved reports. Thousands of teachers were not aware of their existence. These new-type report cards included all phases of pupil growth--

⁸⁸Mary C. Austin, "Report Cards and Parents," The Reading Teacher, XVIII (May, 1965), 660-661.

intellectual, social, emotional, and physical. The percentage mark was absurd. All traditional marking systems were arbitrary and unrelated to functional learning. The symbol became the important thing and stood for facts memorized, orders followed, and nonfunctional skills performed. Cramming, cheating, catering to teacher whims and views, and even open flattery became ways of receiving desired marks regardless of learning achievement. These evils were not useful when functional learning achievements were evaluated and "marked" on the report card by means of descriptions. The passing-grade concept was ridiculous and contrary to known facts regarding growth and facts about desirable learning achievement. The passing-grade was not based upon functional learning but upon subject-matter standards. The "get by" complex resulted. The possibility of receiving high marks which did not represent real learning was strong. Teachers frequently failed to clarify desirable outcomes of learning when the "passing grade," the mark, "the course credit," "passing the college boards," and "passing the Regents" became the real aims for many.⁸⁹

Confusion in building and interpreting report cards resulted from confusion regarding the definitions of success according to Hanson. Of the two kinds of success--"success in competition" and "fulfillment of self," parents stressed success in competition. Both kinds of success must be important in educational programs.⁹⁰

⁸⁹Burton, op. cit., pp. 118-522.

⁹⁰Earl H. Hanson, "What is Success and How Should We Report to Parents?" Education, LXXXII (October, 1961), 126.

Rolf reported that emphasis had been placed on personal competition in recent years, and competition with others had been abolished. According to him, this was not theoretically sound because a pupil's ego was falsely inflated by not informing him of his limitations and weaknesses in a competitive world. The type of competition which a pupil faced had its limitations. If every pupil attained his potential, then the educational purposes would be accomplished.⁹¹

A consensus among principals regarding reporting methods was summarized by Varner. Principals agreed that the report techniques must be clear and simple, faculty members must thoroughly understand the reporting procedures, reporting practices must not be too time-consuming, and scholastic achievement and other characteristics must be evaluated and reported. Further findings indicated that parent-teacher conferences were effective when relief time was utilized and that the interests of the pupils were the all-important factor in the administration of evaluating and reporting methods.⁹²

Cutler observed that some schools have gone "full circle" on updating and revising pupil report cards, but they have returned to earlier report card formats. A prime aim of the report card was total understandability. Some schools realized that their new report cards were refined

⁹¹Rolf, loc. cit.

⁹²Glen F. Varner, "How Do We Best Appraise and Report Individual Pupil Progress in the Junior High School?" The Bulletin of the National Association of Secondary-School Principals, XLVI (October, 1962), 23.

to the extent that total understandability was lost. Earlier forms served their needs with maximum understanding and minimum confusion.⁹³

Anastasiow stated that report cards were still with us and tended to be very controversial. The report card was utilized by school districts to communicate a vast amount of complex data regarding pupil progress to parents and to pupils. It was only one means of evaluation. Too many schools used it as the only means of evaluation. The crucial question was concerned with the real purpose of the report card. Teachers became frustrated when they tried to use the report card because no knowledge regarding the scoring of the reporting categories was available. Anastasiow conducted a three-year study in order to evaluate an existing report card, to guide the development of a new report card, and to evaluate the new card. The criteria used in the evaluation were the following:

1. Teachers. Teachers must be able to use the report card to communicate the data they have accumulated to parents. The categories included in the reporting device must be flexible enough to allow the teacher to present information that is an accurate appraisal of the student at all ability levels. The report card must not require judgments beyond the information possessed by the teacher.
2. Parents. A report card should convey clearly to parents what it is a teacher wants to report about a child's progress and the child's current standing in both academic and social areas. The parent should be able to interpret the card in the manner intended by the teacher and school district. A report card is essentially a progress report to parents and, therefore, must be easily understood.
3. Pupils. Evaluation implies the wish on the part of the evaluator to encourage improvement, or the consolidation

⁹³Cutler, op. cit., p. 60.

and continuation of desirable habits and attitudes. While the report card is not intended as a learning device to motivate children, it should not, however, have a negative effect on future achievements.⁹⁴

No all-purpose, perfect reporting system existed because pupils were not alike. Reports were concerned with intangibles such as character and attitudes as well as easy-to-measure arithmetic skills, and educational programs differed from community to community. It would have been simple to adopt a universal reporting system if pupils were alike, if educational programs were concerned with limited factual knowledge, and if all schools were identical.⁹⁵

Kingston and Wash pointed out that only a few scientific research studies on grading and reporting pupil progress had been reported even though there had been vital and continuous interest among educators. Most of the published articles reflected professional and personal bias and opinion. Only a few described the grading and reporting of pupil progress procedures used in a specific school. The research studies dealing with the effect of grading practices on motivation or on pupil achievement were generally on a high school or college level. Trying to apply these findings to the elementary school was extremely difficult. Parents were more satisfied with most of the present reporting systems than educators realized. Schools welcomed parental assistance in devising the reporting system not so much for seeking solutions to the technical

⁹⁴Nicholas J. Anastasiow, "Research and Development of an Elementary School Report Card," California Journal of Educational Research, XVII (November, 1966), 209-210.

⁹⁵"Report on Reports," (Adapted from NEA leaflet), NEA Journal, LII (December, 1963), 14.

problems of report cards, but for improving public relations. Teacher variations in marking had been widely discussed. Improvement in the reliability and validity of teacher marking could be accomplished through continuous and systematic effort by faculties under the leadership and guidance of principals. Reviewing teachers' reports and aiding them to improve their marking practices was an important and continuing responsibility for each principal.⁹⁶

It was believed by Johnson that it made little difference whether marks were in the form of letters, numbers or even geometric figures provided that their meaning was understood. In order to lessen the "mess in marking" and to clarify some of the confusion about marks, schools must have a formulated policy stating the purposes of the marks, the sources of evidence used, the basis of comparison, the attributes to be marked, and the curricular reference of the marks.⁹⁷

Grades were viewed by Doak as deterrents to learning. Therefore, the primary task for educators was to discover how to replace the false emphasis on grades with a meaningful search for knowledge.⁹⁸

Grades were a threat, caused anxiety and defensiveness, and were not intrinsically related to learning, observed Cummins. Learning was

⁹⁶Albert J. Kingston and James A. Wash, Jr., "Research on Reporting Systems," The National Elementary Principal, XLV (May, 1966), 36-40.

⁹⁷Mauritz Johnson, Jr., "Solving the Mess in Marks," New York State Education, XLIX (November, 1961), 30.

⁹⁸E. Dale Doak, "Grading - A Deterrent to Learning," The Clearing House, XXXVII (December, 1962), 247.

directed toward positive growth and development; grades were not. Current indignation was against widespread cheating. Research validated the influence grades had upon the social relationships and emotional health of the pupils. The temptation was strong to "damn grading." Because grades would be with us at the present, Cummins suggested that grades be abolished gradually by beginning now to eliminate grades in courses such as PE activity courses, laboratory courses, and student teaching. Grades that would be given in other courses should be placed in proper perspective in the total evaluation process. To involve pupils in the selection of the criteria to be used as a basis for grading and to determine the objectives of the course were goals which educators believed might help to achieve a proper perspective.⁹⁹

The traditional A-B-C-D-E report card must be abolished and replaced by the parent conference maintained Nicholson. In a questionnaire-type study undertaken in Ohio, she discovered that teaching and non-teaching parents did not want to abolish the traditional report card. Sixty-one parents out of a hundred wanted parent conferences but only in addition to the traditional report card. Parents cling to the traditional five-letter system because they grew up with it while teachers cling to it because they had not received adequate training in conference techniques.¹⁰⁰

⁹⁹Cummins, op. cit., p. 403.

¹⁰⁰Mary Louise E. Nicholson, "Reporting to Parents," Ohio Schools, XXXIX (November, 1961), 23.

Wilson advocated the abolishment of report cards and grades because they were so meaningless. The A does not tell us

How good is the class, the school, or the teacher? It does not tell us whether the student is performing up to his ability; whether he is creative, imaginative or industrious; whether he is developing his capacity to think; or whether he is acquiring a genuine desire to learn. It tells us only that compared with others in his class, the student possesses a good memory. Conversely, the C, D, or F tells us that the student has a comparatively poor memory - that and little more.¹⁰¹

Wilson stated further that because many students were unable to parrot facts and figures on examinations, society was being deprived of many otherwise competent achievers. The root of the educational sickness was not the examination per se, but its misuse. Examinations and their resultant grades were used as weapons of punishment and discipline, and as a measurement of achievement. They should have been used for measurements of teaching and learning. The importance of the examination and a good memory was not to be minimized. Wilson stated:

What is so patently evil about the emphasis on both in our American grading system is the fraudulent claim that we teach our students to think, reason, and learn - indeed, that we teach them at all. With the examination as our obstacle course and our grade as the reward, we smugly assure ourselves that we have done our duty to all - where in fact we may have done nothing for anyone.¹⁰²

In characterizing a good report, Crosby declared:

. . .if a report helps a child, his parents, and his teachers learn more about what he is capable of, where he is reaching his potential, what is being done to

¹⁰¹Charles H. Wilson, "Educational Innovation: Are Public Schools Going Overboard?" Nation's Schools, LXXX (November, 1967), 66-67.

¹⁰²Ibid.

foster his learning, and perhaps most significant of all, if it provides 'that something to grow on,' it is a good report.¹⁰³

In 1966, the NEA Research Division sent a questionnaire on administrative practices to school systems. The respondents consisted of 232 systems with enrollments from 300 to 2,999 pupils, 234 systems had 3,000 to 24,999 pupils, and 150 systems had 25,000 or more pupils. Results showed that the traditional A-F type of report card was used in the majority of school systems. School systems with 300 or more pupils reported that eight in ten school systems used some type of classified scale such as the A-F or the 1-5 for senior high school students. The figure was slightly higher for junior high school pupils. Between eight and nine in ten systems used it for fourth graders. These traditional report cards usually listed the subjects and provided space for the letters A-F, the numbers 1-%, or percentages which indicated the amount or the quality of the pupil's achievement. School systems not using the traditional cards used the simplified version of the classified scale called the two-part scale of pass-fail or satisfactory-unsatisfactory, the written description method, or the conference method. All four types had limitations, but the traditional report card had been the most severely criticized. Some school systems supplemented the report card with letters or conferences.¹⁰⁴

¹⁰³Muriel E. Crosby, "Good Report," NEA Journal, LI (April, 1962), 47.

¹⁰⁴Anonymous, "Reports to Parents," NEA Research Bulletin, XLV (May, 1967), 51-52.

Chansky conducted a research study to determine whether teachers believed that various reporting forms of the elementary school were equally effective in giving functional information of pupil growth. Twenty-five report cards used in New York were judged by seventy teachers. The number of entries on the report cards proved to be significant. Report cards with many entries reflected progress in many facets of learning. The more pupil characteristics listed on the report card, the higher was its rating. The range of pupil behavior covered by the report card was significant. Higher rated cards dealt with subject-matter, health, social adjustment, and work habits, while lower rated cards dealt only with subject-matter. The highest ratings went to the cards which stimulated teachers to observe the "whole child." Another important factor was the system used in informing parents of pupil progress. The lower rated cards used the numerical or letter systems of grading while higher rated cards used positive adjectives entirely or in part. The highest ratings went to the card which stimulated teachers to describe pupil achievement.¹⁰⁵

Halliwell and Robitaille undertook a research study to determine whether teachers employed in school systems using the dual reporting program actually evaluated pupil progress according to the

¹⁰⁵Norman M. Chansky, "Elementary School Teachers Rate Report Cards," The Journal of Educational Research, LVI (July-August, 1963), 523-528.

dual reporting theory, and to ascertain the relationship between the subjective and objective report card grades given by sixth-grade teachers. The finding showed that the teachers graded the pupils in the traditional manner on the individualized part of the dual marking report card. The stated philosophy of the individualized section was not utilized. There was a positive relationship between the scores on the objective and subjective parts of the dual marking system of reporting. The halo effect was apparent. Bright pupils were rewarded twice with good grades. Slow pupils were punished twice with poor grades.¹⁰⁶

Coogan reported that parental concern regarding academic achievement had increased. Adequate reporting systems giving parents an understandable and accurate view of the academic achievement as well as the academic potential of their child was important. Parents demanded a reporting system that presented their child's achievement in relation to himself and in relation to his classmates. Some schools returned to the "A-B-C" method. The dual-marking system became a trend. It must be noted, however, that parents were not necessarily better-informed because of the changes in reporting procedures. The reporting methods must be evaluated by the degree of parental understanding conveyed by the report.¹⁰⁷

¹⁰⁶Joseph W. Halliwell and Joseph P. Robitaille, "The Relationship Between Theory and Practice in a Dual Reporting Program," The Journal of Educational Research, LVII (November, 1963), 137-141.

¹⁰⁷Dale Roger Coogan, "Reporting Pupil Progress: An Analysis of the Relationship Between the Administrative Methods Used for Change and the Degree of Parental Understanding" (unpublished Doctor's Dissertation, The University of Southern California, 1963), pp. 1-3.

The research study of Coogan was undertaken in order to examine the degree of parental understanding of pupil progress. Only school districts that had made recent major changes in the elementary school reporting system and had met the established criteria for adequacy were used in the study. Following the launching of Sputnik, parents demanded that the reporting system give a comparative view of their child in relation to his classmates. Political and legislative events affected the emphasis on subject matter and pupil achievement in this subject matter. Most school districts had not abolished the report card because it, though viewed as both good and bad, was accepted and expected by parents.¹⁰⁸

Coogan, in summarizing the findings from the literature, stated that the history of reporting to parents revealed that parents and teachers should work together for the educational development of the child. Increased knowledge of child growth and development had demanded it. Knowledge of individual differences led to the realization that the individual was not to be measured against a norm. Parents, however, insisted on information that served as a basis for comparison with other classmates. Parents and educators agreed that the individual parent-teacher conference was the best method of reporting basic skill progress. Surveys showed that, aside from the parent-teacher conference, the traditional A-B-C-D-F written report card was preferred by parents.¹⁰⁹

¹⁰⁸ibid., pp. 15-29.

¹⁰⁹ibid., pp. 52-53.

Coogan recommended that schools should adjust their reporting practices in order to utilize the most understandable methods. If written reports were demanded, the A-B-C-D-F marking symbols were recommended for middle and upper grade elementary pupils. His recommendations also included a plea for extensive in-service training for teachers.¹¹⁰

Cohen, in a 1965 research study which measured the effects report cards had on students receiving grades concluded that report cards affect students. Expectancies for the next semester generally increased after the students received their report cards. The importance of getting good grades generally devaluated during this same period. Whether these student reactions were harmful or beneficial remained debatable after this research study.¹¹¹

The changes or variations some individual students made concerning estimations of academic ability and future grade expectancies were related to receiving report cards. Some even distorted their memory in the directions of their original estimations. Original grade expectancies did not affect the changes in students' attitudes towards grades and school nor their general estimations of themselves. Thus, expectancy for future grades, memory of grades, and estimation of the highest grades the students think they could have earned (all

¹¹⁰ibid., pp. 277-278.

¹¹¹Karen Sue Cotzin Cohen, "Reactions of Students to Receiving Grades" (unpublished Doctor's Dissertation, Johns Hopkins University, Baltimore, 1965), p. 56.

were items most closely related to grades) were influenced by the particular grades students received on their report card. General attitudes towards grades and school (items less closely related to grades) were not affected by the particular grade received.¹¹²

Jensen's 1966 research study revealed the effects of grades on learning. Almost no research had been undertaken to determine the effects that low grades had on classroom learning. Grades were a powerful stimuli for they influenced the academic, personal, social, and vocational life of a student. General evaluations of the pupils were determined from the grades received on quizzes, projects, daily tests, and homework. Teachers would benefit from the knowledge of the specific effects of grades, but the lack of research prevented a real understanding of grades. The immediate result of low grades may produce less learning yet many teachers were advised to grade low in order to increase learning.¹¹³

Regarding the immediate effects of grades on learning in a lecture-type classroom, Jensen concluded:

1. Grades were followed by less cognitive learning among both male and females.
2. Grades were followed by less affective learning among females.¹¹⁴

¹¹²ibid., pp. 57-58.

¹¹³Larry Cyril Jensen, "The Effects of Grades on Cognitive and Affective Learning" (unpublished Doctor's Dissertation, Michigan State University, Lansing, 1966), pp. 1-3.

¹¹⁴ibid., pp. 54-55.

A study reported in 1968 was undertaken by Edmunds¹¹⁵ to investigate the relationship between the general self image of students and the report card marks achieved by these students in eleventh year English classes.¹¹⁵

Although much literature supported the belief that a positive relationship existed between the self-concept and academic achievement, certain findings by Edmunds refuted this belief.¹¹⁶

Certain doubts had been cast upon the validity of the self-report in education by the findings of Edmunds regarding the relationship between a high school student's report card achievement and his self image.¹¹⁷

In examining the findings from the literature of the past sixty years, the writer observed that research had not been undertaken to determine the effects that report card forms had on pupil achievement. This lack of research was surprising. The writer drew the following conclusions based on the literature findings:

1. Of all the phases of American education, reporting had made the least advancement.
2. Although the traditional report cards received the severest criticism, it remained the most widely used form of reporting in the schools.
3. The report card had not been abolished in many schools because parents demanded it.

¹¹⁵Brehaut R. Edmunds, "The Relationship Between the Self Image and the Report Card Achievement of Eleventh Year English Students" (unpublished Doctor's Dissertation, The University of Southern California, Los Angeles, 1968), p. 1.

¹¹⁶ibid., p. 30.

¹¹⁷ibid., p. 82.

4. There was a demand for a report card in written form. Even advocates of the informal parent-teacher conference method recommended a written duplicate copy of comments made during the conference because parents preferred a written form to serve as a reference in the home.
5. Parents and teachers clung to the traditional system of reporting because of their familiarity with the age-old system.
6. A report card providing information of pupil strengths and weaknesses received attention, but was time-consuming.
7. Discussion of teacher variations in marking remained widespread.
8. Report cards tended to be controversial.

CHAPTER III

DESIGN AND PROCEDURE

SAMPLING

The object of this research was to investigate whether pupil achievement differed because of the type of report card received. For this purpose the pupils used were tested prior to and following the experiment to determine the effect report card forms had on achievement.

The evidence gleaned from the research literature and from the writer's experiences suggested the desirability of such an experiment. Because the writer was aware of other factors in the classroom that influenced achievement, the experiment was adapted to a specific basis by confining the inquiry to the report card form factor.

During the period of the experiment the two groups received exactly the same treatment and performed the same tasks at the same time of day. One exception was evident--the control group received the traditional report card and only letter grades on their papers while the experimental group received the diagnostic report card and written teacher remarks on their papers.

Selection of School and Pupils

The sampling for the present study was selected from the sixth grade class of Seven Dolors Grade School, Manhattan, Kansas. At the time of the study, Seven Dolors Grade School, a Catholic parochial

school, had an enrollment of three hundred and seventy-five pupils. A Superintendent, a Principal, ten Sisters, and nine lay teachers formed a competent team of faculty. With Standard accreditation by the State Department of Education, Seven Dolors has served the Manhattan area for sixty-two years. Manhattan is a mid-western college town of approximately 27,200 plus an additional 13,000 university students.

Equating the Groups

The parallel-group design was used in which an attempt was made to match pupils from the sixth grade for a control group and an experimental group within the limits defined by the study. Sex, intelligence, and achievement scores were factors considered in the matching process. After obtaining a pair of matched subjects, the writer then randomly selected one from each pair for the control group or the experimental group. Matching was preferred to simple random sampling because more comparability between groups was achieved on the matching variables.

Table I contains pre-test grade scores, composites, totals, and means resulting from the use of the Stanford Achievement Test - Form X, relevant to the control group (fourteen pupils); while Table II provides the same information relevant to the experimental group (fourteen pupils). The t-test of the control and experimental pre-test grade score means was non-significant at the .05 level ($t = .538$).

Table III contains information relevant to the control group (fourteen pupils) while Table IV provides necessary information regarding the experimental group (fourteen pupils). Code numbers, sex, intelligence

TABLE I

CONTROL GROUP PRE-TEST SCORES
(STANFORD - FORM X)

Code number	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
1A	10.0	10.6	11.5	10.8	6.6	7.8	8.0	9.0	6.7	9.0
2A	7.3	7.3	8.5	6.5	7.9	7.0	8.6	7.9	6.6	7.5
3A	7.8	9.2	6.0	7.5	7.7	7.0	8.3	10.0	6.9	7.8
4A	7.5	7.8	7.5	7.3	7.9	8.2	7.4	8.3	7.2	7.7
5A	7.5	8.2	6.3	6.6	5.9	6.5	9.1	6.8	10.0	7.4
6A	5.4	6.6	6.8	4.5	9.6	6.8	7.4	6.4	6.7	6.7
7A	8.0	7.2	6.8	5.5	5.0	6.3	4.9	5.0	8.1	5.4
8A	8.0	8.2	7.6	5.7	7.7	6.5	6.1	6.3	7.2	6.1
9A	6.2	6.4	7.5	4.8	5.8	5.9	5.4	5.4	5.8	5.9
10A	5.9	5.3	6.4	5.8	5.6	6.3	5.6	5.9	4.1	5.7
11A	7.5	7.0	4.8	6.2	5.8	5.6	4.9	5.4	4.6	5.7
12A	5.9	6.2	4.5	4.8	8.2	6.8	6.3	6.3	4.4	5.9
13A	7.6	7.5	9.2	6.1	6.6	6.3	5.4	5.6	7.2	6.8
14A	6.2	6.9	3.6	5.2	3.6	4.9	5.9	7.7	5.0	5.6
TOTALS	100.8	104.4	97.0	87.3	93.9	91.9	93.3	96.0	90.5	93.2
MEANS	7.2	7.5	6.9	6.2	6.7	6.6	6.7	6.9	6.5	6.7

TABLE II

EXPERIMENTAL GROUP PRE-TEST GRADE SCORES
(STANFORD - FORM X)

Code number	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
1B	8.0	8.0	6.8	9.8	7.7	8.8	7.7	7.4	6.9	7.9
2B	7.8	9.6	8.2	10.5	7.7	7.6	7.4	9.0	6.2	8.2
3B	8.3	5.7	10.5	10.4	7.1	8.2	9.6	8.3	8.1	8.5
4B	7.3	8.2	8.2	7.7	7.4	7.8	6.8	7.6	7.5	7.6
5B	7.5	6.5	6.8	4.5	8.4	7.3	8.6	5.8	9.2	7.1
6B	7.3	8.2	6.4	5.1	10.0	7.8	8.3	7.0	7.5	7.5
7B	6.6	5.3	4.8	3.9	5.4	5.6	5.7	5.4	5.7	5.5
8B	6.9	6.9	8.8	8.0	6.3	6.8	6.6	6.6	6.0	7.0
9B	6.9	5.7	6.0	4.5	5.8	5.4	4.4	5.9	6.6	5.7
10B	6.7	6.7	6.7	5.9	6.8	6.5	4.6	4.5	5.2	5.9
11B	6.0	6.6	6.0	3.4	6.6	5.6	7.1	6.5	6.4	6.0
12B	5.7	6.7	6.2	5.2	6.3	5.4	5.2	4.2	6.0	5.7
13B	7.8	7.0	7.3	7.5	7.1	6.3	5.9	6.2	6.4	7.8
14B	5.2	4.9	3.7	3.7	5.0	6.3	6.5	4.6	4.3	4.9
TOTALS	98.0	96.0	96.4	90.1	97.6	95.4	94.4	89.0	92.0	95.3
MEANS	7.0	6.9	6.9	6.4	7.0	6.8	6.7	6.4	6.6	6.8

TABLE III

DESCRIPTION OF THE CONTROL GROUP

Code number	Sex	Kuhlmann-Anderson I.Q. Test - 1969	1969 Stanford Achievement Mean	1968 Metropolitan Achievement Mean	Stanford-Metropolitan Mean	Fathers' occupation
1A	F	132	9.0	7.9	8.5	Professional
2A	F	128	7.5	7.6	7.6	Semi-Prof.
3A	F	123	7.8	8.5	8.2	Professional
4A	M	123	7.7	7.4	7.6	Managerial
5A	M	124	7.4	7.1	7.3	Semi-Prof.
6A	M	125	6.7	6.8	6.8	Semi-Prof.
7A	M	122	5.4	5.6	5.5	Semi-Prof.
8A	F	118	6.1	7.2	6.8	Semi-Prof.
9A	M	119	5.9	5.6	5.8	Professional
10A	F	110	5.7	6.3	6.1	Semi-Prof.
11A	M	109	5.7	6.3	6.0	Professional
12A	F	114	5.9	5.9	5.9	Professional
13A	F	103	6.8	7.1	7.0	Semi-Prof.
14A	M	98	5.6	6.4	6.0	Professional
TOTALS		1648	93.2	95.7	95.1	

scores, means of achievement scores obtained prior to the matching, means of previous year's achievement scores, means of the combined Stanford and Metropolitan Achievement scores, and fathers' occupation categorized as Professional, Managerial, or Semi-Professional were incorporated within the Tables. These items were secured from recently administered tests and from available school records.

The t-test was used to determine whether the differences between the score totals of the Kuhlmann-Anderson Intelligence Test, the Stanford Achievement Test, the Metropolitan Achievement Test, and the combined mean score totals of the Stanford and the Metropolitan Achievement Tests (Tables III and IV) was significant. The analysis showed that the difference was non-significant at the .05 level ($t = .362$). Thus, both sampling groups were considered equal.

Table V presents the pre-test grade score means, composite, difference and t-test of the control and experimental group. The t-test showed Test Two and Test Eight as significant at the .05 level while Tests One, Three, Four, Five, Six, Seven, Nine and the composite were non-significant.

Description of Pupils

Twenty-eight pupils (seven pairs of boys and seven pairs of girls) from the sixth grade of Seven Dolors Grade School, Manhattan, Kansas, during the academic year of 1969-1970 were matched on the basis of sex, intelligence and achievement scores. These pupils ranged from 98-134 in intelligence quotient figures and from 4.9-9.0 in achievement mean figures. Personality traits varied--the introverts, the extroverts;

TABLE IV

DESCRIPTION OF THE EXPERIMENTAL GROUP

Code number	Sex	Kuhlmann-Anderson I.Q. Test - 1969	1969 Stanford Achievement Mean	1968 Metropolitan Achievement Mean	Stanford-Metropolitan Mean	Fathers' occupation
1B	F	134	7.9	7.6	7.8	Professional
2B	F	127	8.2	8.5	8.4	Semi-Prof.
3B	F	126	8.5	7.3	7.9	Managerial
4B	M	122	7.6	7.5	7.6	Professional
5B	M	121	7.1	7.3	7.2	Professional
6B	M	124	7.5	6.2	6.9	Semi-Prof.
7B	M	123	5.5	5.4	5.5	Semi-Prof.
8B	F	114	7.0	5.9	6.5	Deceased
9B	M	116	5.7	5.2	5.5	Professional
10B	F	113	5.9	6.2	6.1	Managerial
11B	M	113	6.0	6.1	6.0	Semi-Prof.
12B	F	112	5.7	6.1	5.9	Professional
13B	F	104	7.8	6.9	7.4	Professional
14B	M	<u>98</u>	<u>4.9</u>	<u>8.9</u>	<u>6.9</u>	Semi-Prof.
TOTALS		1647	95.3	95.1	95.6	.

TABLE V
PRE TEST GRADE SCORE MEANS OF CONTROL AND EXPERIMENTAL GROUP
(STANFORD - FORM X)

Group	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
Control	7.2	7.5	6.9	6.2	6.7	6.6	6.7	6.9	6.5	6.7
Experimental	<u>7.0</u>	<u>6.9</u>	<u>6.9</u>	<u>6.4</u>	<u>7.0</u>	<u>6.8</u>	<u>6.7</u>	<u>6.4</u>	<u>6.6</u>	<u>6.8</u>
Difference	.2	.6	.0	-.2	-.3	-.2	.0	.5	-.1	-.1
t-test	.990	3.410*	0.	-.990	-.939	-.990	0.	2.546*	-.990	-.990

* significant difference.

the carefree, happy-go-lucky, and the serious; the independent, the dependent; and the motivated, the non-motivated. However, all possessed scholastic potential.

The following daily program of studies was utilized by all the pupils.

Daily Program of Studies

8:30 - 9:30	Reading	M-T-W-Th-F
9:30 - 10:00	English	M-T-W-Th-F
10:00 - 10:15	Recess	M-T-W-Th-F
10:15 - 11:00	Math	M-T-W-Th-F
11:00 - 11:35	Music	M - W
11:00 - 11:35	Independent Study	T Th-F
11:35 - 11:50	Free	M-T-W-Th-F
11:50 - 12:20	Lunch Period	M-T-W-Th-F
12:20 - 1:00	Social Studies	M-T-W-Th-F
1:00 - 1:45	Physical Education	M-T-W-Th-F
1:45 - 2:10	Ethics	M-T-W- -F
2:15 - 2:45	Science	M-T-W-Th
2:45 - 3:00	Spelling	M-T-W-Th
2:15 - 3:15	Art	F
3:00 - 3:25	Study	M-T-W-Th

MEASURING DEVICES

The Stanford Achievement Test, Form X (Intermediate II Battery), and the Stanford Achievement Test, Form W (Intermediate II Battery), standardized and nationally known, reliable and valid, were used to measure pupil achievement. The Intermediate II Battery, designed primarily for use from the middle of Grade 5 to the end of Grade 6, was comprised of nine categories: (1) Word Meaning, (2) Paragraph Meaning, (3) Spelling, (4) Language, (5) Arithmetic Computation, (6) Arithmetic Concepts, (7) Arithmetic Applications, (8) Social Studies, and (9) Science.

According to Buros the Stanford Test is recommended for various uses

...in the analysis of group differences among subjects and also of the differences in the abilities of individual pupils in the various subjects for purposes of planning individualized instruction, grouping pupils for instructional purposes, determining and evaluating rate of progress, and evaluating achievement.¹

Although the tests were time-orientated, the time limits were sufficient to allow practically all pupils to complete known questions. The tests are "fundamentally power tests and not speed tests."²

¹Oscar Krisen Buros (ed.), The Sixth Mental Measurement Year Book (Highland Park, New Jersey: The Gryphon Press, 1965), 26.

²Truman L. Kelley and others, Stanford Achievement Test, Directions for Administering (New York: Harcourt, Brace and World, Inc., 1964), 4.

The Kuhlmann-Anderson Intelligence Test (Seventh Edition), Form EF, was used to measure native ability (intelligence quotient, I.Q.). The EF Form is designed for grades 5-6-7.

Michael, in reviewing the Kuhlmann-Anderson Test for The Sixth Mental Measurement Yearbook, found little to criticize either on the adequacy of the interpretative data furnished or the standardization and construction of the test. In many research studies it is desirable to have a quantitative (Q) score as well as a verbal (V) score; both scores may be derived from this test.³

The seventh edition of this test, according to Michael, "introduces several new features that are in agreement with a great deal of modern thinking about the assessment of intelligence."⁴

Pidgeon, in reviewing this test, stated that the seventh edition "incorporates most of the desirable features that the user of modern test requires."⁵ The estimates of the reliability for the total score (after factor analyses of subtests were made) ranged from .85 to .95. The coefficients quoted in the test manuals are satisfactory. Even with two grades between testing, the test-retest coefficients ranged from .83 to .92. Adjacent form tests produced correlations from .77 to .89.⁶

³Buros, op. cit., p. 735.

⁴Ibid.

⁵Ibid., p. 738.

⁶Ibid.

The present study required two types of pupil progress reports for the control group and the experimental group, respectively, to determine whether pupil achievement differed because of the type of report card received. Thus, the traditional report card, the Pupil Progress Report, (Appendix A), was given to the pupils of the control group. This progress report was currently used in grades four through eight throughout the Diocesan school systems of Salina, Kansas. The progress report card incorporated the Scholastic Record as well as Personality and Attendance Records. Because the school filed duplicate copies of the quarterly report, parents could keep the Progress report and use it as a reference in guiding their children.

The Diagnostic Report Card (Appendix B) was used with pupils in the experimental group. The diagnostic report was developed by the writer and interested associates. The purpose of the diagnostic report of pupil progress was to serve as an aid to evaluate the strengths and/or weaknesses of the pupil in the various subject areas and to promote understanding and cooperation between the home and the school. No traditional letter grades were used, but rather comments regarding the strengths and weaknesses of each pupil as observed by the teacher were recorded.

PROCEDURE

Letter to Parents (Appendix C) was mailed prior to the testing which preceded the matching of the pupils.

The procedure for this parallel-group experimental research on report card forms affecting achievement was conducted for a period of twenty-seven weeks (three nine-week marking periods) beginning in August, 1969, and terminating in March, 1970. The Kuhlmann-Anderson Test was administered to serve as an aid to equate the experimental and control groups. The Stanford Achievement Test, Form X, was also given as a pre-test in May before the study commenced as an aid for matching the sample and as a means of measuring achievement prior to the study. After the study, the Stanford Achievement Test, Form W, was administered as a post-test to determine whether report card forms would affect achievement. Separate forms of the Stanford Achievement Test were administered as a reliability measure.

The pupils in both groups were taught simultaneously by the experimenter who also administered all tests and scored them manually.

The traditional report card was given to the control group in November, January, and March while the experimental group received the diagnostic report card in the same sequence of time.

During the experimental period the techniques for securing the desired information included the following:

1. A record was maintained for each pupil of the control group, showing only the letter grades.
2. A record was kept for each pupil of the experimental group, noticing strengths and weaknesses.
3. The control group received only grades. All teacher comments regarding strengths and weaknesses were excluded.

4. Teacher remarks for the experimental group include such comments on pupil papers as "Jack, division of decimals is one of your strong points," or "Mary, you need to improve in multiplication of fractions."

CHAPTER IV

REPORT OF FINDINGS

ANALYSIS TECHNIQUES

This study was designed to investigate whether pupil achievement differed because of the type of report card used. The following section reports results of the statistical analysis utilized in making this comparison. Grade scores, composites, totals, means, and t-tests were used in making this analysis.

The test of the hypothesis--that there is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards--lies in the comparison of the performance of its two groups of individuals who differed only in the report card received quarterly and the type of comments made by the teacher on the paper of the pupil.

To determine the acceptance or rejection of the null hypothesis at the .05 level of confidence, the probability values in Fisher's special table of t for small sample were consulted.¹

DESCRIPTION OF FINDINGS

Table VI shows the post-test mean scores (Stanford Achievement - Form W), the composite, the difference of the means, and the t-test

¹G. Melton Smith, Ph.D., A Simplified Guide to Statistics (New York: Rinehart and Company, Inc., 1946), p. 69.

TABLE VI

POST-TEST GRADE SCORE MEANS OF CONTROL AND EXPERIMENTAL GROUP
(STANFORD - FORM W)

Group	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
Control	6.9	6.3	7.8	6.9	6.1	7.0	7.9	7.3	8.9	7.5
Experimental	<u>7.0</u>	<u>7.7</u>	<u>7.9</u>	<u>8.0</u>	<u>6.5</u>	<u>7.2</u>	<u>7.6</u>	<u>7.6</u>	<u>7.2</u>	<u>7.4</u>
Difference	-.1	-1.4	-.1	-1.1	-.4	-.2	.3	-.3	1.7	.1
t-test	-.990	-1.000	-.990	-1.003	-1.022	-.990	.939	-.939	.998	.990

All non-significant.

results of the control and experimental group. The means ranged from 6.1 to 8.9 for the control group and 6.5 to 8.0 for the experimental group. The difference ranged from -1.4 in Test 2 to 1.7 in Test 9 and the t ranged from -1.022 to 1.000. The composite difference was .1 and the $t = .990$. The null hypothesis was sustained for each of the nine tests and composite score at the .05 level of confidence.

Table VII presents the mean gains of the control and experimental group following the twenty-seven weeks of experimentation. The control group made gains in six of the nine tests while the experimental group made gains in seven of the nine tests. The control group composite gain was .8 while the experimental group composite gain was .6.

Tables VIII and IX list the Stanford Achievement-Form W Grade Scores of Tests 1-9 (Test 1, Word Meaning; Test 2, Paragraph Meaning; Test 3, Spelling; Test 4, Language; Test 5, Arithmetic Computation; Test 6, Arithmetic Concepts; Test 7, Arithmetic Applications; Test 8, Social Studies; and Test 9, Science). The composite scores, the totals and the means of the control and experimental group are also listed on the tables.

The t -test was used to determine whether the difference between the post-test means of the nine sub-tests of the control group (Table VIII) and the post-test means of the nine sub-tests of the experimental group (Table IX) was significant. The analysis showed that the difference was non-significant at the .05 level ($t = .544$). Thus, the null hypothesis was sustained for each of the nine post-sub-tests of the control and experimental group at the .05 level of confidence.

TABLE VII

MEAN GAINS OF CONTROL AND EXPERIMENTAL GROUP
(Grade Score Means of the Stanford Achievement - Form W - Post-Test of 1970)
(Grade Score Means of the Stanford Achievement - Form X - Pre-Test of 1969)

Group	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
A 1970	6.9	6.3	7.8	6.9	6.1	7.0	7.9	7.3	8.9	7.5
A 1969	<u>7.2</u>	<u>7.5</u>	<u>6.9</u>	<u>6.2</u>	<u>6.7</u>	<u>6.6</u>	<u>6.7</u>	<u>6.9</u>	<u>6.5</u>	<u>6.7</u>
Mean gain	-.3	-1.2	.9	.7	-.6	.4	1.2	.4	2.4	.8
B 1970	7.0	7.7	7.9	8.0	6.5	7.2	7.6	7.6	7.2	7.4
B 1969	<u>7.0</u>	<u>6.9</u>	<u>6.9</u>	<u>6.4</u>	<u>7.0</u>	<u>6.8</u>	<u>6.7</u>	<u>6.4</u>	<u>6.6</u>	<u>6.8</u>
Mean gain	0.0	.8	1.0	1.6	-.5	.4	.9	1.2	.6	.6

A = Control group.

B = Experimental group.

TABLE VIII

CONTROL GROUP POST-TEST GRADE SCORES
(STANFORD - FORM W)

Code number	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
1A	8.8	10.0	12.9	10.9	8.2	8.5	11.1	9.0	10.4	10.0
2A	8.3	8.4	9.2	6.0	7.7	8.0	8.0	5.9	8.1	7.7
3A	9.5	8.4	8.0	7.7	5.8	7.6	9.6	6.4	9.6	8.1
4A	7.5	10.4	7.3	7.5	8.2	8.5	11.9	11.9	9.6	9.3
5A	8.0	7.8	7.1	9.5	6.5	8.0	8.3	7.4	10.0	8.1
6A	6.4	7.3	7.3	6.6	7.7	7.8	9.1	7.2	7.5	7.4
7A	6.9	7.5	6.7	6.0	4.4	6.6	8.3	7.6	9.6	7.1
8A	8.0	7.2	8.8	8.2	5.9	6.1	5.4	7.0	9.6	7.4
9A	6.8	6.7	7.5	3.8	5.8	5.9	6.6	7.0	9.2	6.6
10A	7.1	6.3	6.4	7.7	2.9	4.9	4.2	6.6	7.5	6.0
11A	8.0	8.2	7.5	6.1	5.9	5.9	7.7	7.4	9.9	7.4
12A	7.5	6.3	6.3	5.6	6.5	7.8	8.6	7.0	10.7	7.4
13A	7.8	6.5	10.2	7.1	6.3	6.5	6.8	5.8	7.2	7.6
14A	<u>6.6</u>	<u>6.9</u>	<u>4.5</u>	<u>3.6</u>	<u>3.7</u>	<u>5.9</u>	<u>4.9</u>	<u>6.6</u>	<u>6.2</u>	<u>5.4</u>
TOTALS	97.2	87.9	109.7	96.3	85.5	98.0	110.5	102.8	125.1	105.5
MEANS	6.9	6.3	7.8	6.9	6.1	7.0	7.9	7.3	8.9	7.5

TABLE IX
EXPERIMENTAL GROUP POST-TEST GRADE SCORES
(STANFORD - FORM W)

Code number	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
1B	7.6	8.7	7.6	9.3	9.9	8.5	8.6	7.9	7.5	8.4
2B	8.8	10.4	10.2	11.3	7.4	8.5	8.6	7.9	7.2	8.9
3B	8.8	9.6	10.5	10.8	5.9	8.5	8.6	9.0	8.1	8.9
4B	6.9	8.2	7.7	9.8	5.0	8.0	8.3	7.2	9.2	7.8
5B	6.8	8.2	7.6	7.5	8.2	7.8	11.1	9.6	10.4	8.6
6B	7.5	7.8	7.3	6.9	8.2	7.3	10.6	11.0	9.6	8.7
7B	5.6	5.7	5.3	5.5	4.6	6.5	6.6	4.6	5.7	5.6
8B	6.7	7.2	10.5	9.0	7.4	7.6	8.0	6.3	7.5	7.8
9B	7.3	6.4	7.3	4.8	4.6	5.6	5.1	5.2	7.5	6.0
10B	6.8	6.4	7.7	7.7	6.3	6.1	6.3	12.7	4.7	7.2
11B	6.6	6.8	7.7	6.9	4.4	6.5	5.9	5.9	6.6	6.4
12B	6.4	7.5	6.3	8.0	6.5	6.3	5.7	6.2	5.6	6.5
13B	6.9	7.7	9.7	10.0	5.9	6.8	6.8	5.8	6.0	7.3
14B	<u>5.2</u>	<u>7.3</u>	<u>4.7</u>	<u>4.6</u>	<u>6.2</u>	<u>6.1</u>	<u>6.8</u>	<u>6.5</u>	<u>5.8</u>	<u>5.9</u>
TOTALS	97.9	107.9	110.1	112.1	90.5	100.1	107.0	105.8	101.4	104.0
MEANS	7.0	7.7	7.9	8.0	6.5	7.2	7.6	7.6	7.2	7.4

CHAPTER V

SUMMARY AND CONCLUSIONS

PURPOSE OF THE STUDY

The purpose of this study was to investigate and obtain objective evidence on the effect the report card form has on achievement.

PROCEDURES

In order to determine whether pupil achievement differed because of the type of report card received, the pupils used were tested prior to, and following the experiment, to determine the effect report card forms had on achievement.

For the experimentation, the parallel-group design was used in which an attempt was made to match pupils from the sixth grade for a control group and an experimental group within the limits defined by the study. Sex, intelligence, and achievement scores were the factors considered in the matching process.

The sampling for the present study was selected from the sixth grade class of Seven Dolors Grade School, Manhattan, Kansas. During the period of the study experiment the two groups received exactly the same treatment and performed the same tasks at the same time of the day. One exception was evident--the control group received the traditional report card and only letter grades on their papers while

the experimental group received the diagnostic report card and written remarks of the teacher on their papers.

The study was conducted the first twenty-seven weeks of school. Report cards were given quarterly.

The data evaluated consisted of scores obtained in the Stanford Achievement Test (Form X) and the Stanford Achievement Test (Form W). Fisher's Table of t for small sample determined the non-significance of the t-test and indicated the acceptance of the null hypothesis.

SUMMARY OF THE RESULTS

The summary of findings is presented in tabular form (Table X). Test Two (Paragraph Meaning) and Test Eight (Social Studies) t-test results on the pre-test showed a significant difference favoring the control group. The same results were expected on the post-test. However, the t-tests of Test Two and Test Eight on the post-test were non-significant. Gains were not made in Test Five (Arithmetic Computation). Perhaps less stress on computation by the teacher and modern math texts used by the students could account for this decrease in computational skills. Greater emphasis was placed on Arithmetic concepts and applications by the teacher and the texts. Gains were not made in Test One (Word Meaning) by both groups nor did the control group gain in Test Two (Paragraph Meaning). Reading deficiencies perhaps become more obvious as the child matures and develops. It seems that advanced readers tend to level off their reading scores during pre-and-early adolescence. For example, pupil 1A scored a 10.0 on the pre-test word

TABLE X

SUMMARY OF TEST RESULTS FOR CONTROL AND EXPERIMENTAL GROUP
(Grade Score Means and t-test)

Group	Pre-test of Stanford Achievement Form X									
	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
Control	7.2	7.5	6.9	6.2	6.7	6.6	6.7	6.9	6.5	6.7
Experimental	7.0	6.9	6.9	6.4	7.0	6.8	6.7	6.4	6.6	6.8
t-test	.990	3.410*	.000	-.990	-.939	-.990	.000	2.546*	-.990	-.990
* Significant differences										
Group	Post-test of Stanford Achievement Form W									
	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Comp.
Control	6.9	6.3	7.8	6.9	6.1	7.0	7.9	7.3	8.9	7.5
Experimental	7.0	-7.7	7.9	8.0	6.5	7.2	7.6	7.6	7.2	7.4
t-test	-.990	-1.000	-.990	-1.003	-1.022	-.990	.939	-.939	.998	.990
All non-significant										

meaning (Test 1) and scored a 8.8 on the post-test.

Sixth grade pupils tested during the month of March should have a grade placement of 6.6. The experimental group post-test means surpassed the required grade placement in all tests except Test 5 (6.5). The control group post-test means surpassed the required grade placements of 6.6 in all tests except Test 2 (6.3) and Test 5 (6.1).

CONCLUSIONS, RECOMMENDATIONS AND IMPLICATIONS

Basically, this research was an attempt to study the effect the report card form had on achievement. This study tested the hypothesis: There is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards. Statistical analysis showed no significant difference between the control group receiving the traditional report card (letter grades) and the experimental group receiving the diagnostic report card (written descriptive analysis of pupil strength and weakness). This evidence contradicted the notion that a change in report card forms, in itself will result in higher academic achievement.

As a result of this experimentation, the writer seems justified in stating that there is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards.

The writer feels justified in placing considerable confidence in the findings of this study. Care was taken to match carefully the sampling population according to sex, I.Q. and achievement.

Some suggestions for further research and study include:

1. Development of a similar study using either a larger population sampling, exceptional children sampling, other grade levels, or longer periods of time.
2. Development of a similar study of the psychological effects of motivational drives, tensions, home pressures, or attitudes of the teacher and/or pupils toward the reporting form.

Controversy will always cloud this problem. Much more time and effort must be spent in experimentation in order to meet effectively this challenge in education. It should be noted that much verbalization occurs but very little research has been done on reporting.

Research should provide some answers to the question whether report cards will be necessary in our future educational system. Perhaps each school (staff, parents, and student representatives) needs to adapt its own reporting system to fit its own philosophy and objectives.

The writer feels that educators may never find a perfect report card form, but as Jones once stated, "We can always have the stimulation, pleasure, and profit of trying."¹

¹John A. Jones, "Grading, Marking, and Reporting in the Modern Elementary School," Educational Forum, XIX (November, 1954), 54.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Abell, E. L. "Grading From the Median," Journal of Educational Research, XVIII, (December, 1928), 425-428.
- Anastasiow, Nicholas J. "Research and Development of an Elementary School Report Card," California Journal of Educational Research, XVII (November, 1966), 209-219.
- Anonymous. "Report on Reports," (Adapted from NEA Leaflet), NEA Journal, LII (December, 1963), 14-15.
- Anonymous. "Report to Parents," NEA Research Bulletin, XLV (May, 1967), 51-53.
- ✕ Austin, Mary C. "Report Cards and Parents," The Reading Teacher, XVIII (May, 1965), 660-663.
- Berman, Samuel. "Revising the Junior High-School Report Card," National Association of Secondary-School Principals Bulletin, XXVII (May, 1943), 49-62.
- Bolmeier, Edward C. "Principles Pertaining to Marking and Reporting Pupil Progress," The School Review, LVIX (January, 1951), 15-24.
- Bolton, Frederick E. "Do Teachers' Marks Vary as Much as Supposed?" Education, XLVIII (September, 1927), 23-39.
- ✕ Brimm, R. P. "Report Cards - Yesterday and Today," The Clearing House, XXXIII (September, 1958), 17-19.
- Buros, Oscar Krisen (ed.). The Sixth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1965.
- ✕ Burton, William H. The Guidance of Learning Activities. (3rd Ed.). New York: Appleton-Century-Crofts, Inc., 1962.
- Cagle, Dan F. "How May We Make the Evaluation and Reporting of Student Achievement More Meaningful?" National Association of Secondary-School Principals Bulletin, XXXIX (April, 1955), 24-27.
- Cajori, F. "A New Marking System and Means of Measuring Mathematical Abilities," Science, XXXIX (1914), 874.
- ✕ Campbell, Laurence R. "So Pupils May Know," Schools and Society, XXXII (December, 1930), 762-763.
- Carter, Robert S. "How Invalid Are Marks Assigned by Teachers?" Journal of Educational Psychology, XLIII (April, 1952), 218-228.
- Caudle, James B. "Educators Look at Grading," Part III. Elementary Level - "A Lingering Problem," The Texas Outlook, VI (July, 1961), 14-29.

- X Chadwick, Ruth E. and others. "The Report Card in a Nongraded School," The National Elementary Principal, XLVII (January, 1968), 22-28.
- Chansky, Norman M. "Elementary School Teachers Rate Report Cards," The Journal of Educational Research, LVI (July-August, 1963), 523-528.
- Cohen, Karen Sue Cotzin, Ph.D. "Reactions of Students to Receiving Grades." Unpublished Doctoral Dissertation, The Johns Hopkins University, Baltimore, 1965.
- Coogan, Dale Roger, Ed.D. "Reporting Pupil Progress: An Analysis of the Relationship Between the Administrative Methods Used for Change and the Degree of Parental Understanding." Unpublished Doctoral Dissertation, The University of Southern California, Los Angeles, 1963.
- Crosby, Muriel E. "Good Report," NEA Journal, LI (April, 1962), 45-47.
- Cummins, Robert E. "Evaluating and Grading," Education, LXXXII (March, 1962), 403-405.
- Cutler, Marilyn H. "Does Your Report Card Format Rate an A?" The Nation's Schools, LXXII (September, 1963), 56-86.
- Davis, J. DeWitt. "The Effect of the 6-22-44-22-6 Normal Curve System on Failure and Grade Levels," Journal of Educational Psychology, XXII (November, 1931), 636-640.
- Dearborn, Walter Fenno. School and University Grades. Wisconsin: University of Wisconsin Press, 1910.
- X DePencier, Ida B. "Trends in Reporting Pupil Progress in the Elementary Grades, 1938-1949," Elementary School Journal, LI (May, 1951), 519-523.
- X Detter, Shirley. "Reporting Progress to Parents," School Activities, XXX (November, 1958), 81-83.
- Doak, E. Dale. "Grading - A Deterrent to Learning," The Clearing House, XXXVII (December, 1962), 245-248.
- X Dobbin, John E. and Ann Z. Smith. "Marks and Marking Systems," Encyclopedia of Educational Research (3rd ed.), New York: The Macmillan Company, 1960.
- Drake, C. Elwood. "Trends in the Field of Evaluating Secondary Education," Educational Administration and Supervision, XXVI (April, 1940), 241-256.

- Edmunds, Brehaut R. Ed.D. "The Relationship Between the Self Image and the Report Card Achievement of Eleventh Year English Students." Unpublished Doctoral Dissertation, The University of Southern California, Los Angeles, 1968.
- Eells, Walter Crosby. "An Improvement in the Theoretical Basis of Five Point Grading Systems Based on the Normal Probability Curve," Journal of Educational Psychology, XXI (February, 1930), 128-135.
- Fay, Paul J. "The Effect of the Knowledge of Marks on the Subsequent Achievement of College Students," The Journal of Educational Psychology, XXVIII, (October, 1937), 548-554.
- X Forman, W. O. "The Gradeless Era in High School," Journal of Education, CXI (May, 1930), 501-502.
- Freyd, Max. "The Graphic Rating Scale," Journal of Educational Psychology, XIV (February, 1923), 94.
- X Good, Warren R. "Should School Marks Be Abolished?" The Education Digest, XI (December, 1945), 11-12.
- Gould, George. "Practices in Marking and Examination," The School Review, XL (February, 1932), 142-146.
- Hadley, S. Trevor. "A School Mark - Fact or Fancy," Educational Administration and Supervision, XL (May, 1954), 305-312.
- Halliwell, Joseph W. and Joseph P. Robitaille, "The Relationship Between Theory and Practice in a Dual Reporting Program," The Journal of Educational Research, LVII (November, 1963), 137-141.
- Hammel, John A. "Report Cards: A Rationale," The National Elementary Principal, XLIII (May, 1964), 50-52.
- Hanson, Earl H. "What Is Success and How Should We Report to Parents?" Education, LXXXII (October, 1961), 126.
- Hillbrand, Earl K. "College Marks - and Remarks," School and Society, XXXIII, (May, 1931), 725-729.
- Jaggard, Guy H. "Improving the Marking System," Educational Administration and Supervision, V (1919), 25-35.
- Jensen, Larry Cyril, Ph.D. "The Effects of Grades on Cognitive and Affective Learning," Unpublished Doctoral Dissertation, Michigan State University, Lansing, 1966.

- Johnson, Franklin W. "A Study of High-School Grades," School Review, XIX (January, 1911), 13-24.
- Johnson, Mauritz, Jr. "Solving the Mess in Marks," New York State Education, XLIX (November, 1961), 12-30.
- Johnson, Roswell H. "The Coefficient Marking System," School and Society, VII (June, 1918), 714-716.
- Jones, John A. "Grading, Marking, and Reporting in the Modern Elementary School," Educational Forum, XIX (November, 1954), 54.
- Karrer, Enack. "Reflections on a New Method of Grading," School and Society, XXIV (November, 1926), 582-584.
- Kelley, Truman L. and others. Stanford Achievement Test, Directions for Administering. New York: Harcourt, Brace and World, Inc., 1964.
- Kelly, Frederick James, Ph.D. Teachers' Marks Their Variability and Standardization. New York: Teachers College, Columbia University Press, 1914.
- Kingston, Albert J. and James A. Wash, Jr. "Research on Reporting Systems," The National Elementary Principal, XLV (May, 1966), 36-40.
- Lawson, Douglas E. "Scoring of Subjective Tests with Several Variables Controlled," Elementary School Journal, XXXVIII (February, 1938), 450-457.
- Link, Frances R. "To Grade or Not to Grade," The PTA Magazine, LXII (November, 1967), 10-12.
- Maxson, Wilbur B. "Grading, a Serious Matter," NEA Journal, LIII (October, 1964), 56-57.
- Michaelis, John U. and Charles Howard. "Current Practices in Evaluating in City School Systems in California," Journal of Educational Research, XLIII (December, 1949), 250-260.
- Middleton, Warren C. "Some General Trends in Grading Procedure," Education, LIV (September, 1933), 5-10.
- X Misner, Paul J. "The Restoration of Report Cards - A Report on Reporting," The PTA Magazine, LVIII (February, 1964), 10-12.
- Monroe, Walter S. Directing Learning in High School. New York: Doubleday, Page and Company, 1927.

X Morris, Lucile. "Evaluating and Reporting Pupil Progress," Elementary School Journal, LIII (November, 1952), 144-149.

Nicholson, Mary Louise E. "Reporting to Parents," Ohio Schools, XXXIX (November, 1961), 23.

Norsted, Roy A. "To Mark or Not to Mark?" The Journal of Education, CXXI (March, 1938), 81-84.

Odell, Charles W. Ph.D. Educational Measurement in High School. New York: The Century Company, 1930.

Ohlson, David. "School Marks vs. Intelligence Rating," Educational Administration and Supervision, XIII (1927), 90-102.

Roelfs, R. M. "Trends in Junior High-School Progress Reporting," Journal of Educational Research, XLIX (December, 1955), 241-249.

Rolf, Fred J. "Reporting to Parents," Ohio Schools, XXXIX (May, 1961), 10-11.

Ruch, Giles Murrell. The Objective or New-Type Examination. New York: Scott, Foresman and Company, 1929.

X Rugg, Harold O. "Teachers' Marks and Marking Systems," Educational Administration and Supervision, I (January, 1915), 117-142.

_____. "Teachers' Marks and the Reconstruction of the Marking System," Elementary School Journal, XVIII (May, 1918), 701-719.

Russell, Ivan L. and Wellington A. Thalman. "Personality: Does It Influence Teachers' Marks?" Journal of Educational Research, XLVIII, (April, 1955), 561-564.

X Sanders, Eugene. ". . . behind the Report Card," Nation's Schools, XXXI (February, 1943), 32.

Smith, Eugene Randolph, and others. Appraising and Recording Student Progress. New York: Harper and Brothers, 1942.

Smith, Henry L. and Wendell W. Wright. Tests and Measurements. New York: Silver, Burdett and Company, 1928.

- Smutter, Faith. "Report Cards. . . . Problems and Possibilities," The National Elementary Principal, XL (September, 1960), 166-171.
- Starch, Daniel, Ph.D. Educational Measurements. New York: The Macmillan Company, 1916.
- Stockard, Jimmy. "Case for Dual Grading. . . . Ability Versus Achievement," Texas Outlook, XLVI (May, 1962), 30.
- × Symonds, Percival M. Measurement in Secondary Education. New York: The Macmillan Company, 1927.
- Symonds, Percival M. "Notes on Rating," Journal of Applied Psychology, IX (1925), 188-195.
- Thorndike, E. L. and E. O. Bregman. "On the Form of Distribution of Intellect in the Ninth Grade," Journal of Educational Research, X (November, 1924), 271-278.
- Tiegs, Ernest W. Ph.D. Tests and Measurements for Teachers. New York: Houghton Mifflin Company, 1931.
- Trabue, Marion Rex, Ph.D. Measuring Results in Education. New York: American Book Company, 1924.
- Traxler, Arthur E. Techniques of Guidance. New York: Harper and Brothers, Publishers, 1957.
- Varner, Glenn F. "How Do We Best Appraise and Report Individual Pupil Progress in the Junior High School?" The Bulletin of the National Association of Secondary-School Principals, XLVI (October, 1962), 23-25.
- Vredevoe, Lawrence E. "How May We Make the Recording and Reporting of Pupil Achievement More Meaningful?" National Association of Secondary-School Principals Bulletin, XXXVII (April, 1953), 179-182.
- Wakeham, Glen. "Humanizing Grades," School and Society, XXXIV (October, 1931), 596-598.
- Wherry, Joseph E. "What Are Current Trends in Reporting Student Growth and Achievement to Parents?" The Bulletin of the National Association of Secondary-School Principals, XLIII (April, 1959), 153-157.
- × White, Mary Alice and Ann E. Boehm. "The Child's World of Marks," NEA Journal, LVII (January, 1968), 12-13.

Williams, Lois. "Teachers and Parents: Did You Know That Your Children Feel This Way?" Childhood Education, XXXV (October, 1958), 60-64.

Williamson, E. G. "The Cooperative Guidance Movement," The School Review, XLIII (April, 1935), 273-280.

X Wilson, Charles H. "Educational Innovation: Are Public Schools Going Overboard?" Nation's Schools, LXXX (November, 1967), 66-68.

X Wrinkle, William L. Improving Marking and Reporting Practices in Elementary and Secondary Schools. New York: Rinehart and Company, 1947.

_____. "School Marks - Why, What and How?" Educational Administration and Supervision, XXI (March, 1935), 218-225.

APPENDICES

APPENDIX A

DIOCESE OF SALINA - PUPIL PROGRESS REPORT

GRADES 4-5-6-7-8

19__ 19__

Name _____ Grade _____

School _____

	QUARTERS			
	1	2	3	4
SCHOLASTIC RECORD				
KNOWLEDGE OF RELIGION				
READING				
ENGLISH				
HANDWRITING				
SPELLING				
MATHEMATICS				
SOCIAL STUDIES				
SCIENCE				
MUSIC				
ART				
PHYSICAL EDUCATION				

Teacher _____

Principal _____

	QUARTERS			
	1	2	3	4
PERSONALITY RECORD				
CONDUCT				
STUDY HABITS				
Beginning or completing work on time				
Working independently				
Making an effort to improve				

	QUARTERS			
	1	2	3	4
ATTENDANCE RECORD				
Days Absent				
Times Tardy				

MEANING OF MARKS

A - Outstanding achievement (95 - 100)
 B - Above average achievement (88 - 94)
 C - Average achievement (77 - 87)
 D - Below average achievement (70 - 76)
 U - Unacceptable achievement (69 or below)

Key for subheadings:

A check (✓) indicates improvement is needed.

Grade assignment for next year: _____

APPENDIX B

S A M P L E

SEVEN DOLORS GRADE SCHOOL
Manhattan, Kansas
1969 - 1970

93

DIAGNOSTIC REPORT OF PUPIL PROGRESS

To: Mr. & Mrs. John Doe
(Parent's Name)

Date: Sept. 2, 1969

Jim Doe
(Pupil's Name)

Teacher: Sister Mary Lou Pfannenstiel

Grade: 6

Principal: Sister Mary Barbara Bader

This Diagnostic Report of Pupil Progress serves as an aid to evaluate the strengths and/or weaknesses of your child in the various subject areas and to promote understanding and cooperation between the home and the school.

ETHICS Jim is very interested in his search for God. With encouragement, his participation in class will increase.

READING Much extra reading is necessary for Jim to advance in comprehension. His vocabulary is improving.

ENGLISH Jim has made remarkable improvement for constructing good sentences. Jim's weakness lies in good study habit organization.

SPELLING Jim needs much help in spelling as he generally misspells one out of five words. He finds defining words difficult.

SOCIAL STUDIES Jim is doing very well in class participation and volunteers for extra reports, projects and information. His comprehension is good and will improve, I am sure. He is showing interest and deserves much praise for this interest.

MATHEMATICS Math is very challenging for Jim. He is very weak in comprehending problems and how to solve them. Computation is weak (multiplication).

SCIENCE Jim has manifested a weakness in scientific concepts that involves Math (Chap. I). He understands concepts of heat & molecules.

HANDWRITING
Jim has very good handwriting. I hope he will keep it.

ART Jim has creative ability which will take time to develop. He lacks confidence in this ability.

MUSIC Jim enjoys music very much and is very enthusiastic.

PHYSICAL EDUCATION
Jim is showing improvement in sportsmanship and in physical skill.

ATTENDANCE RECORD

QUARTERS

Days Absent:

1

2

3

4

APPENDIX C

Seven Dolors Grade School
Manhattan, Kansas
April 25, 1969

Dear Parents of the 1969-1970 Sixth Graders:

What other subject is more frequently discussed than the report card? Teachers, administrators, and parents are concerned with the relative merits of the various reporting systems now in vogue, and writers seem to intensify rather than subdue the controversy.

Research is an important and necessary aspect of education; yet very little research has been done in the area of evaluating pupil progress. Whether pupil achievement is aided or hindered by receiving a traditional report card in comparison with the achievement of pupils receiving a diagnostic report card containing no grades remains an unanswered question. As a teacher I am interested in helping to search for an answer. It is with this view in mind that I am conducting a comparative study in achievement of pupils receiving traditional report cards and pupils receiving a more recently developed diagnostic report card. Sister Mary Barbara, the principal, has given her approval and full support to this research project.

We are using the parallel-group design in which pupils of the control group will receive the traditional report cards and pupils of the experimental group will receive the diagnostic report card. Parents of the pupils of the experimental group (approximately fifteen pupils) will receive more information in the fall regarding the diagnostic report card.

It is hoped that this Diagnostic Report of Pupil Progress will serve as an aid to the pupil's growth. The purpose of this report may be summarized as twofold:

- (1) to impart information to you, the parents, of your child's progress regarding the strengths and/or weaknesses in the various subject areas.
- (2) to promote understanding and cooperation between the home and the school.

By helping parents and children to understand the individual strong and/or weak points, we hope to show where progress is being made and where greater effort will be needed.

This study, under the direction of Dr. Charles Peccolo of K-State University, is being made in partial fulfillment for my Master's Degree. Your cooperation and understanding in the forthcoming year will be greatly appreciated. I shall be happy to answer any questions that you may have regarding this subject. Please feel free to contact me at any time. Next week, as you know, we are having our open-house for Parents, and I shall be available at that time.

With prayers and best wishes for a successful school year,
I am

Sincerely,

Sister Mary Lou Pfannenstiel

Sister Mary Lou Pfannenstiel

REPORT CARD FORMS AND ACHIEVEMENT

by

SISTER MARY LOU PFANNENSTIEL

B. A., Marymount College, 1967

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Education

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1970

This study attempted 1) to review available literature on the scope and functions of report cards as well as present a historical review of report cards, and 2) determine if the report card form would result in significantly greater achievement by the pupil. More specifically, the primary consideration of this study was to test the following hypothesis:

There is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards.

The study was conducted the first twenty-seven weeks of the school year 1969-1970 from August to March in one sixth grade classroom at Seven Dolors Grade School located in Manhattan, Kansas. The parallel-group design was used in which an attempt was made to match twenty-eight pupils from the sixth grade on the basis of sex, intelligence and achievement scores for a control and an experimental group within the limits defined by the study. Through the use of the t-test, both sampling groups were considered equal prior to the experiment.

During the period of the experiment the two groups received exactly the same treatment and performed the same tasks at the same time of day. One exception was evident--the control group received the traditional report card and only letter grades on their papers, while the experimental group received the diagnostic report card and written teacher remarks on their papers. This diagnostic card sought analytically to identify strong and weak aspects of a pupil's performance, so appropriate corrective measures may be taken.

The data evaluated consisted of scores obtained in Stanford Achievement Tests (Form X and Form W). A null hypothesis was tested concerning the difference among the means using Fisher's Table of t and a .05 level of confidence.

The study was designed to determine the difference, if any, in the achievement between a control group receiving a traditional report card and an experimental group receiving a diagnostic report card. Achievement was measured by the Stanford Achievement Test - Form X (pre-test) and the Stanford Achievement Test - Form W (post-test).

Statistical analysis showed no significant difference between the control group receiving the traditional report card (letter grades) and the experimental group receiving the diagnostic report card (written descriptive analysis of pupil strength and weakness). This evidence contradicted the notion that a change in report card forms, in itself, will result in higher academic achievement.

As a result of this experimentation, the writer seems justified in stating that there is no significant difference in achievement of pupils receiving traditional report cards and those receiving diagnostic report cards.