

STATEMENT OF PERMISSION TO COPY

In presenting this professional paper in partial fulfillment of the requirements for an advanced degree at Montana State University, I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this professional paper for scholarly purposes may be granted by my major professor, or, in his absence, by the Director of Libraries. It is understood that any copying or publication of this professional paper for financial gain shall not be allowed without my written permission.

Signature

Marley Douglas M. Dowall

Date

Aug 9, 1973

THE DEVELOPMENT OF AN ACCOUNTABILITY MODEL FOR
THE TEACHING OF READING

by

MARLEY DOUGLAS ^{✓✓}MC DOWALL

A professional paper submitted to the Graduate Faculty in
partial fulfillment of the requirements for the degree

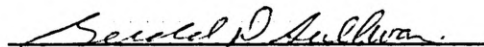
of

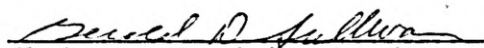
MASTER OF EDUCATION

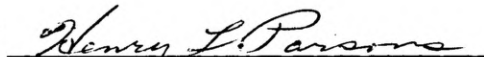
with concentration in

Elementary Administration

Approved:


Head, Major Department


Chairman, Examining Committee


Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

August, 1973

ACKNOWLEDGMENT

The writer wishes to express his appreciation to Dr. Gerald Sullivan for his advice, patience, and help in preparing this paper. Appreciation is also extended to Dr. Eric Strohmeyer for his technical advice. A special thanks goes to my wife, Janice, who not only gave me encouragement, but also provided valuable assistance in the construction of this paper.

TABLE OF CONTENTS

	Page
LIST OF TABLES.	vi
ABSTRACT.	vii
Chapter	
1. INTRODUCTION	1
PURPOSE AND NEED	1
QUESTIONS TO BE ANSWERED	3
GENERAL PROCEDURES	3
LIMITATIONS.	4
DEFINITIONS OF TERMS	5
SUMMARY.	6
2. RELATED LITERATURE CONCERNING ACCOUNTABILITY IN EDUCATION.	7
EVALUATION IN GENERAL.	7
EVALUATION AND THE READING PROGRAM	9
ACCOUNTABILITY IN EDUCATION.	12
SUMMARY.	15
3. FORMULATION OF THE MODEL	16
POPULATION DESCRIPTION	16
DEFINITION OF CATEGORIES	18
COLLECTING AND ORGANIZING DATA	19
HYPOTHESIS	20
METHOD OF DATA ANALYSIS.	20

Chapter	Page
Construction of the Model.	20
Use of the Model	21
SUMMARY.	22
4. RESULTS AND INTERPRETATION OF DATA ANALYSIS.	23
CORRELATION OF VARIABLES	23
USE OF MODEL EQUATION.	27
SUMMARY.	28
5. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.	30
SUMMARY.	30
CONCLUSIONS.	33
RECOMMENDATIONS.	34
LITERATURE CITED	36

LIST OF TABLES

Table	Page
1. Variable Student Data	24
2. Correlation of Variables.	25

ABSTRACT

Evaluation is inherent in every meaningful enterprise in which people engage. As soon as goals or purposes for an activity or enterprise are announced, the question is raised, "To what extent are the purposes or objectives being realized?" Here lies the roots of the rapidly growing concern in education referred to as "accountability." An accountability device which may assist in evaluation of reading programs was a major concern of this study.

Several variables were considered in the development of the device or model. The variables thought to have predictive value in determining reading achievement were: past performance, intelligence quotient, age, educational level of parents, sex, and participation in a special reading class.

The variable data from eighteen fourth grade students was tabulated, coded, and key punched. Statistical analysis was facilitated by computer. The first phase in data analysis was finding the correlation among the independent variables. Then, in order to determine the degree of relationship between the post test reading score (criterion) and all independent variables at one time, a multiple regression analysis was accomplished. From these results and the correlation data, a "weighted" factor was assigned to each variable according to its importance in predicting the criterion.

A model equation was formed utilizing variables (X), weighted factors (b), and a constant (k) to predict the post test score. The model took the following form:

$$X_1b_1+X_2b_2+.....X_7b_7+k = \text{Criterion}$$

According to the multiple correlation coefficient obtained (.99), the equation is significantly predictive. However, the recommendations made were:

1. Further investigations should be made with larger sample size, involving many teachers and covering larger geographic area.
2. The variable weights and model arrived at through this study should be used to predict the reading progress in the fourth grade class in the same school. The predicted progress should then be compared to the actual progress made during the year.

Chapter 1

INTRODUCTION

Accountability in education has rapidly become a point of interest and greater concern among parents, educators, and the public in general.¹ Are our schools achieving the goals for which they are held responsible? To what degree are the goals being reached? How effective are the teachers, their methods, and the materials they use?

Thorough evaluation is a very difficult task requiring a great deal of time and energy even for a single program, let alone a school or an entire system. Many conditions will need examination prior to determining the effectiveness of any method, program, or teacher. What was the size of the class? Did their abilities parallel national, state, or school norms? Was there ample equipment and materials available? The list could go on and on. The point is: evaluation is a tough job with few guidelines and objective instruments to help the evaluators as well as the teachers. An accountability device which may assist in evaluation of reading programs was the major concern of this study.

PURPOSE AND NEED

The purpose of this investigation was to develop an accountability model for use in evaluating reading progress. The model was

¹Norman E. Hern, "Evaluation is a Full-Time Job," American Education, V, (April, 1969), 18.

based on student's ability, age, sex, past performance, and parent's educational level.

As was earlier mentioned, evaluation is a many faceted problem. Judgment of teacher performance, as observed by the writer, is largely done in a subjective manner, and perhaps rightly so. However, all too often, teacher evaluation consists of one or two short classroom observations, a look at daily lesson plans, followed by a short verbal critique. The rest of the evaluative information would come from informal conversations and from indirect observation, i.e., pupil behavior in the hall and frequency of disciplinary problems sent to the office. An evaluation of this nature may be a valid one in a few cases, but one needs little imagination to think of situations and occurrences that would paint an all but true picture of a teacher's performance.

Even when a thorough evaluation is contemplated, the model will have value. An evaluator who makes frequent classroom observations uses numerous structured conversations with teachers dealing with methods, materials, and pupil progress will also want to make as much use as possible of objective data. If he can compare, using the model, the progress made by each pupil in a classroom to the average progress made by many students with similar abilities and past performances, then he will have an additional source of information which will aid in a more valid evaluation. However, this model should in no way be used

as a replacement for conventional evaluation procedures. Its purpose is intended only to be an aid to already established methods.

If and when accountability becomes the rule rather than the exception, objective data will be required. At present, much data is prejudiced in some way. With this model, it is hoped that the criterion will not be unfair to any teacher because of the inherent or environmental deficiencies of his students; nor will it be advantageous to teachers of gifted groups.

QUESTIONS TO BE ANSWERED

Can reading progress be accurately predicted by using a combination of these indicators: intelligence quotient, past performances, sex, age, and educational level of the parents? If this is possible, can a model be developed, based on predicting a minimum amount of progress, which will be an effective, efficient evaluatory device for reading instruction and pupil progress?

GENERAL PROCEDURES

A review of literature was completed using materials available at the Montana State University Library. Educational Research Information Center (ERIC) materials dating from 1968 through 1973 were also used for collecting reference data and gathering general background information. Based on the information gained, the use of a multiple

regression equation seemed to be the most accurate predictor of pupil progress. This equation requires use of several variables. This data was obtained from students permanent cumulative school records.

Once the variable data was collected and recorded in table form, the information, along with a multiple regression equation, was synthesized by computer. It calculated the inter-correlation among variables and total correlation of the sum of the variables with the criterion. The equation to be used was: $X_1b_1 + X_2b_2 + \dots$
 $k = \text{Criterion (expected raw score one year later)},$ where X_1 is the past reading raw score and b_1 is an appropriate "weight" computed as a result of its correlation with the criterion. X_2 indicates I.Q. and b_2 an appropriate "weight." The variables are multiplied times their "weights." The final sum of the products is the criterion.

LIMITATIONS

The greatest inherent weaknesses of this investigation lie in the use of standardized achievement test results. However, they existed as the only source of objective data available on reading ability in the school system and they are the most widely used tests in the area. The study took place in a small school system in which rural children make up 75 per cent of the total enrollment. The socio-economic average is considered somewhat lower than one would expect to find in the upper midwest in general. The review of related literature and

research was conducted at Montana State University.

A larger sample size, over a larger geographic area would, no doubt, result in a more reliable model. However, this investigation was only a pilot study from which other studies with larger scopes may evolve.

DEFINITIONS OF TERMS

The terms throughout this study which were used in such a manner as to necessitate clarification are defined below.

Accountability. In education, accountability is the idea of holding liable the teachers and/or administration for their methods, programs, time, and cost as related to the end product of pupil performance.

Model or equation. This is a mathematical formula in which variables (i.e., I.Q., past performance, age, etc.) are multiplied times a computed factor or "weight." The products are added to a constant to find the predicted post test score (criterion).

Weight. Weight is an appropriate decimal factor computed on the basis of importance of each variable in predicting the criterion.

Behavioral objectives. These are goals stated in such a manner that they indicate what the student will know or be able to do following

instruction.

Variables. Variables are factors thought to affect reading progress and included: I.Q., age, sex, special reading class, parents' educational level, and past performance.

SUMMARY

The need for more efficient and effective methods of evaluation was evident not because schools are failing but because educators are continually striving for improvement. The purpose of this study was to develop an aid or model which may be of some assistance in objectively evaluating reading progress in light of past performance and ability. In general, the model is in the form of an equation in which the variables (i.e., I.Q., past reading scores, etc.) would be correlated with each other and the criterion. Computations arrived at through the use of a computer were assigned as appropriate "weights" to be multiplied times each variable. The sum of these products will give the expected amount of progress that any particular student should make in a one-year period. Past performance and expected progress use the raw reading score as the means of measurement.

The first step, as indicated in the preceding section, was a review of literature on accountability and evaluation of reading programs. The results of this review are presented in Chapter Two.

Chapter 2

RELATED LITERATURE CONCERNING ACCOUNTABILITY IN EDUCATION

The following summary of the review of literature was based on three major divisions. From general to specific, they are: evaluation in general, evaluation and the reading program, and accountability in education.

A broad view of evaluation in general was needed for a deeper understanding of present accountability problems confronting our schools. The need for accountability and awareness of its positive as well as negative factors is essential in an investigation in which the purpose is to develop a model formula for use in the evaluatory process.

EVALUATION IN GENERAL

Evaluation is inherent in every meaningful enterprise in which people engage. As soon as goals or purposes for an activity or enterprise are announced, the question is raised: "To what extent are the purposes or objectives being realized?"¹

Evaluation of educational procedures, in an informal way, has probably taken place as long as education has been in existence. It is a very necessary factor in effective teaching, for without it one has no indication of the effectiveness of the method used. The problem

¹Paul J. Misner, Elementary School Administration (Columbus, Ohio: Charles E. Merrill Books, Inc., 1963) p. 163.

for centuries has been: What factors of instruction leads to the greatest amount of learning with relation to the time and cost of material or equipment used. Was the teacher the all important variable with success or was it the instructional method used, sophistication of material, or perhaps, length of time under instruction?² Thousands of investigations have taken place in hopes of answering the above question. Their results seemed often elusive, contradictory, or prejudiced because of uncontrolled variables. Research is essential to find good, reliable measuring instruments. James E. Allen, former U.S. commissioner of education, said, "Unless we develop the capacity to assess the value of one instructional program over another, real accountability is impossible."³

The concentration of population and reorganization of school districts in recent years has brought about larger and more complex schools creating an even greater need for more formal, objective evaluation.

Supervisors generally have more teachers, instructing over a wider range of curriculum than ever before. They are dealing with more diverse abilities, personalities, needs, and interests of students

²Walter N. Durost, "Evaluation of Reading Progress," The Reading Teacher, XXIV (January, 1971), 290.

³Donald D. Woodington, "Accountability from the Viewpoint of a State Commissioner of Education," Phi Delta Kappan, LIV (October, 1972), 95.

because of the wide socio-economic gap which makes the evaluatory process even more complex. Evaluatory procedures to secure advances in education may surely be part of any plan of accountability and may involve teacher improvement, responsibility for results, public information systems, fiscal auditing, or any one of a number of other possibilities.⁴

EVALUATION AND THE READING PROGRAM

The goals and objectives in many schools still emphasize what the instructional "input" will be, especially on short term objectives, rather than what this student will know or be able to do following the instruction. Behaviorally stated objectives are quickly becoming an important aspect of a new trend in education.⁵ If teachers know that as a result of their instruction the students will be able to recognize and say the letters of the alphabet, or add common fractions with like denominators, then he will not only have a better guide for his instruction, but evaluation will become a much easier task.

One of the most important problems in this procedure is the

⁴Jacob Landers, "Accountability and Progress by Nomenclature: Old Ideas in New Bottles," Phi Delta Kappan, LIV (April, 1973), 539.

⁵Raymond Bernabei, "Instructional Accountability," Speech presented to the 104th Annual Convention of American Association of School Administrators. Atlantic City, New Jersey, February 11, 1972.

writing of precise, measurable objectives. The evaluative device then would be produced on which the students must perform specific tasks which are clearly stated in the objectives. Criterion scores would be established as a cooperative effort within the educational unit.

In a Staff Performance Improvement Appraisal Plan (SPI&A) conducted in the Newport-Mesa Unified School District in southern California, the following format is used: (1) Each teacher completes an appraisal form which contains the stated objective(s), results of the pretest, the performance criteria, the criterion measures (pretest and posttest), and evaluation (data collection) plan. (2) The form is reviewed by a team of fellow teachers and principal. (3) At the end of the semester, the teacher posttests the class relative to the objectives and reports the results on a form to the team and principal. (4) A decision is reached whether or not the objectives were attained. One class "improvement" type evaluation is also made where the teacher states objectives, is observed presenting the lesson (by team and principal), posttests the class to decide whether objectives were met, and observers meet with teacher to provide suggestions that are aimed at improvement of teacher's competencies in promoting pupil learning.⁶

A second problem exists which opponents of the procedure clearly

⁶Fred Niedermeyer and Stephen Klein, "An Empirical Evaluation of a District Teachers' Accountability Program," Phi Delta Kappan, LIV (October, 1972), 100.

bring out. That is one where behaviorally stated objectives are impossible to state and evaluation is nearly non-existent. A common example would be affective behavior. Was attitude or appreciation of art or music changed because of instruction? Is he now more empathetic toward the negro child and his problems? Did instruction help him define a system of values? Was instruction valuable in establishing better skills in human relations? These are difficult objectives not only to reach but also to measure to what degree they were reached. Landers stated, "It is necessary to be conscious of the very real danger that the aims of education will be increasingly restricted to those which can be most easily measured, rather than those which are most important."⁷

Others viewed the evaluation process in a broader spectrum. They would have reading specialists from nearby universities or colleges come in, meet with administrators and teachers, explain procedures and purposes, answer questions and try to minimize anxieties. The specialists would make a classroom observation followed by a personal interview with the teacher. They would also review opinion surveys which had been previously completed by the teacher on their reading program and evaluate standard achievement test results. They would then again meet with teachers and administration to give a

⁷Landers, op. cit., p. 541.

general evaluation, make suggestions, and constructively criticize the entire reading program. Care would be taken not to rate any individual teacher. The key to this type evaluation is the confidence and cooperation shared by the participants.⁸

Cautions were made that evaluators must be aware of day-to-day problems faced by the teacher, that short term observations may be misleading, and that the test scores reflect performance on a particular day, at a particular time, and under particular conditions.

Durost indicated the importance of evaluation in the following statement.

The quality of instruction is directly related to the effectiveness of the evaluative program. If evaluation steps are omitted, much testing is wasted and valuable insights are lost. Evaluative process should be used to improve reading ability through improvement of instruction so that the chances of every pupil becoming an effective reader are increased.⁹

ACCOUNTABILITY IN EDUCATION

In all articles and papers reviewed concerning accountability, there was a general consensus of opinions that: fewer and fewer people will remain idle and willingly finance schools with their increasing costs without some objective evidence of their effectiveness. Not only

⁸Sidney J. Rauch, "How to Evaluate a Reading Program," The Reading Teacher, XXIV (December, 1970), 244.

⁹Durost, op. cit., p. 291.

because of cost but because of the importance of a more lengthy education on the future life and livelihood of young people are parents, students, and the community in general becoming more interested in the results of education.

According to a Gallup poll taking in 1970, approximately two of every three people believed that teachers and school administrators should be held more accountable for the progress of their children.¹⁰

Employers are wanting to know what specific skills and knowledge a potential employee possesses. A diploma must certify the quality of performance in reading, writing, mathematics, and other skills rather than it being a certificate of longevity or endurance of twelve years. In short, education will become more accountable for relating process, procedures, and methods to results.

Even the federal government has taken the position that recipients of federal funds for education show measurable, objective data that the funds used are, in effect, helping to change behavior commensurate with their amounts.

These requirements require a logical, systematic method for collecting information of educational growth in light of ability and past performance so that educators may take appropriate action to drop inept programs and to strengthen those that seem most promising.

¹⁰Landers, op. cit., p. 539.

Woodington further emphasizes this idea through his statement that,

Accountability is rather a cyclic process in that it concerns: establishing goals, setting specific objectives, devising programs to meet objectives, carrying out the programs, measuring their degree of success, comparing costs and performance under alternate programs, revising and trying again.¹¹

Opponents of accountability indicated that if it is to be used in the reading program, evaluators must proceed with caution and awareness of the many problems. Erratic learning due to situations beyond the control of the instructor affects the reliability of test performances and other evaluative devices. Examples of outside situations that affect learning are: long term illness, family trouble, and mothers who work temporarily.

Sutton found the following examples of erratic learning in a study of 129 primary students as was evidenced by grade equivalent scores on standardized achievement tests. While under the same instructor, Kim gained thirty-two months in the first semester followed by a gain of two months in the second. Diane gained twenty-three months in the first grade, twenty months in the second, and showed no improvement in the third grade.¹²

Even with its negative aspects, many critics of accountability

¹¹Woodington, op. cit., p. 95.

¹²Marjorie H. Sutton, "Accountability and Teaching Reading," The Reading Teacher, XXV (March, 1972), 528.

agree that if it will improve general reading instruction, get rid of outmoded classroom practices and make teachers more aware of better methods and techniques, then the positive points far outweigh the negative.

SUMMARY

A major question in education today as well as centuries ago is: What factors of an educational program bring about the most desired results with the least amount of time and expense? Some authorities feel that clearly stated, behavioral objectives be used and evaluation geared to them. However, affective behavior such as appreciation is nearly impossible to measure objectively. Evaluation by "reading specialists" might prove very satisfactory, but they are spread far too thin for thorough work. Accountability in education is relatively new but growing rapidly. Objective results are being required by the community. They want information on educational growth in light of ability and past performance so that they can help determine which programs are inept and which show the most promise. To date, little has been done in the area of objective accountability in education and finding authoritative articles on objectively oriented evaluatory devices or models is rare.

The next step in this study was to describe procedures and formulate a model. This is discussed in Chapter Three.

Chapter 3

FORMULATION OF THE MODEL

The development of an accountability model for evaluation of elementary reading necessarily includes students, teachers, parents, and school with characteristics common only to them.¹ These variables, no doubt, have influenced the result of the investigation so an attempt has been made in this chapter to familiarize the reader with the school and the population of the community in general and includes a more detailed description of students involved, their parents, and the teacher. The methods employed in (1) gathering and organization of data, (2) analysis of data and formation of a predictor equation, and (3) how this equation may be used are explained.

POPULATION DESCRIPTION

The study investigated the reading comprehension progress of eighteen fourth grade students in the Wing Elementary School, Wing, North Dakota. Of the eighteen students involved, thirteen were rural in the true sense. These families lived on farms and derived a major portion of their income from production of agricultural products. The remainder of the student's parents were involved in occupations which

¹Raymond B. Cattell and H. J. Butcher, The Prediction of Achievement and Creativity (New York, New York: Bobbs - Merrill Co., 1968), p. 163.

served the agricultural community. The socio-economic status was considered lower than in a more urban area. The average I.Q. of the eighteen students was 101, as measured by the Otis Quick Scoring Test of Intelligence. The average educational level completed by their parents was tenth grade.

The school was an accredited institution employing seventeen qualified teachers and one teacher aid. Of these, seven were utilized entirely at the elementary level. Grades seven through twelve and administration offices were contained in a connected building.

The average length of total teaching experience of the elementary teachers was fifteen years with a minimum of ten years and a maximum of twenty-nine years. The average college level completed by the elementary teachers was 146 quarter hours or the equivalent of three years. A teacher without a four-year degree may seem unusual. This was the result of a North Dakota law which allowed elementary teachers who completed a two-year teaching course prior to 1964 to continue teaching.

The fourth grade teacher of the students involved in the study had sixteen years of elementary experience, had completed 151 quarter hours of college work in the elementary education field, and was forty-seven years old.

The school system was rather unique in that grades one through six were self contained and were located in an older building which

was connected to a newer building containing grades seven through twelve, administrative offices, and lunch room facilities. The school was well kept and attractive.

DEFINITION OF CATEGORIES

Below are several variables which were considered in the development of the model. Each was assigned a weight in proportion to its predictability through the use of a regression equation which is treated in the section on construction of the model.

1. The raw scores on the reading comprehension sub test of the Iowa Tests of Basic Skills which was administered in October was used as the first variable.

2. The intelligence quotient as indicated by the group intelligence test (Otis Quick Scoring Test of Intelligence, form B) was used as the second variable.

3. The number of years of formal education completed by the father was the third variable.

4. The number of years of formal education completed by the mother was the fourth variable.

5. The fifth variable considered was the chronological age of the student in months as of September of the current school year.

6. Sex was coded by 1 for female and 2 for male and was used as the sixth variable.

7. The last variable included was the number of months each student was fully involved in a special reading program which took place at some other time than during the regular reading period in his home room. If a student was involved part time, i.e., less than forty-five minutes per day, five days a week, a proportional amount was used.

8. The criterion was the raw score on reading comprehension sub test as measured by the Iowa Test of Basic Skills one year after the preliminary test.

COLLECTING AND ORGANIZING DATA

All data was collected from information contained in each student's permanent, cumulative records. The Iowa Test of Basic Skills subtest on reading comprehension had a reliability of .95.² The tests were machine scored at the company's central office. The company furnished individual results which had an adhesive back. These were affixed to the permanent record which eliminated errors in transferring results.

The intelligence quotient was derived from the Otis Quick Scoring Mental Ability Test which had a reliability of .91.³

²H. H. Remmers, "Achievement Tests," Fifth Mental Measurement Yearbook, ed. Oscar K. Buros (Highland Park, N. J.: Gryphon Press, 1959), pp. 16-19.

³B. J. Johns, "Mental Ability Tests," Fifth Mental Measurement Yearbook, ed. Oscar K. Buros (Highland Park, N. J.: Gryphon Press, 1959), pp. 96-99.

The data was organized on a table where each student was coded and all variable data was entered on the same horizontal line.

HYPOTHESIS

The null hypothesis held was that reading comprehension progress cannot be predicted through the use of variables $X_1 \dots 7$ at the .05 level of significance. The alternative hypothesis was that reading comprehension progress can be predicted through the use of $X_1 \dots 7$ variables at the .05 level of significance.

METHOD OF DATA ANALYSIS

The analysis of the data was divided into two main categories.

1. The construction of a regression equation model.
2. The use of the model.

Construction of the Model

When the data was compiled in table form, multiple regression was used to find the relationship between the criterion and the predictor variables by the following equation:

$$Y = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + k$$

Where "Y" is the criterion (raw score in October subtest in reading comprehension). The $b_{1 \dots 6}$ indicate appropriate weights which are assigned to each of the predictor variables. The $x_1 \dots x_6$ are the variables, i.e., I.Q., educational level of parents, etc. K indicates

a constant.

A coefficient of multiple correlation (R^2) was obtained to find the amount of correlation between the criterion and the sum of the products of the weighted variables.

An "F" test was completed which indicated rejection of the null hypothesis and acceptance of the alternative at the .05 level of significance by the following equation:

$$F = \frac{(R^2 - 0)/df_1}{(1 - R^2)/df_2}$$

Where df_1 is the number of independent predictor variables (7) or degrees of freedom and df_2 is the total number of students (N)- df_1 -1.

If an F of 1.96 or greater is obtained, the null should be rejected and the alternative hypothesis accepted. The "F" computed with the data derived for this investigation was 58.

Computations were made through the use of a Sigma Seven computer located at Montana State University, Bozeman, Montana.

Use of the Model

Once the appropriate weights have been assigned to the variables, they may be employed by multiplying them times the variables and using the sum to predict the progress for any fourth grade student whose variables are known providing data from the same test is available.

SUMMARY

The study took place in the Wing Elementary School, Wing, North Dakota and involved eighteen fourth grade students.

The variables involved in the development of the model were: pretest scores, intelligence quotient, parents' educational level, age, sex, and participation in special reading classes.

The null hypothesis held was that reading comprehension progress cannot be predicted through the use of the variables $X_1 \dots X_7$. The alternative was that prediction could be made with the use of the seven variables. A multiple regression equation was constructed which involved the variables (X) and appropriate "weights" (b) which were to be found through data analysis. The equation took the following form:

$$(\text{Criterion}) \quad Y = X_1 b_1 + \dots + X_7 b_7 + k$$

An "F" test was devised which could be used for rejection or acceptance of the hypotheses.

The next chapter deals with the variable data that was compiled, the correlation among the variables, and computation of the appropriate "weights" for the model.

Chapter 4

RESULTS AND INTERPRETATION OF DATA ANALYSIS

This chapter is devoted to the presentation of the results of data analysis, construction, and use of the model.

Having established the variable information required and the model equation to be used, the data was gathered and formed into Table 1, page 24.

- X_1 September 71 Raw Reading Score (pretest)
- X_2 Intelligence Quotient
- X_3 Educational Level of Father in Years
- X_4 Educational Level of Mother in Years
- X_5 Chronological Age
- X_6 Sex (1-female, 2-male)
- X_7 Special Reading Class (months)
- Y September 72 Raw Reading Score (Criterion)

The data was coded and key punched. The statistical analysis was facilitated by use of the Sigma Seven computer at Montana State University.

CORRELATION OF VARIABLES

The Pearson paired analysis was used as the first phase in finding the correlation among the independent variables. The amount of correlation is indicated by a correlation coefficient (r) which is

Table 1

Variable Student Data

Students	Variables							Y
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	
A	60	110	12	12	9.6	1	0	70
B	36	93	9	10	9.5	1	0	46
C	26	84	8	8	10.5	2	9	37
D	56	111	8	8	9.0	2	0	61
E	36	103	13	16	10.4	2	9	47
F	17	92	8	8	10.9	2	0	34
G	67	120	17	13	9.5	2	0	84
H	35	86	8	8	11.8	2	9	35
I	23	90	8	8	11.8	2	9	32
J	47	105	8	12	9.8	1	0	60
K	30	80	8	12	9.8	2	9	34
L	54	117	17	12	9.5	1	0	72
M	32	102	8	10	10.9	2	0	45
N	55	115	12	12	9.5	1	0	75
O	65	118	8	12	9.2	2	0	74
P	23	85	12	8	11.8	2	9	34
Q	50	105	12	13	9.4	1	0	68
R	50	108	13	12	9.2	2	0	60

simply an index of the relation between two variables, expressed in such a form that a perfect positive relation will yield a correlation coefficient of +1.0; a perfect negative relation, a coefficient of -1.0; and complete independence of the two variables, a value of zero. The results of correlation analysis are shown in Table 2.

Table 2
Correlation of Variables

	2	3	4	5	6	7	8
1. Pretest score	.89	.49	.54	-.78	-.37	-.63	.96
2. Intelligence		-.56	.54	-.69	-.34	-.75	.94
3. Father's Education			.55	-.31	-.27	-.23	.60
4. Mother's Education				-.55	-.33	-.24	.58
5. Chronological Age					.43	.68	-.77
6. Sex						.50	-.48
7. Special Reading Class							-.72
8. Posttest Score (criterion)							

Using a two-tailed test, the critical value of r is .468, when the level of significance has been predetermined at .05 and there are 18 degrees of freedom.

As one might expect, there is high correlation between pretest score and intelligence quotient (.89), pretest score and criterion (.96),

and intelligence quotient and criterion (.94). A lower, yet positive, correlation exists between parents' educational level and criterion (.58 and .60). The negative correlation between age and intelligence quotient (-.69), age and pretest (-.78), and age and criterion (-.77) may at first seem unusual. However, in the class studied, there were seven students who had at some time been retained. Their mean I.Q. test score of 91 is somewhat lower than the class mean of 101. Their mean age of 11.1 is greater than the mean of 10.1 for the entire class. This accounts for the negative correlations. Five of these older students were also involved in the special reading class which has a negative correlation with the criterion. The same reasoning accounts for this negative correlation as in the age correlations above.

The Pearson correlation shown in Table 2 indicated the degree of relationship (r) between all possible pairs of variables. In order to determine the degree of relationship between the criterion and all independent variable at one time, a multiple regression analysis was accomplished. The degree of relationship is known as the coefficient of multiple correlation (R), and is an index of how far a whole series of variables (e.g., tests) will serve to predict another variable (the criterion).

Raymond Cattell helps to make this clear through this example:

If one has, say, seven variables and one wishes to discover the combined power of these to predict a measure of achievement, the appropriate statistic is R . Unlike the ordinary or "zero order" coefficient of correlation, the coefficient of multiple

correlation can never be negative, but ranges between 0, showing no prediction, and +1, showing perfect prediction. It operates by forming a weighted composite of the predicting variables. The method of multiple correlation finds this maximally predictive combination and expresses the relation between this composite and the criterion.¹

According to Table 2, the variables that most accurately predicted the criterion was the pretest scores (.96) and intelligence quotient (.94). The addition of the other "weighted" variables produced a multiple correlation coefficient (R) of .99 and an R^2 of .98 (1.0 indicates perfect correlation).

USE OF MODEL EQUATION

The beta "weights" to be used with the equation were found through computer use of correlation data and are: pretest .60, I.Q. .39, educational level of father .69, educational level of mother .66, age -.45, sex -3.28, and special reading class -.30, and a constant of -8.04. Each of the beta "weights" is then multiplied times its appropriate variable. The products of all variables and weights are added to a constant to arrive at the posttest score (criterion). The following example should help clarify. A male student (X_6) has a pretest score of 40(X_1), an I.Q. score of 100(X_2), educational level of parents is twelve years for each (X_3 and X_4), chronological age is

¹Raymond B. Cattell and H. J. Butcher, The Prediction of Achievement and Creativity (New York, New York: Bobbs-Merrill Co., 1968, p. 161.

10(X_5), and had no special reading class (X_7). Using the equation $Y = X_1b_1 + X_2b_2 + X_3b_3 + X_4b_4 + X_5b_5 + X_6b_6 + X_7b_7 + k$ and substituting numerical data, a criterion should equal $40(.60) + 100(.39) + 12(.69) + 12(.66) + 10(-.45) + 2(-3.28) + -8.04$, which has a sum of 60. Under normal conditions then, his posttest reading score should be near 60. The same procedure would be followed for each student in the class or classes under evaluation.

Although the equation and "weights" are used to predict individual achievement, that is not intended as its primary purpose. Individual evaluation made with the use of the equation will not consider erratic progress due to influences beyond the control of the pupil or the teacher and will, therefore, be unfair. Nor is it intended as a comparison of the teaching abilities of individual teachers. A much more suitable use would be made if it were employed as an aid to a more complete, general evaluation. An example of such an evaluation might be to determine the effectiveness of new methods in a classroom or different programs that had been initiated throughout a school or system.

SUMMARY

Having established the variable information required and the model equation to be used, tabulation of variable data was made. Pearson paired analysis was used to find the relationship between all

pairs of variables. Computations were facilitated by the Sigma Seven computer at Montana State University. Once all paired correlations were found, a multiple regression analysis was accomplished to determine the degree of relationship between the criterion and all independent variables at one time. The resulting multiple correlation coefficient was .99 on a scale where 1.0 indicates perfect correlation. With the use of paired correlation, multiple correlation, and the model equation, appropriate "weights" were computed with the facilities of the computer center at Montana State University. The individual "weights" found for each variable then could be multiplied times the variables and their products added to a constant. The resulting sum is the predicted posttest score (criterion). The same procedure would be followed for each student in a class. From the information gathered as a result of a comparison of the actual post test score and the predicted posttest score, some evaluations can be made. These comparisons, however, are more suitable to large groups and general evaluation. An effective use of it would be made where the reading progress of a class using new methods were compared to a class using a more traditional approach. It is an inappropriate device for individual evaluation in that it does not have the capacity to consider outside influences on progress as do subjective, human evaluations.

Chapter 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

SUMMARY

A major question in education today as well as centuries ago is: what factors of an educational program bring about the most desired results with the least amount of time and expense? Some authorities feel that clearly stated behavioral objectives be used and evaluation geared to them. However, affective behavior such as appreciation is nearly impossible to measure objectively. Evaluation by "reading specialists" might prove very satisfactory, but they are spread far too thin for thorough work. Accountability in education is relatively new but growing rapidly. Objective results are being required by the community. They want information on educational growth in light of ability and past performance so that they can help determine which programs are inept and which show the most promise.

The need for more efficient and effective methods of evaluation is evident not because schools are failing but because educators are continually striving for improvement. The purpose of this study was to develop an aid or model which may be of some assistance in objectively evaluating reading progress in light of past performance and ability. In general, the model is in the form of an equation in which the variables (i.e., I.Q., past reading scores, etc.) would be correlated with each other and the criterion. Computations arrived at through the

use of a computer were assigned as appropriate "weights" to be multiplied times each variable. The sum of these products will give the expected amount of progress that any particular student should make in a one year period. Past performance and expected progress use the raw reading score as the means of measurement.

Having established the need and possible pattern for construction of an evaluatory aid, the collection of data was initiated. Variable data was collected from eighteen fourth grade students from the Wing Elementary School, Wing, North Dakota.

The variables involved in the development of the model were: pretest scores, intelligence quotient, parents' educational level, age, sex, and participation in special reading classes.

The null hypothesis held was that reading comprehension progress cannot be predicted through the use of the variables $X_1 \dots X_7$. The alternative was that prediction could be made with the use of the seven variables. A multiple regression equation was constructed which involved the variables (X) and appropriate "weights" (b) which were to be found through data analysis. The equation took the following form:

$$\text{(Criterion)} \quad Y = X_1b_1 + \dots X_7b_7 + k$$

An "F" test was devised which could be used for rejection or acceptance of the hypotheses.

The next step was analysis of the tabulated data. Pearson paired analysis was used to find the relationship between all pairs of

variables. Computations were facilitated by the Sigma Seven computer at Montana State University. Once all paired correlations were found, a multiple regression analysis was accomplished to determine the degree of relationship between the criterion and all independent variables at one time. The resulting multiple correlation coefficient was .99 on a scale where 1.0 indicates perfect correlation. With the use of paired correlation, multiple correlation, and the model equation, appropriate "weights" were computed with the facilities of the computer center at Montana State University. The individual "weights" found for each variable then could be multiplied times the variables and their products added to a constant. The resulting sum is the predicted post test score (criterion). The same procedure would be followed for each student in a class. From the information gathered as a result of a comparison of the actual post test score and the predicted post test score, some evaluations can be made. These comparisons, however, are more suitable to large groups and general evaluation. An effective use of it would be made where the reading progress of a class using new methods were compared to a class using a more traditional approach. It is an inappropriate device for individual evaluation in that it does not have the capacity to consider outside influences on progress as do subjective human evaluations.

CONCLUSIONS

Conclusions based on the review of literature and on analysis of the data are summarized in the following statements:

1. Processes involved in accountability in education are growing rapidly as a result of greater concern among parents, educators, and the public in general.

2. Adequate instruments that would aid in objective evaluations appear to be very much in need.

3. Affective behavior is difficult to measure by any means and definitely does not lend itself to objective evaluation.

4. A significant positive correlation existed between: (a) pretest score and--intelligence, parents' educational level, post test score; (b) intelligence quotient and--mothers' educational level, post test score; (c) fathers' educational level and--mothers' educational level, post test score; (d) mothers' educational level and post test score; (e) chronological age and special reading class; and (f) male sex and special reading class.

5. A significant negative correlation existed between: (a) pretest score and--age, special reading class; (b) intelligence quotient and--age, special reading class; (c) age and post test score; and (d) special reading class and post test score.

6. Reading achievement appears to be predictable through the use of a "weighted" equation. The equation involves use of the

following predictor variables: pretest score, intelligence quotient, age, sex, parents' educational level, and participation in special reading class.

7. Care must be taken to consider outside influences that would greatly affect the processes of learning when using the model for individual evaluation.

RECOMMENDATIONS

The following recommendations are made:

1. Further investigations should be made with larger sample size, involving many teachers and covering larger geographic areas. Length of teaching experience and educational level of teacher should be included as variables in a more comprehensive study.

2. The variable weights and model arrived at through this study should be used to predict the reading progress in the fourth grade class in the same school. The predicted progress should then be compared to the actual progress made during the year. This should be a means of establishing some degree of reliability.

Walter N. Durost, an outstanding educator and reading specialist, said, "The quality of education is directly related to the effectiveness of the evaluative program. Properly done it should improve reading ability through improvement of instruction."¹

¹Walter N. Durost, "Evaluation of Reading Progress," The Reading Teacher, XXIV (January, 1971), 290.

If this investigation only plays a very small role in the formation of a more effective evaluation of reading programs, its purpose will have been fulfilled.

LITERATURE CITED

LITERATURE CITED

- Bernabei, Raymond. "Instructional Accountability," Speech presented to the 104th Annual Convention of American Association of School Administrators. Atlantic City, New Jersey, February 11, 1972.
- Bond, Guy L. and Robert Dykstra. "The Cooperative Research Program in First Grade Reading Instruction," Reading Research Quarterly, 2:5, Summer, 1967.
- Buros, Oscar K. (ed.). Fifth Mental Measurement Yearbook. 5th ed. Highland Park: Gryphon Press, 1959.
- Cattell, Raymond B. and H. J. Butcher. The Prediction of Achievement and Creativity. New York: Bobbs-Merrill Co., 1968.
- Durost, Walter N. "Accountability: The Task, the Tools, and the Pitfalls," The Reading Teacher, 24:291, January, 1971.
- _____. "Evaluating Reading Progress," The Reading Teacher, 24:270, January, 1971.
- Hern, Norman E. and David Iwamoto. "Evaluation is a Full Time Job," American Education, 5:18, April, 1969.
- Landers, Jacob. "Accountability and Progress by Nomenclature: Old Ideas in New Bottles," Phi Delta Kappan, 54:539, April, 1973.
- Lessinger, Leon M. Every Kid a Winner: Accountability in Education. New York: Science Research Associates, Inc., 1970.
- Misner, Paul J. Elementary School Administration. Columbus: Charles E. Merrill Books, Inc., 1963.
- Niedermeyer, Fred and Stephen Klein. "An Empirical Evaluation of a District Teachers' Accountability Program," Phi Delta Kappan, 54:100, October, 1972.
- Rauch, Sidney J. "How to Evaluate a Reading Program," The Reading Teacher, 24:244, December, 1970.
- Sutton, Marjorie Hunt. "Accountability and Teaching Reading," The Reading Teacher, 25:528, March, 1972.
- Woodington, Donald D. "Accountability from the Viewpoint of a State Commissioner of Education," Phi Delta Kappan, 54:95, October, 1972.