

THE UNIVERSITY OF MANITOBA

A STUDY IN THE PREDICTION OF TEACHING EFFICIENCY

BEING A THESIS SUBMITTED TO THE COMMITTEE

ON POST-GRADUATE STUDIES IN PARTIAL

FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE

OF

MASTER OF EDUCATION

BY

JOHN MICHAEL KOCHAN

CANORA, SASK.

APRIL, 1956



A STUDY  
IN THE PREDICTION OF TEACHING EFFICIENCY

BY  
JOHN MICHAEL KOCHAN

ABSTRACT

This thesis arose out of a segment of a comprehensive study undertaken by the Faculty of Education, University of Manitoba, for the purpose of critically examining and perhaps improving the faculty's teacher-training program. The core of this segment consisted of the administering of a battery of tests to newly-enrolled candidates in Education I. The original battery, which was administered in the fall of 1953 and to which other tests have been added later, included seven tests: the ACE Psychological, Cooperative Mathematics, Watson-Glaser, the Michigan Vocabulary Profile, Essentials of English, the SRA Reading Record, and the Minnesota Multiphasic Personality Inventory. The measures from these tests were used primarily for educational guidance purposes during institutional training. The faculty conceived the idea that these same measures could perhaps be used for purposes of selective admission and selective retention of potential teachers. The need for assessing the potential value of these measures for prediction of teaching efficiency suggested this thesis.

In order to pave the way for further and more extensive research in this field, locally, the thesis problem assumed a threefold character. First, in order to gain some insight into the extent of research already done in the field, to appraise the problems involved, and to arrive at some consensus of opinion about effective techniques and procedures, previous research in the field had to be reviewed and evaluated. Second, based on the findings of the review and adjusted to the local situation, techniques and procedures for securing measures of prediction variables, securing ratings of teaching performance, classifying and recording data, and statistically treating the data had to be set up. Third, as a test of the effectiveness of the proposed techniques and procedures a "pilot study" had to be run, to form a basis for recommendations for further research.

The review of previous investigations constitutes the bulk of the thesis. Directly or indirectly, approximately five hundred studies have been evaluated, revealing a multiplicity of approaches and a multitude of problems. Fortunately, despite the many disagreements over methods and findings, it was possible to select the more effective techniques and procedures and with modifications to embody them into the pattern of the pilot study.

Using a sample of thirty-five teachers, in their second year of teaching, who had graduated from Education I two years earlier, the services of supervisors (principals,

inspectors, faculty members, etc.), and a simplified categorizing scale, a dichotomy of twenty-eight successful and seven unsuccessful teachers was obtained. Five prediction variables were secured from the measures on the battery of tests administered to the subjects two years earlier. These five were: scores on the ACE test, scores on the Coop Math test, scores on the Watson-Glaser Appraisal, composite English Usage scores (Michigan Vocabulary, Essentials of English, SRA Reading), and the total T-scores from nine scales of the MMPI. Both the biserial coefficient technique and the discriminant equation method were then applied in order to determine the predictive strength of the variables, individually and in combinations.

Within the limitations of this somewhat exploratory study, particularly within the limits of the small sample used, the proposed techniques and procedures were applied with considerable success. Although no sweeping generalizations are warranted, the findings of the pilot study are revealing, encouraging and quite consistent with those of previous studies in the field. Individually, the ACE measures of intelligence and the composite measures of English Usage show closest relationships to teaching success, with a coefficient of .507 significant at the  $3\frac{1}{2}\%$  level for the former, and a coefficient of .458 significant at the 5% level for the latter. Some relationships, significant at the 9% or 10% level, are found for measures on the Watson-

Glaser and the Coop Math tests. No significant relationship is found between teaching success and measures of personality, as estimated by the MMPI. No highly significant relationship is found between teaching success and any combination of variables. Maximum separation in the dichotomous variable, significant only at the 9% level, was obtained from a combination of ACE and English Usage measures. The significance levels for other combinations ranged from 10% and up. Much of this, of course, could be attributed to the limited sample and to the relatively high positive intercorrelations between the more potent variables.

Irrespective of the findings of the pilot study, the techniques and procedures used were judged practical and powerful tools for conducting studies of this nature in the local situation. On that basis definite recommendations are laid out for a continuation and extension of this type of research for the purpose of furthering the efforts of the teacher-training study undertaken by the faculty.

Approved, March , 1956.

~~Dr. H. L. Stein, Adviser~~

TABLE OF CONTENTS

	Page
LIST OF TABLES . . . . .	iv
ACKNOWLEDGMENT . . . . .	vi
 Chapter	
I. INTRODUCTION . . . . .	1
A. The Problem Setting	
B. The Specific Problem	
C. The Thesis Pattern	
II. REVIEW OF PREVIOUS STUDIES IN THE FIELD . . . . .	10
A. General Aspects and Problems of Measurement and Evaluation of Teaching Efficiency	
B. General Aspects and Problems of Prediction of Teaching Efficiency	
C. Conclusions and Generalizations	
III. DESCRIPTION OF THE SAMPLE . . . . .	60
IV. PROCEDURES AND TECHNIQUES IN DATA- GATHERING . . . . .	74
A. Securing Measures of Predic- tion Variables	
B. Securing a Rating (Categor- ization) of Teaching Performance	
C. Classifying and Recording of Data	
D. Important Limitations and Basic Assumptions	
V. PROCEDURES AND TECHNIQUES IN INTERPRETING DATA . . . . .	90
A. Biserial Coefficient Technique	

	Page
B. Discriminant Analysis Technique	
VI. FINDINGS AND CONCLUSIONS . . . . .	107
A. Interpretation and Value of Findings	
B. Recommendations for Further Research	
BIBLIOGRAPHY . . . . .	121
APPENDIX . . . . .	126
1. Copy of "California Statement of Teaching Competence"	
2. Proposed individual filing card for teacher-training study	
3. Copy of scale for categorizing teaching efficiency	
4. Copy of letter accompanying scale for categorizing teaching efficiency	
5. Copy of follow-up letter from Dean Scarfe's office to the "raters."	

## LIST OF TABLES

Table		Page
1.	List of Teacher Qualities Used by Mead . . . . .	20
2.	Criteria for Selection Used at the Syracuse University . . . . .	21
3.	Items Used by Herbert I. Von Haden in a Five-Point Rating Scale . . . . .	22
4.	Aspects of Teachers or Teaching Used by Barr in His Summary, 1948 . . . . .	23
5.	Raw Data for the Description of the Sample Background Information . . . . .	62
6.	Types of Schools Taught in by the Subjects at the Time of Rating . . . . .	65
7.	Major Subjects Taught by the Teachers During 1955-56 . . . . .	65
8.	Data on the Sample from Four Areas of Ability - General Scholastic, Mathematics, Critical Thinking and English Usage . . . . .	67
9.	Comparison of Measures of Central Tendency and Variability for Sample Used in Pilot Study with Those of the Whole Group of Fifty-four and with Some Accepted Norms . . . . .	69
10.	T-scores and Combinations of T- scores for the Sample from Scales on the MMPI . . . . .	70
11.	Five Sets of Measures Used as Prediction Variables in the Pilot Study . . . . .	75
12.	Individual and Composite Categorical Ratings of Teaching Efficiency of the Sample . . . . .	83
13.	Prediction Variables and Basic Statistical Measures in Dichotomy . . . . .	84

Table	Page
14. Intercorrelation Coefficients Among the Five Prediction Variables for the Whole Sample of 35 Subjects (by the Raw-Score Method) . . . . .	86
15. Analysis Data for Application of Biserial Correlation Formula . . . . .	91
16. Biserial Correlation Coefficients and Their Significance . . . . .	93
17. Analysis Data from Raw Scores of Prediction Variables for the Whole Sample Showing Sums, Squares of Sums, Sums of Squares and Sums of Cross-Products . . . . .	95
18. Analysis Data from Deviation Scores of Prediction Variables for the Whole Sample Showing Sums of Squares and Sums of Cross-Products . . . . .	97
19. Values of Coefficients as Obtained by Solving Various Sets of Discriminant Equations . . . . .	102
20. Multiple Biserial Correlation Coefficients of Different Combinations of Variables and Their Relative Significances . . . . .	103
21. The 20, 10, 5 and 1 Percent Values of "F" . . . . .	104
22. Relative Effectiveness of Individual Variables in Their Contribution to the Total Predictive Value . . . . .	105

## ACKNOWLEDGMENT

This study was undertaken at the suggestion of Dr. H. L. Stein, Professor of Educational Psychology, who was the author's principal adviser throughout the study. His kind assistance and much-needed guidance have been most valuable in the execution of the study.

The writer wishes also to acknowledge and extend thanks to the other members of his thesis committee, Professor B. R. Whiting, Dr. W. C. Lorimer, for their assistance, and particularly to Dean N. V. Scarfe, Faculty of Education, without whose cooperation and timely suggestions this study would not have been possible.

The writer further wishes to acknowledge the help received from the offices of the Registrar, of the Chief Inspector, and of the Administration Branch of the Manitoba Department of Education, and of the Manitoba Teachers' Society, in locating the names and addresses of the subjects and of those who assisted in rating the teaching performance of the subjects.

Last, but not least, the services of the principals, inspectors, superintendents, and faculty members in rating the subjects' teaching efficiency, to whom is due whatever of value may appear in the results of this study, are gratefully acknowledged.

John Michael Kochan.

## CHAPTER I

### INTRODUCTION

#### A. The Problem Setting

With the recent general increase of critical interest in educational philosophies and practices, it is not surprising to find, as a review of recent literature bears out, that North American universities and teachers' colleges have become increasingly conscious of the importance of their teacher-training programs. It is generally recognized that any proposed improvements in education must first of all be reflected in the training of teachers. Faculties of Education have therefore been adopting a more scientific attitude in evaluating their teacher-training programs. Such attitudes have found expression in six different areas of study. These are:

- |                             |  |
|-----------------------------|--|
| (1) pre-training guidance;  | (4) employment;                                    |
| (2) selection;              | (5) in-service improvement;                        |
| (3) institutional training; | (6) retention, promotion<br>and general follow-up. |

The teacher-training study undertaken by the faculty of education of the University of Manitoba, in the area of selection, perhaps typifies that renewal of interest in an important but difficult problem. This thesis deals with but one aspect of

a comprehensive attempt at tackling selection, though not necessarily the most significant one.

Aiming eventually at a more selective admission, more selective retention and sounder training, the faculty, some three years ago, along with other methods of attack, initiated a teacher-training study whose immediate concern was the study of prediction of teaching efficiency. The core of the study consists of a battery of tests administered upon enrolment to prospective teacher candidates in an effort to identify some source of data that could be useful not only in guiding the candidates through their institutional training but also in choosing from prospective candidates those most likely to be effective in the teaching profession. The original battery, to which other tests are being added from year to year, was administered in the fall of 1953 and consisted of seven tests, namely: (1) the College Edition of the American Council on Education Psychological Examination (ACE); (2) the Cooperative General Achievement Test, Form XX, Test 111 - Mathematics; (3) the Watson-Glaser Critical Thinking Appraisal; (4) the Michigan Vocabulary Profile Test; (5) the SRA Reading Record; (6) the Essentials of English Test, Form A; and (7) the Minnesota Multiphasic Personality Inventory (MMPI).

This thesis was conceived in conjunction with this

particular aspect of the teacher-training study. It is hoped that in reviewing previous studies in the field of prediction of teaching efficiency, in provisionally setting up suitable statistical procedures for the treatment of data that may be used for prediction, and by testing the predictive value of the data available from the above-mentioned battery of tests, the faculty will be in a better position to guide their teacher-training study along most profitable channels. It is hoped, too, that the thesis will be of some assistance in setting up more intensive studies, if deemed necessary, for the re-examination of the existing teacher-training program, which the faculty is anxious to maintain at the highest possible level of efficiency.

#### B. The Specific Problem

The scope of this thesis is indicated by the following outline:

1. To evaluate critically previous investigations in the field of prediction of teaching success in order to determine, for the purpose of this study and other studies for which this is the precursor, just where current studies will fit into the levels of generalizations of reflective thinking with respect to the implications of the project that is undertaken. More specifically, such a review should

help to answer the following questions:

- (a) Why predict teaching efficiency?
- (b) Can teaching efficiency be predicted?
- (c) Can the efficiency be predicted as early as the freshman year?
- (d) What criteria of teaching success have been used as a basis in judging the quality of teaching?
- (e) What factors seem to condition success in teaching?
- (f) Of what relative value are subjective and objective data in predicting teaching efficiency?
- (g) What characteristics of prospective teachers are related to teaching success?
- (h) What techniques have been found useful in gathering data and in interpreting results?
- (i) What conditions favor valid and reliable prediction?
- (j) What conditions limit prediction in pre-service selection?
- (k) What teacher-selection programs have been attempted and with what success?
- (l) What major research programs have been undertaken recently and with what success?
- (m) Does it seem worthwhile to continue with research in this field?

2. To set up proposed techniques for:

- (a) securing as objective a rating as possible of teacher performance in the field; this study will attempt to set up a procedure for obtaining a valid composite rating of experts in the field - principals, inspectors, supervisors, faculty members.

- (b) classifying and recording data from which predictive variables may be selected for statistical treatment. This study will limit itself to data obtainable from the battery of tests mentioned in part A.
- (c) statistical treatment of data in an effort to identify those factors which are most closely related to teaching success, and which may be used for predicting the teaching efficiency of prospective candidates for the profession.

3. To make a "pilot" study as an illustration of the use of those techniques recommended in (a), (b) and (c) above, and to test their effectiveness under local conditions.

4. In the light of the findings of the review of previous studies and of the pilot study, to make recommendations for further research in the field, particularly in the local situation, and for further progress in the teacher-training study undertaken by the faculty.

It should perhaps be reiterated that the emphasis of this study is not so much on the experimental aspect of the problem with a view to arriving at some sweeping generalizations concerning prediction of teaching efficiency but rather on the clearing away of some of the initial problems that usually surround a newly-initiated program in a locality. Thus the primary objective is that of setting up procedures and, in a limited sense, testing the effectiveness of the procedures set up. It will be noted that the selection of procedures and the recommendations will be closely related to the consensus of opinion arrived at in the findings of previous studies in the field and adjusted to the local situation.

Despite the fact that this study is primarily of a descriptive and exploratory nature, it is, nevertheless, of value to point out certain broad assumptions or hypotheses that underlie it. That it is important to know what characteristics of prospective teachers are related to success in the field is attested by the volume of studies in this area. These studies have met with varying degrees of success in terms of prediction and conflicting results have often appeared. It is believed that one of the reasons for this lack of uniformity in results is the complexity of the profession of teaching. Over a million people on this continent call themselves "teachers" and are engaged in hundreds of activities, and "success in teaching" must therefore be defined in numerous ways. It is postulated here, however, that in spite of the disagreement and confusion, as will be indicated later, there may yet be something in common to all "teachers" which causes them to be described as successful or unsuccessful. If this something could be identified at least in broad terms then the teacher-training programs can be more objectively re-examined and perhaps the processes of teacher recruitment, selection, placement, training and induction might be improved. Such is the importance of the problem and such is the anticipation for its value in initiating further studies in this locality.

### C. The Thesis Pattern

Having outlined the origin and the scope of the problem for this study, the writer, as the table of contents indicates, is devoting the second chapter of the thesis exclusively to a critical review of previous studies in the field of prediction of teaching efficiency. In addition to examining the findings of various studies, an effort will be made to arrive at some general consensus of opinion, in the light of local conditions, regarding the choice of a suitable criterion of teaching success, of factors related to teaching success and of measuring and evaluating devices. In the area of actual prediction the reader will be made aware of the many problems that beset research work in this field, of the many types of statistical methods employed with varying degrees of success, and of the many limitations that must be considered in any attempt at generalization. With the ultimate value of this study in mind, the reader's attention will be brought to some current research programs in the field and to some selective programs recently adopted by universities. General conclusions and inferences will be drawn out, on which much of the experimental portion of this study will be based.

The third chapter is devoted to a description of the sample used in the pilot study. Since the experimental portion of this study is designed for the purpose of testing the ef-

fectiveness of the proposed techniques it is important that the reader and subsequent research workers are made aware of any limiting characteristics of the sample used. Therefore, all possible data descriptive of the sample will be presented so that no generalizations will be made of the findings which are not warranted by the particular and somewhat selective sample used.

In Chapter IV the reader will find a description of the procedures used in obtaining and classifying the predictive variables, in arriving at composite scores, in securing ratings of teaching performance, in arriving at composite ratings, and in preparing the data in a form conducive to statistical treatment. Here, too, the subsequent research worker in this field will be able to sense the problems that must be faced in a study of this nature, to test the effectiveness of the methods employed, and to become cognizant of some of the major limitations that plague both the data and the adopted procedures.

Chapter V will present the procedures and techniques for treating the data statistically in order to estimate the predictive value of the variables used in the study. Not only will the findings of the pilot study be given here but also some indication of the effectiveness of the tested method in treating other variables that could be used in similar and more extensive future studies. Again, certain

limitations and assumptions will have to be recognized before any generalizations are extracted from this limited experimental study.

In Chapter VI, in which a summary and an interpretation of the findings of the pilot study will be presented, the reader will find not only a critical analysis of the results of the study but also an indication of the extent to which the results agree or disagree with findings of previous studies in the field.

The final chapter will make more specific the value of the findings of the whole study with some recommendations for further research in the local sphere as well as recommendations that may be of value to the faculty in the continuation of their teacher-training study from which this thesis originated.

## CHAPTER II

### REVIEW OF PREVIOUS STUDIES IN THE FIELD

#### A. General Aspects and Problems of Measurement and Evaluation of Teaching Efficiency

Measurement and evaluation of teaching efficiency must precede prediction of teaching success. When considering the former, three problem areas come to the fore:

1. Choice of criterion of teaching success;
2. Consideration of factors related to teaching success;
3. Selection of appropriate measuring and evaluating devices.

##### 1. A Criterion of Teaching Success

This is a basis used in judging the quality of teaching. Since teaching success is conditioned by many teacher traits and acquirements, the criterion is necessarily complex. In studies of this nature as distinguished from purely administrative procedures, three criteria have been used extensively:

- (a) Supervisors' ratings, or judgment of persons in the field (experts), such as superintendents, principals, departmental inspectors, supervisors, and teaching colleagues.
- (b) Pupil achievement.
- (c) Pupil evaluation of teaching efficiency.

Authorities are in quite general agreement that the judgment of experts, excluding teacher colleagues, is the best available criterion of teaching success. One of the most careful studies which extensively employed the judgment of experts was conducted by Sandiford and others<sup>1</sup> at the University of Toronto. Using the Spearman-Brown formula, they obtained reliability coefficients of .888 and .929 respectively for two groups of experts and of .945 and .899 respectively for two groups of other judges. Correlations between the ratings were .748 and .707 respectively, indicating validities satisfactory to the investigators. Subsequent similar studies have reported somewhat lower but mostly positive correlations.

Pupil achievement would seem to be a justifiable criterion of teaching success. However, its use is accompanied by numerous hazards, not the least of which is an answer to the question, "What achievement?" Answers are varied and include such items as information and knowledge, attitudes, appreciations, and skills. Further difficulties are created by the lack of agreement upon what information, what knowledge, and so on; the absence of valid and reliable instruments for measuring specified achievements; the possibility that pupil achievement as ordinarily measured is

---

<sup>1</sup>Peter Sandiford and others, Forecasting Teaching Ability, University of Toronto, Department of Educational Research, Bulletin No. 8, 1937, p. 93.

nearly valueless because it may be merely a measure of the efficiency with which the pupil retained factual information long enough to pass the tests; the lack of comparability between some of the measuring instruments and the recognized objectives of education; and the rather well-supported suspicion that the pupil's gain in at least information and knowledge is due more to his habits of study than the instruction offered by a teacher. Even if it were assumed that both the tangible and intangible outcomes of education may be evaluated by valid and reliable instruments of measuring pupil achievement, yet, as Magee<sup>1</sup> pointed out in his study, many questions remain unanswered. For example, to what extent can the changes observed in children be attributed primarily to the efforts of teachers who, after all, work in competition and cooperation with the home, the playgrounds, the church, the radio and other influences that exert impacts upon children? The studies which have employed this criterion have not produced significantly high correlations between pupil gains as measured by achievement tests and independent evaluations of teachers. Barr,<sup>2</sup> in his review of studies in the field prior to 1948, substantiates the foregoing observations and reiterates most of the above-mentioned difficulties

---

<sup>1</sup>Robert M. Magee, "Selection of Candidates for Teacher Education," Journal of Teacher Education, 3, September, 1952, 168.

<sup>2</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," Journal of Experimental Education, 16, June 1948, 295-306.

inherent in the use of this criterion.

Much attention has been focused upon the use of pupil evaluation as a criterion of teaching efficiency, but recently this criterion has received much adverse criticism. Critics maintain that this criterion, despite the objectivity of instruments used, is the least objective of the three. It is conceivable that much pre-training of the pupils would be required, that it would be difficult to remove tensions between pupils and teachers concerned, that it would be difficult for the pupils to give proper weighting to factors such as teaching loads, nature of subjects, and other factors which influence teaching in the eyes of the pupil. It is also conceivable that pupils' versions of teaching success will vary not only from pupil to pupil but also from time to time and from age to age within the same pupil. As an illustration, Witty<sup>1</sup> analyzed approximately 12,000 letters of pupils in grades 2 to 12 and found the four most frequently mentioned factors in describing the teacher who had helped them most to be: cooperative, democratic attitude; kindness and consideration for the individual; patience; and wide interest. Definitely, all of these factors are related to teaching effectiveness, but they refer to only one specific area of the teaching process and not necessarily

---

<sup>1</sup>Paul Witty, "An Analysis of the Personality Traits of the Effective Teacher," Journal of Educational Research, (1947), 40, 662-671.

the most important area at that. Of course, not all studies agree with this contention. One recent study of some significance, by Symonds,<sup>1</sup> using the Spearman-Brown formula, found a coefficient of reliability of .933 for this method of evaluation. In considering the merits of this criterion, it should not be overlooked that teachers, for obvious reasons, generally frown on the suggestion of being evaluated by their pupils.

The most unfortunate aspect of the choice of criterion of teaching success is the definite lack of agreement among the criteria. Consensus of opinion is to the effect that each of these seems to measure something entirely different in the complex teaching act. Studies by Lins,<sup>2</sup> Reed,<sup>3</sup> Von Haden,<sup>4</sup> and numerous others bring attention to

---

<sup>1</sup>Percival M. Symonds, "Characteristics of the Effective Teacher Based on Pupil Evaluation," Journal of Experimental Education, (June 1955), 23, 289-310.

<sup>2</sup>Leo Joseph Lins, "The Prediction of Teaching Efficiency," Journal of Experimental Education, (September 1946), 15, 59.

<sup>3</sup>Harold J. Reed, "An Investigation of the Relationship Between Teaching Effectiveness and the Teacher's Attitude of Acceptance," Journal of Experimental Education, (June 1953) pp. 277-325.

<sup>4</sup>Herbert I. Von Haden, "An Evaluation of Certain Types of Personal Data Employed in the Prediction of Teaching Efficiency," Journal of Experimental Education, (September 1946), 15, 83-84.

this lack of consistency. Barr, in his comments on the studies reported on in 1946, repeats: "Attention is called again to the lack of agreement among the criteria employed in the evaluation of teaching efficiency."<sup>1</sup> Then, again, in summarizing the studies in the field up to 1948, Barr states: "Apparently, the several criteria, such as pupil growth and achievement, supervisory ratings, pupil ratings, and tests of qualities commonly associated with teaching success, measure different things."<sup>2</sup> In his latest summary of work in the field, Barr, in 1955, repeats the same observation: "It was found that different criteria of teacher efficiency gave different results; that is, no significant correlation was found between pupils' ratings, supervisors' ratings, and ratings based on pupil growth and achievement."<sup>3</sup>

In terms of the foregoing discussion it seems reasonable to conclude that rating by experts (principals, inspectors, supervisors) is the best available criterion. To add to the validity of this criterion a composite rating of a number of ratings by different judges is desirable.

---

<sup>1</sup>A.S. Barr, "Summary and Comments," Journal of Experimental Education, (September 1946), 15, 99.

<sup>2</sup>A.S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," Journal of Experimental Education, (June 1948), 16, 195-226.

<sup>3</sup>A.S. Barr, "Measurement and Prediction of Teaching Efficiency," Review of Educational Research, (June 1955), 25, 261.

Barr, in his comments on the lack of agreement between criteria adds this suggestion in 1946:

With this thought in mind, more attention should probably be given to subjective rating. While the means of objectifying subjective ratings seem to fail to give improved results it would seem that the situation might be somewhat improved through careful definition of the aspects of teaching to be evaluated, through the careful selection and training of judges, and through the use of 2, 3 or more judges.<sup>1</sup>

## 2. Factors Related to Teaching Success

It is clear from the literature in the field that a solution to the problem of predicting teaching success is intimately connected with the discovery of reasons why teachers succeed or fail. At this point, it seems wise to observe that prior to the last three decades, the general conception of the measurement of teaching ability was a very narrow one. Teacher rating was largely an administration activity performed by supervisors in connection with the tenure and promotion of teachers in service. Today, as observed by some of our authorities,

Educators are seeking valid and reliable measures of teaching ability as a basis for: (a) the selection of teachers to be trained; (b) the evaluation of teacher-training programs; (c) the placement of teachers after training; and (d) a more scientific program for the improvement of teachers in service.<sup>2</sup>

---

<sup>1</sup>Barr, "Summary and Comments," op.cit., p. 99.

<sup>2</sup>T. L. Torgerson, "The Measurement and Prediction of Teaching Ability," Review of Educational Research, (June 1934), 4, 261.

Likewise, it should be noted that education is no longer regarded as merely an implement for transmitting cultural values from one generation to the next, and that the teacher is no longer regarded just as the director of learning, but also as the guide to growing children and the friend and counselor of pupils. As a member of a group of professional workers, he is expected to have qualities that are conducive to group work. He also has his responsibility as a neighbor and citizen in the community. The broadened concept of the teacher's role in society poses fresh problems of prediction, selection and education of a new order. Thus, whereas earlier studies emphasized factors or qualities within three areas: ability to transmit knowledge, grasp of subject matter, and the personality of the teacher, gradually, in later studies, more attention is given to factors within two additional areas: the teacher as a member of his profession, and the teacher as an organizer and administrator of school and community resources.

Again, both the earlier and the later studies show little agreement on the kinds and number of factors that should be included in the areas chosen. Lack of standardization in vocabulary is no doubt one of the reasons for this. Some rating scales, for example, use a staggering number of factors with varying degrees of success at telescoping them into distinct areas. It is beyond the scope of this thesis to go into any detail of describing the lists of factors used by the various studies. In the few pages that follow,

some typical attempts at selection of factors and grouping of these factors into areas are given. These include a list of qualities used by Mead in his study in 1929,<sup>1</sup> qualities proposed by the Selection Committee of Syracuse University,<sup>2</sup> eight items telescoped from forty-three and used in a rating scale by Von Haden.<sup>3</sup> Barr,<sup>4</sup> in summarizing the studies in the field up to 1948, uses a somewhat unique grouping of factors. He refers to these three categories as "ways of talking about teaching efficiency." They are

- (a) in terms of character and personality traits - qualities of the person;
- (b) in terms of desired competencies, performances, abilities to do; and
- (c) in terms of behaviour controls - knowledge, generalized skills, attitudes, interests and ideals.

To further illustrate this lack of agreement and the tendency towards confusion, Barr, in the same summary, found it necessary to list some 75 qualities for the purpose of

---

<sup>1</sup>A. R. Mead, "Questions of Merit in Good and Poor Teachers," Journal of Educational Research, (November 1929), 20, 257-258.

<sup>2</sup>Miss Verna White, "Selection of Prospective Teachers at Syracuse University," Journal of Teacher Education, (March 1950), I, 24.

<sup>3</sup>Herbert I. Von Haden, op.cit., p. 63.

<sup>4</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," op.cit., p. 204.

summarizing the studies, and some 200 factors for summarizing the use of rating scales in the studies. As a final illustration of the diversity of opinion in the field and the staggering number of qualities that can be listed, a copy of the California Statement of Teaching Competence prepared by the California Council on Teacher Education and distributed locally by the Manitoba Teachers' Society is given in the Appendix. No useful purpose can be served by repeating any further lists used.

Despite the lack of agreement in this field, experience and observation indicate that certain human characteristics and competencies are related to success to such an extent that we would expect prospective teachers to possess them to a reasonable degree. Nearly every factor which it is thought may condition success in teaching has been studied, but the investigations have brought us very little closer to any agreement on a particular list of factors. A reason for this, in addition to those given earlier, is the lack of any valid and reliable criteria of instructional efficiency. It is not possible to identify the factors prior to the determination of a criterion. The relative importance of these factors as revealed by previous studies will be dealt with in the review of studies relative to actual prediction, which is the central purpose here, rather than simply evaluation.

TABLE 1

LIST OF TEACHER QUALITIES  
USED BY MEAD<sup>1</sup>

- |                                     |                                     |
|-------------------------------------|-------------------------------------|
| 1. Cleanliness in person and habits | 32. Sympathy                        |
| 2. Cooperation                      | 33. Tactfulness                     |
| 3. Courage                          | 34. Vigor                           |
| 4. Decision                         | 35. Vivacity                        |
| 5. Enthusiasm                       | 36. Voice                           |
| 6. Expressive countenance           | 37. Address                         |
| 7. Good follower                    | 38. Appearance                      |
| 8. Good loser                       | 39. Good mental and physical health |
| 9. Good mixer                       | 40. Sense of justice                |
| 10. Idealism                        | 41. Power                           |
| 11. Industry                        | 42. Taste                           |
| 12. Initiative                      | 43. Judicial sense                  |
| 13. Intelligence                    | 44. Love                            |
| 14. Leadership                      | 45. Effort                          |
| 15. Open-mindedness                 | 46. Kindness                        |
| 16. Loyalty                         | 47. Self-reliance                   |
| 17. Manner                          | 48. Self-confidence                 |
| 18. Modesty                         | 49. Promptness                      |
| 19. Neatness                        | 50. Faith                           |
| 20. Optimism                        | 51. Patience                        |
| 21. Persistence                     | 52. Friendship                      |
| 22. Personality likeable            | 53. Interest in others              |
| 23. Poise and carriage              | 54. Politeness                      |
| 24. Presence commanding attention   | 55. Cheerfulness                    |
| 25. Reliability                     | 56. Firmness                        |
| 26. Reserve and dignity             | 57. Vitality                        |
| 27. Adaptability                    | 58. Fairness                        |
| 28. Alertness                       | 59. Integrity                       |
| 29. Sense of order                  | 60. Self-control                    |
| 30. Sense of humor                  | 61. Resourcefulness                 |
| 31. Sensitive to social proprieties | 62. Good nature                     |
|                                     | 63. Reverence                       |

---

<sup>1</sup> Mead, op.cit., p. 258.

TABLE 2

CRITERIA FOR SELECTION USED AT THE  
SYRACUSE UNIVERSITY<sup>1</sup>

1. Those who would teach should have the physical stamina necessary for long hours of strenuous physical and mental activity, should be free from contagious and infectious diseases, and should exhibit physical vitality.
2. Teachers should be emotionally stable - should be able to resolve their conflicts rather than allow them to tone and disturb their everyday activities.
3. A high degree of social competence is important. Contributing factors are ability to cooperate with others, social graces, and personal appearance.
4. Those who would teach should be academically competent. More important than what one learns is evidence of ability to so order one's mental activities that he can learn that which he sets out to learn.
5. The mental abilities of teachers should be such as to render them capable of civic and educational leadership among professional men as well as laymen.
6. Teachers should have an abiding interest in reading and discussing everyday occurrences in fields such as politics, economics, international affairs, social relationships, literature, music, art, science, sports, and movies.
7. Prospective teachers should be sufficiently skilled in oral and written expression that they may accurately set forth their own thoughts and aid others to develop the ability to do likewise.
8. Those desiring to teach should be capable of understanding and extending the psychological, philosophical and social bases of education.

---

<sup>1</sup>Miss Verna White, op.cit.

TABLE 3

ITEMS USED BY HERBERT I. VON HADEN IN A  
FIVE-POINT RATING SCALE<sup>1</sup>

1. Adaptability - ability and willingness to make the necessary adjustments to varying conditions and situations, open-mindedness, suggestibility, flexibility.
2. Considerateness- concern for the feelings and well-being of others, sympathy, understanding, cooperation, unselfishness, patience, helpfulness.
3. Energy - habitual readiness for effective action, force, vigor, drive, vitality, endurance.
4. Initiative - capacity for taking hold of things and assuming leadership, originality, resourcefulness.
5. Professional judgment - wisdom in the selection of appropriate course of action, techniques and materials; discretion in dealing with others; foresight; prudence.
6. Social adequacy - ability to attract others and to make easy and enjoyable contacts, adjustment to one's fellows, cheerfulness, friendliness, pleasantness.
7. System of values - one's philosophy of life, including ethical, social and educational values; eagerness to succeed; goals in life; ambitions; motivations; interests.
8. Work habits - industriousness, efficiency, perseverance in overcoming difficulties, accuracy, ability to organize, general evaluation as a prospective teacher.

---

<sup>1</sup>Von Haden, op.cit.

TABLE 4

ASPECTS OF TEACHERS OR TEACHING USED BY  
BARR IN HIS SUMMARY, 1948

I. Personal Qualities

- A. Teaching aptitude
- B. Resourcefulness - originality, versatility, imagination, etc.
- C. Intelligence - foresight, judgment, insight, common sense, etc.
- D. Emotional stability - poise, self-control, sobriety, loyalty, etc.
- E. Considerateness - kindness, friendliness, sympathy, tact, etc.
- F. Buoyancy - optimism, enthusiasm, carefreeness, wittiness, etc.
- G. Objectivity - fairness, impartiality, open-mindedness, etc.
- H. Drive - physical vigor, perseverance, motivation, ambition, etc.
- I. Dominance - self-confidence, decisiveness, independence, etc.
- J. Attractiveness - dress, physique, personal magnetism, posture, etc.
- K. Refinement - good taste, modesty, culture, morality, polish, etc.
- L. Cooperativeness - geniality, generousness, responsiveness, etc.
- M. Reliability - accuracy, dependability, punctuality, honesty, etc.
- N. Personality, General.

II. Competencies (abilities to do)

- A. As a director of learning - skill in identifying pupil needs, setting and defining goals, creating favourable mind sets, choosing learning experiences, following the learning process, using learning aids, teacher-pupil relations, appraising pupil growth and achievement, management instruction.
- B. As a counselor and friend of pupils.
- C. As member of a profession.
- D. As member of a community.

TABLE 4 (continued)

III. Effects of teacher leadership

IV. Behaviour Controls

- A. Knowledges - of the subject matter taught or activity directed, of child behaviour and development, of professional practices and techniques; general cultural background; scholarship, grade point average.
- B. Generalized skills - skill in problem solving; work habits; skill in human relationships; skill in the use of language (speech, reading, English usage).
- C. Interests, attitudes and ideals - interest in pupils; interest in subject or activity; interest in community; social attitudes; professional attitudes; efforts towards self-improvement; general interests; interests in extra-curricular activities.
- D. Health
- E. Morale
- F. Status facts - age, height, weight, training, experience, sex, salary, recommendations, photographs, socio-economic status, tenure (in present position), applications.

Just how much agreement is there in this field? Barr,<sup>1</sup> in reviewing the studies up to 1946, points to a somewhat general agreement on the qualities essential to successful participation at the college level in the areas of: teaching ability, scholarship and scholarly ability, experience, and personal qualities as being of primary importance, while standing in the profession, public and community services, and membership in learned societies as being of secondary importance. An extensive study by a graduate student at the University of Manitoba concludes with this note:

Possibly the only noticeable agreement is in some of the areas of the qualities chosen. The areas are those which emphasize: techniques or means of instruction, the amount a teacher knows about the subject or subjects, the interest of teachers in their profession and their attitudes to others in the profession, a teacher's relation with the community, and all personal qualities such as health, speech, manners, vitality, etc.<sup>2</sup>

Pointing to any agreement closer than the foregoing statements bear out would be extremely precarious. In all his reviews of studies in this field, Barr has consistently shown his disappointment in the results of research, and, at the same time, he has always pointed to the need for

---

<sup>1</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency," Review of Educational Research, (1946) 16, 203.

<sup>2</sup>R. H. A. Gray, "Teacher Evaluation," p. 87. Unpublished Master's thesis, Department of Education, University of Manitoba, 1954.

a more adequate listing of qualities considered essential to teaching efficiency and the need for an adequate definition of the several qualities used, which might lead to a better agreement and better results.

### 3. Measuring and Evaluating Devices

Evaluation devices, irrespective of the degree of objectivity claimed for them, may be classified under two headings:

- (a) those which are used directly in the classroom to judge teachers and their teaching; and
- (b) those which are used anywhere to judge teachers and their capabilities indirectly.

Among the former are: check lists and rating scales, stenographic reports, diaries, anecdotal records, recording devices, and measures of accomplishment after exposure to learning situations. Among the latter are: interviews, various devices to determine characteristics and capabilities of teachers, and all data referring to training, experience, health, background, etc. Both types of evidence are required for a complete picture of the teacher's efficiency in the complex teaching act.

Barr, Burton and Brueckner prepared a list, classified according to the nature of the instruments or procedures. The list reads as follows:

#### 1. Check Lists:

##### A, Question check-lists

- B. Activity check-lists
2. Other Types of Records and Recording Devices:
    - A. Written records - stenographic reports, diary records, anecdotal records
    - B. Mechanical Recording Devices - time-recording devices, sound recording, sound motion pictures
    - C. Personal Data Records - records of training (institutional and non-institutional), records of experience professional and non-professional)
  3. Rating Scales:
    - A. Point scales
    - B. Quality scales
    - C. Diagnostic scales
    - D. Graphic scales
    - E. Human scales
    - F. Conduct scales
  4. Tests:
    - A. Growth and achievement tests
    - B. Intelligence tests
    - C. Tests of teaching aptitude
    - D. Tests of character, personality, etc.
  5. Interviews, questionnaires, and inventories
  6. Measures of Pupil Growth<sup>1</sup>

From this list alone, one may conclude that many different sorts of data-gathering devices are being employed

---

<sup>1</sup>Barr, Burton, and Brueckner, Supervision, New York: Appleton-Century Croft Inc., 1947, p. 342.

in studying teachers and teaching. A more comprehensive and up-to-date list of some 209 rating devices and other measuring instruments is given by Barr in his summary of investigations in the field prior to 1948.<sup>1</sup> Tests, rating scales and inventories predominate in the list. With so much overlapping in terminology, little can be accomplished by stating who used what device with what success. As yet no uniform set of expressions to describe the devices has been adopted. No two authors will use the same term to describe a similar device. Furthermore, variations from study to study in criteria of teaching success, in purpose of study, in factors related to teaching efficiency, in samples, in conditions, would make any attempt at further classification of little value. Barr, who is no doubt in the best position to pass judgment on the matter, makes, among other comments, the following observations in his summary referred to above:

1. The reliability of these various devices seems to be relatively high; their validities, relatively unknown.
2. The data are many times inconsistent.
3. The reliability of examiners should be known as well as that of the devices.
4. Many persons confuse the terms "measure" (collect data about) and "evaluate". Measuring is only one step in the much more complex activity of evaluation.

---

<sup>1</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," op.cit., pp.217-223

5. The collection of data relative to teachers and teaching might be facilitated if the number of aspects about which information is desired could be reduced.
6. It might help to have all of the instruments revalidated upon the same samples and under comparable conditions.
7. The different types of data-gathering devices appear to supplement each other in a manner to make multiple approach desirable. A combination of so-called objective and subjective devices appears to give, under current conditions, better results than either employed separately.

In concluding this portion of the review, perhaps more than a passing interest should be shown in Barr's seventh observation, above, about the relative merits of objective and subjective devices. Lins,<sup>1</sup> in his study in 1946, and Hale,<sup>2</sup> in his study in 1955, made similar observations. It should be noted that the comparison refers to objective devices as they are now available, and it may be a criticism of the level of objectivity claimed for them rather than of the comparative value of objectivity as such. However, in terms of progress or lack of progress in the field of objective instruments for measuring and evaluating teaching effectiveness, it seems advisable to use simplified subjective procedures rather than objective ones which are conducive to an overestimation of their objectivity. One way of simplifying sub-

---

<sup>1</sup>Lins, op.cit., p. 60.

<sup>2</sup>Peter P. Hale, "Isolating Objective Factors for the Teaching Profession," Journal of Educational Research, March 1955, pp. 497-507.

jective procedures, especially in the case of expert judging, for the purpose of prediction studies, is that of categorizing teaching performance rather than score-rating or ranking. The N. E. A. pamphlet<sup>1</sup> on rating lists recommends that more than one expert rating be used for the purpose of minimizing the error of subjective judgment inherent in the devices used.

B. General Aspects and Problems of Prediction  
of Teaching Efficiency

Having discussed the purpose of reviewing literature of past studies on the measurement and prediction of teaching efficiency, and having noted some of the important aspects and problems in teacher evaluation as revealed by past investigations, the author suggests the next logical step to be that of reviewing some of the important aspects of actually predicting teaching success. But before any attempt is made to crystallize the findings and views of the various studies, it seems appropriate to consider first the reasons for the continued interest in this field of education.

1. Why Predict Teaching Efficiency?

---

<sup>1</sup> National Education Association, Teacher Rating, Discussion Pamphlet No. 10, Washington: National Education Association, 1950, p. 3.

The basic reasons may be found in the list given by Syracuse University, offered as the philosophy behind their selection program:

- (a) Present-day teaching demands a superior person.
- (b) The University's primary obligation is to school children and their parents.
- (c) A teacher-training institution cannot afford to graduate students who are a mediocre advertisement and a threat to professional standards.
- (d) Quality is attracted by quality.
- (e) Training facilities must be developed on the basis of quality rather than quantity.
- (f) Many students accept guidance only when faced with a definite barrier.
- (g) Selection by indirection has proved inadequate.<sup>1</sup>

As has been pointed out earlier in this review, the general conception of the measurement of teaching ability has changed considerably within the last three decades, with the emphasis now on selection, evaluation and placement rather than on tenure and promotion.

However, within the last two decades, the problem has taken on a new twist. With an oversupply of teachers, at one time school systems were in a position to demand personnel of superior capacity, training, and apparent potentialities for teaching. Today, with a definite short-

---

<sup>1</sup>Miss Verna White, op.cit., p. 24.

age, many candidates of inferior quality and a low level of professional training have been admitted to the teaching field. Thus, it has become necessary to attract into the teacher-training institutions students who possess the qualities and capacities for successful teaching. It may be argued, then, in view of the shortage, that selection is not as urgent today, but it may be shown that selection requires even greater care now. The greater tenure that teachers enjoy makes it increasingly less likely that inferior teachers will be replaced after they have taught a few years. Competition has lost a great deal of its former sting. As pointed out by the Teacher Selection and Counseling Service Office of the University of California,

A selection program is important, not only to increase the quality of teaching personnel, but for two other reasons. Students with superior potential should be identified early and constantly encouraged. They are most likely to be drawn away from teaching into competing fields, yet from them should come the educational leadership of the future. In addition, the shortage of teachers makes it essential that we reach out to the other divisions of the university, to the junior colleges and down to the high schools to recruit young men and women of superior potential.<sup>1</sup>

Despite sporadic though serious attempts in that direction, it is unlikely that the picture in the near future will improve. The combined factors of great increases in

---

<sup>1</sup> MacLean, Gowan and Gowan, "A Teacher Selection and Counseling Service," Journal of Educational Research, May, 1955, p. 669.

pupil population without a comparative increase in teacher population and the trend in education toward higher training standards will no doubt add to the problem, but it should stimulate a growing interest in research in this field.

Without going into any greater detail, the argument that preservice selection of teachers is incompatible with democratic ideals may be refuted on the following grounds:

1. The social importance of the work of the teacher justifies selection.
2. Schools expect higher institutions which prepare teachers to have in operation plans of recruitment and selection.
3. An analysis of teacher supply-and-demand indicates the desirability of selection.
4. Preservice selection should not be discrimination but a form of very useful guidance and placement.
5. Selection is a step towards utilizing human resources where they will contribute most to the common good.

## 2. Problems of Prediction

If it is agreed that selection can be easily justified and that prediction is an important phase of a selection program, then, generally speaking, what progress has been made in this field of work? What are some of the main problems? What are the inevitable limitations? Prediction must be preceded by measurement and evaluation; therefore, problems of measurement and evaluation are also problems of

prediction. Attention will now be called to problems of prediction in addition to those mentioned in the preceding pages.

Since individual contributions of various studies to the total picture of research in this field will be dealt with later, only general aspects and problems will be touched on here. Since attempts at both measurement and prediction of teaching ability involve human relationships and social adjustments, and the concomitant "unpredictability" of human nature which is often baffling, a diversity of methods of attack and of findings is inevitable. As pointed out earlier, at least three different criteria have been used in the evaluation of teaching efficiency. Some have used a fourth, namely practice teaching marks and grades. Others have used these marks and grades as variables. Some studies have used somewhat objective measuring devices; some have relied mostly on subjective ones; others have tried both. The choice of predictive variables differs tremendously from study to study. Some use a limited number; others have used as many as 240 in one study. It is not surprising then to find clashes in the findings. Even the choice of statistical methods shows an interesting variety. Used most often with varying degrees of success are correlations of zero order, regression equations, factor analysis, and discriminant functions. Here are some

of the comments made by experts in the field, who have reviewed the findings of many surveys. Torgerson, in his summary, states,

No single measure reported has sufficient validity in terms of the criteria used to establish it as a valid measure of teaching ability. A great deal of time and energy has been expended during the past two decades in an attempt to determine the specific traits, qualities and teaching activities which are related to teaching ability. The results in terms of positive relationships are not highly encouraging. More objective techniques and refinement of statistical procedures have served in many instances to lower the relationships.<sup>1</sup>

Robert M. Magee, in his summary, makes this observation:

Research workers have, as yet, found no reliable objective means of appraising teacher success, nor have they found significant relationship between measurable traits and reputed teaching success. Teaching is a highly complex function, and those who succeed in it do so, not because they possess certain separate traits to a given degree, but rather because the total personality exerts a salutary influence upon those whom it touches.<sup>2</sup>

The teacher-selection committee at the University of California adds this,

The thorny question of what characteristics constitute the potentially good teacher lies at the root of much of our uncertainty. There is first the necessity to identify and analyze good teaching itself. In addition, there is the question of which of these qualities are the product of professional training, and which the

---

<sup>1</sup>Torgerson, op.cit., p. 266.

<sup>2</sup>Magee, op.cit., p. 168.

individual brings with him at the time he enters it. The latter are more generalized and basic and are, by their very nature, less specifically identifiable. Predicting the quality of a teacher's performance in a specific field three years hence, before he has been exposed to training, presents a rather different and quite complex task.<sup>1</sup>

Such is the general picture of findings of studies in the prediction of teaching efficiency. More specific observations will be made later. At this moment the question that arises is: In view of the general dissatisfaction with, and disappointment in the progress to date, is it possible to justify any further research along similar lines? Turning to what seems to be the latest consensus of opinion by experts, Tom Arthur Lamke, chairman of the committee on teacher personnel, in his introduction to a summary of studies in the field from June 1952 to June 1955, compiled by Barr and others, has these encouraging words:

If the research during the last three years were to be wiped out in the fields of medicine, agriculture, physics, or chemistry, our lives would be materially changed. If research in the area of teacher personnel during the last three years should vanish, education and educators would continue much as usual. There are relatively few studies among the some 500 reported here which will, or should, widely affect educational practice.

There are perhaps two reasons for this state of affairs. In recent years it has been the style to emphasize the need for cooperative and coordinated effort in research of all kinds, and surely many of the defects in current educational research can be ascribed to the fragmentary, partial, and

---

<sup>1</sup> MacLean, Gowan and Gowan, op.cit., p. 669

sporadic approaches necessitated by the limited resources of one individual, or even of one institution. There can be little doubt that educational research would benefit by cooperative effort on a scale not heretofore realized.

This truism should not be permitted to obscure the fact that one-man research may also have great value. Perhaps a difficulty even more basic than lack of cooperative effort is the fact that educators are desperately in need of one-man research of a certain kind; we still await a Copernicus to simplify our explanations; a Newton to postulate a few major principles upon which a whole conceptualization may rest - at least for a time; a Mendeleev to order masses of apparently unrelated data; a Descartes, a Leibnitz, a Fisher, to provide us with mathematical models of reality as we see it, models constructed especially for our work, and not necessarily for other disciplines. Such one-man research still remains to be done.<sup>1</sup>

### (3. General Extent of Studies in Prediction

Reporting in the Encyclopedia of Educational Research on studies in the field prior to February, 1948, the authors, Charles W. Sanford and J. Lloyd Trump, point out that approximately 675 articles and monographs have been reported since 1905, on problems incident to the pre-service selection of teachers. The first rather complete summary was reported by Butsch<sup>2</sup> in 1931. This summary was supplemented by Torgerson<sup>3</sup> in 1934 and 1937, by Barr in

---

<sup>1</sup>A. S. Barr, "Measurement and Prediction of Teaching Efficiency," op.cit., p. 192.

<sup>2</sup>R. L. C. Butsch, "Teacher Rating," Review of Educational Research, (1931), 1, 99-107, 156-157.

<sup>3</sup>Torgerson, op.cit., pp. 261-266.

1940,<sup>1</sup> 1943,<sup>2</sup> and 1946.<sup>3</sup> Other excellent summaries would be those of Brammell<sup>4</sup> in 1932, Betts,<sup>5</sup> Yaukey and Anderson,<sup>6</sup> in 1933, Sandiford and Others<sup>7</sup> in 1947, and Eliasson and Martin<sup>8</sup> in 1945. Baker<sup>9</sup> and Troyer and Pace<sup>10</sup> cite some of

---

<sup>1</sup>A. S. Barr, "Measurement of Teaching Ability," and "Prediction of Teaching Success," Review of Educational Research, (1940), 10, 182-190.

<sup>2</sup>A. S. Barr, "Measurement and Prediction of Teaching Efficiency," and "Recruitment, Institutional Selection and Guidance of Teachers," Review of Educational Research, (1943), 13, 218-227.

<sup>3</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency," Review of Educational Research, (1946), 16, 203-208.

<sup>4</sup>P. R. Brammell, "Articulation of High School and College," U. S. Office of Education, Bulletin No. 17, 1932. National Survey of Secondary Education, Monograph No. 10, pp. 96.

<sup>5</sup>G. L. Betts, "The Education of Teachers Evaluated through Measurement of Teaching Ability," Special Survey Studies, U. S. Office of Education, Bulletin No. 10, 1933, National Survey of Education of Teachers, Vol. 5, Part 2, pp. 87-153.

<sup>6</sup>J. V. Yaukey and P. L. Anderson, "A Review of the Literature on the Factors Conditioning Teacher Success," Educational Administration and Supervision, (1933) 19, 511-520.

<sup>7</sup>Sandiford and Others, op.cit., summarized in Elementary School Journal, (January 1938), 38, 326-329.

<sup>8</sup>R. H. Eliasson and R. L. Martin, "Pretraining Selection of Teachers during 1940-43," Journal of Educational Research, (1945), 38, 666-677.

<sup>9</sup>F. E. Baker, "Selective Admission and Selective Promotion in Teacher Education Institutions," Yearbook XXIII, National Society of College Teachers of Education, University of Chicago, 1935, pp. 16-71.

<sup>10</sup>M. E. Troyer and C. R. Pace, "Evaluation in Teacher Education," American Council on Teacher Education, 1944, p. 368.

the usual practices and Barr, Burton and Brueckner<sup>1</sup> analyze various aspects and cite some of the pertinent research. In 1948, in a most comprehensive summary available, Barr presents the highlights of research in the field, classifying some 150 representative studies according to criteria, predictive variables, data-gathering devices and summarizing their findings.<sup>2</sup> Then again, in 1952,<sup>3</sup> Barr revised his reviews, and later, in 1955,<sup>4</sup> he presents a final review of research in the field for the past three years. It is beyond the scope of this review to imitate any of the above-mentioned summaries or to summarize them any further. The most practical thing to do would be to refer specifically to a few representative studies and to repeat some of the comments made by the experts in their summarizing. Mention has already been made of the variety of studies possible from the many criteria, factors related to teaching success, data-gathering

---

<sup>1</sup>Barr, Burton, and Brueckner, *Supervision*, op.cit.

<sup>2</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency; A Summary of Investigations," op.cit., 16, 207-211.

<sup>3</sup>Glen Fulkerson, "A Resumé of Current Teacher Personnel Research," Journal of Educational Research, May, 1954, p. 669.

<sup>4</sup>A. S. Barr, "Measurement and Prediction of Teaching Efficiency," Review of Educational Research, (June 1955), 25, 261-269.

devices, and statistical procedures attempted by research workers. Emphasis will be placed on those studies that have a more direct bearing on the study for which this review is prepared.

With reference to the predictive value of certain variables, research reports up to the end of 1947 indicate that only four of the many factors studied are important in the field of prediction. They are: intelligence, scholarship, personality, and scores earned on professional-information and subject-matter tests. The last, as pointed out by Charles W. Sanford and J. Lloyd Trump, are probably a combination of intelligence and scholarship.<sup>1</sup> Their observations continue in this manner:

The correlation between the four factors and teaching is positive but low. For example, Yaukey and Anderson<sup>2</sup> in a summary of eleven studies, reported a median correlation of .26 between scores on professional tests and teaching success in the field. There is also a positive, but low correlation between practice-teaching marks and estimated success in the field. Sandiford's summary<sup>3</sup> of sixteen studies reports a median correlation of .23 with a range of from .06 to .70 between marks earned in practice teaching and measures of teaching success.

Most of the studies have been concerned with the relationship between only one factor and one criterion. A

---

<sup>1</sup>C. W. Sanford and J. L. Trump, "Teacher Education - Preservice Selection," Encyclopedia of Educational Research, Revised edition, New York: MacMillan Co., 1952, p. 1392.

<sup>2</sup>Yaukey and Anderson, op.cit., p. 513.

<sup>3</sup>Sandiford and Others, op.cit.

more comprehensive study is that of Odenweller.<sup>1</sup> Five hundred and sixty Cleveland teachers, all graduates of the Cleveland School of Education, were ranked by experts as to effectiveness in teaching. Correlations and inter-correlations were computed for a variety of factors, including age, height, handwriting, intelligence, etc. Many positive but low correlations were computed with the highest of .825 between measures of personality and teaching effectiveness. The reader is also referred to a study by Kriner<sup>2</sup> in which a number of positive correlations were obtained for many factors, using ratings by Sanford and Trump: "Although the four factors mentioned above are positively correlated with teaching success, the researches clearly show that the prognostic value of the correlations is not high enough to justify pre-service selection based on these factors."<sup>3</sup>

In 1952, Bach<sup>4</sup> found the relationship between practice-teaching ability and success in the field as neg-

---

<sup>1</sup>A. L. Odenweller, "Predicting the Quality of Teaching," Contributions to Education, No. 676, Teachers' College, 1936, pp. 158-159.

<sup>2</sup>H. L. Kriner, "Five-Year Study of Teacher College Admissions," Educational Administration and Supervision, (1937) 23, 192-199.

<sup>3</sup>Sanford and Trump, op.cit.

<sup>4</sup>Olav Jacob, "Practice Teaching Success in Relation to Other Measures of Teaching Ability," Journal of Experimental Education, (September 1952), pp. 57-80.



ligible, while Ullman<sup>1</sup> in 1930 reported that success in practice teaching is the best single measure of teaching success. He also found that other factors had predictive value. These were: socio-economic status, academic and professional marks, social intelligence, general intelligence, interest in teaching, and others. Lins,<sup>2</sup> in 1946, found the percentile rank in high school as the best single measure for prediction purposes. This view is upheld by a study conducted by Jones.<sup>3</sup> In 1940, Rostker<sup>4</sup> found that the intelligence of the teacher was more closely related to teaching ability than any other factor. Leavitt<sup>5</sup> found

---

<sup>1</sup>Roy R. Ullman, "The Prediction of Teaching Success," Educational Administration and Supervision, (1930), 16, 598-608.

<sup>2</sup>Lins, op.cit., p. 59.

<sup>3</sup>Ronald D. Jones, "The Prediction of Teaching Efficiency from Objective Measures," Journal of Experimental Education, (September 1946), 15, 85-99.

<sup>4</sup>Leon E. Rostker, "The Measurement and Prediction of Teaching Ability," School and Society, (January 1940), 51, p. 31.

<sup>5</sup>Jerome E. Leavitt, "Personal Data and Prediction of Success of Student Teachers," Journal of Teacher Education, (September 1953), 4, 194-197.

that students scoring above the sixtieth percentile on the ACE Psychological Examination were more successful in student teaching than those who ranked below. Gough and Pemberton<sup>1</sup> found no significant correlation between student-teaching success and scores on the MMPI. Michaelis,<sup>2</sup> also, found that none of the scales included in the MMPI had a significant relationship to the rated success of university graduates enrolled in student teaching. Reed<sup>3</sup> found a relationship beyond chance expectancy between the teacher's effectiveness in the classroom and that aspect of the teacher's personality which permits him to be an accepting person. Knoell<sup>4</sup> found positive but small correlations between word fluency and three different ratings of teaching success. Carlile<sup>5</sup> found

---

<sup>1</sup>H. G. Gough and W. H. Pemberton, "Personality Characteristics Related to Success in Practice Teaching," Journal of Applied Psychology, (October 1952), 36, 307-309.

<sup>2</sup>John V. Michaelis, "The Prediction of Success in Student Teaching from Personality and Attitude Inventories," University of California Publications in Education, (1954), II, 415-481.

<sup>3</sup>Harold J. Reed, "An Investigation of the Relationship Between Teaching Effectiveness and the Teacher's Attitude of Acceptance," Journal of Experimental Education, (June 1953), pp. 277-325.

<sup>4</sup>Dorothy M. Knoell, "The Prediction of Teaching Success from Word Fluency Data," Journal of Educational Research, (May 1953), 46, 673-683.

<sup>5</sup>Amos B. Carlile, "Predicting Performance in the Teaching Profession," Journal of Educational Research, (May 1954), 47, 641-668.

that grade-point averages correlated best with student-teaching grades, but all correlations with some twenty-three measures used, though positive, were very low. Ringness<sup>1</sup> found certain attitudes predictive of teaching success, especially interest in a subject-matter field. In one of the most recent and comprehensive studies undertaken in Canada, Henry Bowers<sup>2</sup> in 1952, using literally a hundred different variables, found that the previous academic record of the student-teacher was worth more in prediction than any other single or composite variable.

And so the review may go on. Hundreds of different factors have been tested with a half-dozen different criteria. Results vary. Some studies have even attempted some rather strange factors. Montross,<sup>3</sup> for example, found positive correlations between certain objective measures of temperament such as speed of tapping, reaction time, fluency, right- and left-hand coordination, and teaching success. Perhaps most can be accomplished by falling back again on the expert judgment of Barr, whose observations given below

---

<sup>1</sup>T. A. Ringness, "Relationships Between Certain Attitudes Towards Teaching and Teaching Success," Journal of Experimental Education, (September 1952), 21, 1-55.

<sup>2</sup>Henry Bowers, Research in the Training of Teachers, J. M. Dent & Sons (Canada) Limited, The MacMillan Co. of Canada Ltd., (1952), pp. 1-167.

<sup>3</sup>Harold W. Montross, "Temperament and Teaching Success," Journal of Experimental Education, (September 1954), 23, 73-97.

seem to be justified by the preceding reviews:

The trend has been toward the inclusion of more measures and aspects of teacher ability in prediction studies. The latter studies in this field frequently employ 15 to 20 variable prediction equations. The accuracy of the prediction seems to have improved as more and better measures have been employed and as the criterion of efficiency has been refined. Multiple R's in the 70's and 80's are not uncommon in carefully designed studies. To date, the best predictions seem to be had from combinations of so-called subjective and objective measures. While subjective measures are in general more unreliable than objective measures they appear to provide in some instances data on aspects of efficiency not as yet adequately covered by objective means.<sup>1</sup>

Then, in 1955, Barr adds,

The search continues for a single generalized pattern of qualities or behaviors that characterize good teachers, notwithstanding the possibility that differential studies of teachers teaching different subjects to different sorts of people, under different conditions, and for different purposes might prove worthwhile.<sup>2</sup>

Such is the general picture of the predictive values of the various variables employed in studies until the current year. Perhaps much of the difficulty and inconsistency so noticeable in all summaries may be attributed to a disregard for the problem set forth by Barr as given above.

Turning now to a review of some of the major statis-

---

<sup>1</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," op.cit., pp. 224-225.

<sup>2</sup>A. S. Barr, "Measurement and Prediction of Teaching Efficiency," Review of Educational Research, (June 1955), 25, 266.

tical procedures used, it may be said that the computation of zero-order coefficients of correlation between factors which it is thought may condition success in teaching and the quality of teaching as determined by some criterion is the technique most frequently employed. Less frequently used techniques include the regression equation, partial correlation, multiple correlation, tetrachoric correlation, and more recently factor analysis and the discriminant function.

A zero-order coefficient of correlation is difficult to interpret. A low coefficient is not necessarily of little or no significance if interpreted in the light of the elements operating in the data used. Similarly, a high positive correlation is not necessarily indicative of high causation. Furthermore, the demonstration of a high reliability does not insure a correspondingly high validity. Many uncontrollable human factors, such as bias and the "halo effect" often enter into the subjective ratings used as criteria. Few investigators have used multiple and partial correlations with any success outside of Rostker,<sup>1</sup> Rolfe,<sup>2</sup> and La Duke.<sup>3</sup> Yet an analysis of teaching success indicates that

---

<sup>1</sup>Rostker, op.cit., pp. 30-32.

<sup>2</sup>J. F. Rolfe, "The Measurement of Teaching Ability," Journal of Experimental Education, (1945), 14, 52-74.

<sup>3</sup>C. V. La Duke, "The Measurement of Teaching Ability," Journal of Experimental Education, (1945), 14, 75-100.

more than one factor affect it and the further use of these techniques might prove significant. A regression-equation technique is currently being used in many institutions<sup>1</sup> while the tetrachoric coefficient has been used only spasmodically. Ullman,<sup>2</sup> in 1930, in an extensive study using correlation coefficients of zero-order and regression equations from partial and multiple coefficients, concluded that no surprising revelations could be made, but observed that when several factors found to be related to teaching success are combined in a regression equation, prediction of teaching success can be made. Lins,<sup>3</sup> in 1946, using the biserial coefficient technique, found it very useful in combining predictive variables into a composite measure for predicting success. In 1950, Schmid,<sup>4</sup> using the regression-equation method, drew our attention to the fact that the accuracy of prediction by that method decreases with the inclusion of an increasing number of variables in the equation, and is lost beyond a certain limit. He points out that the stability of regression coefficients is dependent

---

<sup>1</sup>Baker, op.cit.

<sup>2</sup>Ullman, op.cit.

<sup>3</sup>Lins, op.cit., pp. 2-60.

<sup>4</sup>John Schmid, Jr., "Factor Analysis of Prospective Teachers' Differences," Journal of Experimental Education, (June 1950), 18, 295.

upon the ratio of the number of variables to the number of cases in the original sample used to develop the coefficients. He also stresses the dangers involved in using measures that overlap in part or entirely. Lamke,<sup>1</sup> using the discriminant function, claimed that it distinguishes better than any other linear function between specified groups on whom common measurements are available.

One gathers, therefore, that despite many claims for the usefulness of the various techniques, a satisfactory technique for treating any set of data has not yet been produced. The studies by Yaukey and Anderson,<sup>2</sup> Betts,<sup>3</sup> and Sandiford and Others<sup>4</sup> illustrate the application of numerous techniques of investigation, and from that evidence it is relatively clear that no valid and reliable basis is yet available. Speaking of the statistical procedures in his summary of 1948, Barr makes these important observations:

There seems to be certain misuses of statistical techniques that occur with considerable frequency in our thinking and educational writing that could probably be corrected with a little attention to them: (a) coefficients of correlation cannot be taken at their face value; a co-

---

<sup>1</sup>Tom Arthur Lamke, "Personality and Teaching Success," Journal of Experimental Education, December 1951, p. 253.

<sup>2</sup>Yaukey and Anderson, op.cit., pp. 511-520.

<sup>3</sup>Betts, op.cit., pp. 87-153.

<sup>4</sup>Sandiford and Others, op.cit.

efficient of correlation is the product of many factors and needs careful interpretation; (b) the coefficient of correlation represents going togetherness of certain types of data and not causal relations; (c) when our concern is with correlations that represent some positive or negative relationships they must be shown to differ from zero by amounts that are statistically significant; (d) workers should be on the lookout for selective factors and errors of measurement that may influence the size and direction of coefficients of correlation.

Some persons are unduly concerned over the low correlations found in the many studies in this area. Low correlations are to be expected when a limited aspect of teaching is studied even with perfect instruments of measurement. Teaching efficiency is the product of many things. No one may contribute much but taken together they constitute teaching efficiency.<sup>1</sup>

#### 4. Some Selection Programs

The purpose of this section is to indicate the extent of progress made in the field of selection, which is the ultimate aim of all measurement and prediction studies. Before reference is made to any specific and comprehensive selection programs, it seems worthwhile to observe that although research does not furnish a scientific basis for preservice selection, a large number of teacher-training institutions use some form of selective admissions or of selective retentions or a combination of both. Among the factors ordinarily considered by institutions prior to

---

<sup>1</sup>A. S. Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations, (June 1948), 16, 225-226, op.cit.

granting entrance to schools and colleges of education are the following: completion of high school courses, scholarship in high school, scores earned on examinations including entrance examinations administered by the institution, intelligence tests, personality ratings following interviews and tests, health, speech, pattern of college courses completed, and others. Selective retention is ordinarily a matter of maintaining a required scholarship status, passing certain subject-matter examinations, and showing successful progress in practice teaching. The exhibition of certain professional attitudes is required in a few institutions. Most of the institutions that use a plan of selective admissions and of selective retentions include a composite of factors and assign weightings to the several items. The student must obtain and maintain a total score which at least equals the minimum composite score established, and, in addition, he must not fail to exceed a certain minimum on each factor. Below are given a few illustrations of methods of selection. These illustrations are to be considered as representative rather than inclusive.

1. In the selection program used at Syracuse University,<sup>1</sup> data on nine factors are obtained on each student, namely, intelligence, general culture, scholarship, special aptitudes, physical health and vitality, mental balance,

---

<sup>1</sup>Miss Verna White, op.cit., pp. 24-31.

personality, character, and attitudes. The selection of these factors is based on the eight criteria given earlier herein (see page 21). The process of selection has become increasingly one of guidance, including contacts with students, examining the demands of the profession and of their own potentialities and so on. The interviewer rates the student's voice, physical appearance, grooming, speaking ability, initiative, social intelligence and emotional balance. Staff members in group interview then rate the candidate. The testing battery consists of the ACE Psychological Examination, the Coop. General Culture Test, the Coop. Reading Test, Strong's Inventory for men, Kuder's Preference Record for women, and the MMPI. A summary of all the information is then made and discrepancies in the total picture are analyzed and diagnosed for possible remedial work. Admission, in some cases because of weaknesses, is deferred until the necessary remedial steps are taken. Sometimes admission is granted with the provision that the candidate will seek help concurrently with initial work in education. As an indication of the extent to which selection is exercised, it is reported that 577 out of 995 applicants were admitted in the period from 1947 to 1949. As the student progresses through the program and participates in the various experiences provided for the application of theory, there is constant self-revaluation by the student under the

guidance of the faculty. In several instances, students have withdrawn as they have gained insight into their lack of interest in or aptitude for the profession. In rare instances in which such insight is not gained by the student, he has been asked to withdraw from professional education.

2. Selection at Wayne University is based on the following areas of information:<sup>1</sup> general information supplied by the student (past educational history, personal interests, hobbies, extra-curricular activities, professional intentions, experience with children, general cultural and educational status of family, work experience, etc.); physical qualities (including speech, physical examination, etc.); intellectual capacity (college record, ACE); personal and social qualities (subjective estimates by staff workers who know the student, employers' appraisals, etc.); interviews (by staff adviser); group leadership experience (Boy Scouts, agency reports); effort is made to appraise the applicant as a total personality. No system of weighting the various criteria is used and no mathematical score is computed. In the years 1948 to 1952 about 60% of the applicants were admitted on these bases.

Other such examples may be cited. The reader is referred to the teacher selection and counseling service at

---

<sup>1</sup>Magee, op.cit., pp. 169-171.

the University of California.<sup>1</sup> State-wide programs of pre-service selection are used in a number of the states. In New York State, maximum undergraduate enrolments are specified for each state teachers' college. A number of municipal teachers' colleges and universities exercise rigid programs of preservice selection. As a matter of fact, whether or not on a scientific basis, most of the teacher-training colleges in Canada and the U.S.A. have some form of a selection program. It is unfortunate that the provincial normal schools in Canada do not adopt a more rigid program of selection in their institutions.

Although educators do not yet have evaluations of selective programs as they relate to the effectiveness of teaching in the field, there is abundant evidence regarding the relationship between many of the factors mentioned and the scholastic success of students in teacher-training institutions. Research studies such as the one conducted at Wayne University<sup>2</sup> indicate that teacher-training institutions may achieve reasonable success in graduating good teachers if they operate a combined program of selective admission, selective retention, and sound training. In reviewing studies in this field, for the last three years, Bretsch and Jacobson<sup>3</sup> arrive at these conclusions:

---

<sup>1</sup>MaeLean, Gowan and Gowan, op.cit., pp. 669-677.

<sup>2</sup>Magee, op.cit., p. 172.

<sup>3</sup>H. S. Bretsch and G. S. Jacobsen, "Recruitment, Guidance and Screening of Prospective Teachers," Review of Educational Research, June 1955, p. 204.

- (a) that the amount of basic research is meagre;
- (b) that research on teacher recruitment is based largely upon the opinions of students and the experiences of those responsible for recruiting teachers; and
- (c) that identification of factors useful in guidance and screening has been difficult because of lack of agreement as to what constitutes teaching success.

The reader is referred to the studies summarized and the programs evaluated in this most recent review.

#### 5. Some Major Current Research Programs in This Field

As a closing note, the author refers the reader to what may be considered major efforts to contribute to a better understanding of the overall problem of teaching effectiveness. One of these is the Cooperative Study to Predict Effectiveness in Secondary Teaching,<sup>1</sup> directed by Paul A. Hedlund for the New York State Education Department. Still in progress, this study attempts to answer the principal question: what student characteristics, measurable at the time of admission to a program of professional education, are related to success in that program and to effectiveness in teaching? The answer to this question is sought by:

- (a) gathering a comprehensive body of information on all applicants for teacher education in

---

<sup>1</sup> Glen Fulkerson, op.cit., pp. 669-670.

the cooperating colleges;

- (b) following the progress of the students through college and into employment to note the success in the teacher-education program and effectiveness during the first year of teaching; and
- (c) analyzing the prediction data in relation to academic and teaching success to determine the factors which colleges should consider in selecting their education students.

Such a comprehensive study, when completed, should help resolve some of the major problems that have plagued so many research workers in the last half century.

One of the most extensive of recent research programs devoted to this field of study is the Teacher Characteristics Study,<sup>1</sup> a six-year project under the auspices of the American Council on Education. The research has been guided by two principal objectives:

- (a) the identification and analysis of patterns of teacher behaviour; and
- (b) the development of materials useful for the prediction of teacher behaviour.

A number of related studies have been undertaken in the light of these goals. Reports have appeared from time to time describing both the general progress of the study and the results of some of the completed phases. Educators will be anxiously awaiting a complete and comprehensive report on this study.

---

<sup>1</sup> ibid., pp. 670-671.

### C. Conclusions and Generalizations

From the summaries, observations and conclusions presented in the preceding pages of this review, the following generalizations and conclusions appear to be warranted:

1. Research in this field is very extensive, is receiving more attention at the present time than ever before, and is becoming more and more sophisticated in character.

2. In terms of the progress research has made today, of the criteria attempted, supervisors' rating of teaching effectiveness has been used most productively; measurement of student achievement and student evaluation of teaching as criteria, though becoming more popular in recent studies, are still very precarious; the coefficients of reliability between the different criteria show generally no significant relationship.

3. In using the evaluation of judges in the field as a criterion, the tendency is toward composite rather than individual ratings.

4. For prediction purposes, statistical procedures favor categories of teaching effectiveness to ratings or rankings of teachers on their teaching success.

5. The concepts "teacher" and "teaching" have taken on new and additional garments of meaning within the last three decades.

6. There is very little agreement as to what factors

are related to teaching success; yet common sense and many studies indicate that some such relationship must exist.

7. The current trend is to place more emphasis on a general concept of qualities relating to teaching success rather than a determination of specific factors.

8. The relationship between practice-teaching ability and service in the field is negligible.

9. It is conceivable that the teacher's attitude towards teaching conditions his success in the profession.

10. The situation in which a teacher finds herself may have an important bearing on her effectiveness as a teacher.

11. A combination of objective and subjective measuring and evaluating devices appear to give better results than either employed separately; a serious attempt should be made at reducing the number of aspects about which information is required.

12. Teacher selection and recruitment is desirable and possible although research has not yet furnished any scientific bases for such programs.

13. There is no one clear pattern for a successful teacher; it will be many years before accurate scientific measures of teaching potentiality and growth are devised; moderate as the success of past research has been, the indications are that prediction of teaching efficiency is a worthwhile and feasible goal to be sought.

14. Predictions from subjective and objective data seem to be of about equal value and complementary to one another.

15. The current tendency in prediction studies is to use a combination of variables rather than individual ones and to adopt statistical methods to that end; the number of variables used should be governed by the size of the sample employed.

16. The interview is perhaps the best single source of subjective data about prospective teachers.

17. In terms of the foregoing observations, the problem conceived for this thesis and the study as a whole undertaken by the faculty are judged to be worthwhile and feasible.

As a final word it might be observed that the designation in this review of a few researches as outstanding does not, however, detract from the significance and value of those deemed somewhat less important. The eventual progress of research in this field will likely be judged in terms of the cumulative contributions of a great number of minor studies.

As research in the area continues to develop and improve, the number of single investigations likely to qualify as "milestones" seems certain to diminish. The milestones of the future will probably be pooled efforts of many investigators, each attempting to solve a small

segment of a problem that is too complex and extensive for a single study. Such groups of studies would accomplish considerably more if they were directed by some central body and their findings synthesized and interpreted by a few outstanding research experts such as A. S. Barr.

## CHAPTER III

### DESCRIPTION OF THE SAMPLE

The sample for the pilot study consisted of thirty-five teachers who were enrolled in Education I at the University of Manitoba during the 1953-54 term. Originally the enrolled group consisted of fifty-four student-teachers, most of whom went directly into teaching in the fall of 1954. Only thirty-two of the fifty-four remained, in their second year, in the teaching profession in the Province of Manitoba in the fall of 1955 and were thus available for this study. Three others, who had left the profession during or at the end of their first year of teaching, were added to the thirty-two to make up a sample of thirty-five subjects. It had been definitely established that these three left the profession because of their failure as teachers. No other subjects were similarly included because their leaving the profession could not be definitely attributed to failure. It was originally designed to have only those thirty-two who were available for rating of efficiency in the field in their second year of teaching. Reasons for this will be advanced later. However, for obvious reasons, the rating of the other three was such a certainty that it was considered

advisable to include them and thus add strength to the sample for statistical purposes which will be explained later. In all tables based on Chapter III an asterisk (\*) opposite a certain person will indicate that she is one of the above-mentioned three and will explain the absence of certain information available for the other thirty-two. An asterisk used elsewhere in this chapter will have the same denotation. As the description of the sample proceeds, the reader will observe that the inclusion of these three did not in any way jeopardize the quality of the sample in terms of the population from which it was taken. A more detailed description follows. In sections A to I inclusive, references will be made to Table 5, page 62, which supplies the pertinent raw data from which specific observations will be deduced.

A. Sex

Fifteen male and twenty female subjects constitute the sample. It might be noted that in the original group of fifty-four, there were nineteen males and thirty-five females, leaving the proportion approximately the same.

B. Age at Enrolment, September, 1953

As Table 5 indicates, the age range for the whole sample was from 18 years, 11 months to 37 years, 11 months, with a mean age of 23 years,  $4\frac{1}{2}$  months. Breaking this down in terms of sex, the range for the males was from 21 years, 6 months to 37 years, 11 months with the mean at 25 years;



the mean age for the females was 21 years, 6½ months within a range from 18 years, 11 months to 23 years.

C. Years of Teaching Experience Prior to Enrolment in Education I

More than 70 percent of the subjects had no teaching experience prior to their enrolment in the Faculty of Education. Five of the remaining subjects had one year's experience, three had two years' experience, and one had three years' teaching experience.

D. Academic Qualifications Prior to Enrolment in Education I

Twenty of the subjects had graduated with a B. A. degree, four with a B. Sc. degree, and two with an M. A. degree. The remaining nine had no degrees but all of them had completed most of the work required for one. Their deficiencies did not disqualify them from enrolling in Education I.

E. Taught During 1954-55?

It might be of interest to note that only two of the thirty-five subjects did not put in a full year of teaching during the 1954-55 season.

F. Type of School Taught in at Time of Rating (1955-56)

It should be observed here that fortunately with only minor exceptions the "type of school taught in" actually

represents the "grade-level of work taught." Table 6, on page 65, presents a summary of the distribution of this classification.

G. Major Subjects Taught -- 1955-56

It is hoped that the somewhat broad classification of this information will be of some assistance in obtaining a clearer picture of the sample in terms of preference for school subjects and in estimating how representative this sample is of larger populations. This information is summarized in Table 7 on page 65.

(The two items of information that follow - H and I - are given primarily for the purpose of showing that the sample of thirty-five, even in these two respects, does not differ significantly from the whole group of fifty-four. Perhaps it is not too early to observe that these two items of data will not be used as predictive variables in the pilot study because they are secured during the course of training. The predictive variables used in the study are those that can be secured immediately prior to or at enrolment so that they may be used for selection purposes as well as for guidance.)

H. Practice-teaching Rating -- 1953-54

From Table 5, the median rating for the sample was calculated in terms of the grade-ratings used and it was found to lie between B and B-. Of interest is the fact that the median rating for the whole group of fifty-four lies in that same range.

TABLE 6

TYPES OF SCHOOLS TAUGHT  
IN BY THE SUBJECTS AT  
THE TIME OF RATING

Type of School	No.
Elementary . . . . .	5
Elementary and Junior High . . . . .	5
Junior High . . . . .	6
Junior and Senior High . . . . .	9
Senior High . . . . .	7
*Not teaching -- 1955-56 . . . . .	3
Total . . . . .	35

TABLE 7

MAJOR SUBJECTS TAUGHT BY THE  
TEACHERS DURING 1955-56

Major Subjects	No.
Mostly Mathematics and/or Science . . . . .	4
Mostly English and/or Social Studies . . . . .	12
Mostly Foreign Languages . . . . .	3
Mostly Home Economics . . . . .	3
Mostly Physical Training and/or Health .. . . .	2
All subjects of Grade (8) . . . . .	8
*Not teaching -- 1955-56 . . . . .	3
Total . . . . .	35

### I. Average Percentage Standing in Education I Classes

From the average standings in Table 5 the mean and the median were calculated to be 69.8 and 67.75 respectively; for the whole group of fifty-four from which the sample was drawn, the mean and median are 70.2 and 68.5 respectively.

The purpose of the next few paragraphs is to present information that may be useful in comparing the sample with the whole group of fifty-four in terms of data from which the predictive variables will be drawn and also in estimating whether or not the data for the sample shows any significant departure from accepted norms for similar groups of individuals. Table 8, on page 67, supplies the scores for the sample on the ACE-psychological, the Coop Math, the Watson-Glaser, and the English Usage tests. In this table, column J contains total scores (Q and L) on the ACE; column K contains total raw scores on the Coop Math; column L contains total raw scores on the Watson-Glaser; column M contains the sums of the total raw scores on the Michigan Vocabulary, the SRA Reading, and the Essentials of English tests. This last sum of total raw scores is designated as "English Usage." Reasons for such grouping will be advanced in the next chapter on procedures. These four sets of scores, as they are now presented, will be used as four of the five variables in the pilot study. Table 8 also presents the means and standard deviations for each set of scores for both the

TABLE 8

DATA ON THE SAMPLE FROM FOUR AREAS OF  
ABILITY - GENERAL SCHOLASTIC, MATHEMATICS,  
CRITICAL THINKING AND ENGLISH USAGE

	J		K		L		M					
	ACE	Psy	Co- op	Math	Watson- Glaser	English Usage						
1		102		25		59		235				
2		128		28		59		266				
3		154		37		71		264				
4		149		57		82		301				
5		93		20		68		233				
6		154		51		84		282				
7		135		34		63		235				
8		133		33		74		278				
9		137		46		64		243				
10		156		49		67		260				
11		126		28		65		241				
12		126		29		76		280				
13		125		53		72		286				
14		137		37		76		242				
15		118		21		54		218				
16		134		59		71		260				
17		127		30		78		255				
18		128		21		77		245				
19		110		42		64		255				
20		123		27		64		229				
21		117		19		75		227				
22		130		32		70		248				
23		103		29		70		231				
24		131		53		81		271				
25		130		34		64		206				
26		146		56		81		256				
27		174		62		85		313				
28		117		43		73		189				
29		125		42		67		262				
30		109		42		75		184				
31		149		32		74		260				
32		168		45		87		269				
33		96		23		56		172				
34		135		29		67		254				
35		83		22		51		179				
M <sub>s</sub>	128.7		36.86		70.3		246.79					
M <sub>pop</sub>	130.4		36.35		70.3		248.1					
σ <sub>s</sub>	19.8		12.45		8.67		32.66					
σ <sub>pop</sub>	20.5		11.32		8.58		31.77					

sample ( $M_s$  and  $\sigma_s$ ) and the whole group of fifty-four ( $M_{pop}$  and  $\sigma_{pop}$ ). It can be readily seen that there is no significant discrimination in the scores for all four sets between the sample and the population from which the sample was taken.

Table 9, on page 69, shows a comparison between the means and standard deviations, where available, for the sample, for the whole group of fifty-four, and what might be termed as accepted norms for similar populations. In the case of the ACE, the accepted norms were taken from a study made in this province and reported in research bulletin 17.<sup>1</sup> The norms for the remaining two in the table were taken out of the Manual of Instructions supplied with the tests.

Table 10, on page 70, presents the T-scores for the sample for each of the ten scales of the MMPI, the sums of T-scores for all ten scales (10T), and the sums of the T-scores for scales 1, 2, 3, 4, 6, 7, 8, 9, 10 (10T-5). For reasons which will be presented in the chapter to follow, the last column, 10T-5, of scores will be used as the fifth predictive variable in the pilot study. At the bottom of the table will be found comparisons of means and standard deviations between the sample ( $M_s$  and  $\sigma_s$ ) and the whole group of fifty-four ( $M_{pop}$  and  $\sigma_{pop}$ ).

The preceding discussion and the data from Tables 5 to 10 appear to warrant the following general observations about the sample used in the study:

---

<sup>1</sup>H. L. Stein, *Aptitudes and Achievements of Faculty Students*, Research Bulletin No. 17, Faculty of Education, University of Manitoba (December 1953), pp. 13-21.

TABLE 9

COMPARISON OF MEASURES OF CENTRAL TENDENCY  
AND VARIABILITY FOR SAMPLE USED IN PILOT  
STUDY WITH THOSE OF THE WHOLE GROUP  
OF FIFTY-FOUR AND WITH SOME  
ACCEPTED NORMS

Test	Sampling	Mean	S.D.
ACE - Psych	Sample of thirty-five student-teachers used in this pilot study -- 1953	128.7	19.8
	Whole group of fifty-four from which sample was taken	130.4	20.5
	College students - Univ. of Man. -- 1953	131.6	21.3
	Normal School students - Man. -- 1952-53	103.4	23.4
Coop - Math	Sample of thirty-five student-teachers used in this pilot study -- 1953	36.86	12.45
	Whole group of fifty-four from which sample was taken	36.35	11.32
	Grade XII -- as per Manual of Instructions	25.0	--
	College -- as per Manual of Instructions	33.6	--
Watson-Glaser	Sample of thirty-five student-teachers used in this pilot study -- 1953	70.3	8.67
	Whole group of fifty-four from which sample was taken	70.3	8.58
	High School -- as per Manual of Instructions	54.3	10.9
	College -- as per Manual of Instructions	69.3	9.5

TABLE 10  
T-SCORES AND COMBINATIONS OF T-SCORES  
FOR THE SAMPLE  
FROM SCALES ON THE MMPI

	M M P I - T - Scores										Total 1 - 10	10T	- 5
	1 H <sub>s</sub>	2 D	3 H <sub>y</sub>	4 P <sub>d</sub>	5 M <sub>f</sub>	6 P <sub>a</sub>	7 P <sub>t</sub>	8 S <sub>c</sub>	9 M <sub>a</sub>	10 I-E			
1	50	44	61	51	36	47	44	46	52	42	473		437
2	45	39	51	47	69	47	42	51	55	40	486		417
3	52	53	61	68	53	53	56	52	45	49	542		489
4	53	46	52	57	64	61	55	49	58	59	554		490
5	58	53	64	54	54	50	52	57	45	52	539		485
6	53	43	58	59	44	56	51	55	60	41	520		476
7	48	45	50	55	63	53	61	57	62	63	557		494
8	52	47	63	57	53	58	59	59	68	47	563		510
9	47	56	53	53	57	46	49	51	55	56	523		466
10	46	51	58	59	44	56	49	48	52	42	505		461
11	48	45	59	53	36	56	51	50	55	48	501		465
12	52	58	45	48	70	43	52	45	60	57	530		460
13	56	42	57	66	53	53	53	59	65	47	551		498
14	46	45	48	43	52	47	42	42	45	55	465		413
15	39	43	31	41	44	44	51	49	70	64	476		432
16	48	49	54	59	56	56	53	57	60	44	536		480
17	54	58	67	51	53	35	39	46	47	48	498		445
18	45	51	58	53	67	49	47	59	52	42	523		456
19	62	64	64	59	73	61	61	57	66	63	630		557
20	54	47	55	48	44	49	50	42	52	55	496		452
21	50	59	50	68	63	53	59	61	55	45	563		500
22	54	48	58	62	57	59	62	56	50	65	571		514
23	53	55	56	57	62	47	70	61	42	57	560		498
24	61	56	67	52	65	61	53	57	49	43	564		499
25	41	53	41	43	49	48	48	47	45	65	480		431
26	51	65	59	62	63	58	55	51	55	53	572		509
27	59	68	56	64	67	53	62	57	33	52	571		504
28	42	59	58	64	63	53	52	53	58	54	555		492
29	59	47	59	59	48	58	55	56	66	40	547		499
30	49	34	54	55	53	44	44	42	39	50	464		411
31	44	42	53	53	42	53	48	54	56	40	485		443
32	44	40	58	55	57	65	49	51	63	42	524		467
33	54	61	52	64	61	64	58	49	47	51	561		500
34	58	65	66	55	48	62	70	61	58	58	601		553
35	46	55	52	51	53	64	64	62	66	53	566		513
M <sub>s</sub>	50.8	50.9	55.7	55.19	55.19	53.4	53.3	53.24	54.5	50.9	533.9		477.9
M <sub>pop</sub>	51.13	51.0	55.83	56.57	53.8	53.46	54.48	53.69	56.41	50.76	537.6		483.4
σ <sub>s</sub>	5.73	8.19	7.27	7.09	9.2	6.73	7.35	4.99	8.87	7.57	40.28		35.02
σ <sub>pop</sub>	6.65	8.28	7.3	7.97	10.04	7.5	7.86	5.85	10.15	8.17	44.7		39.61

1. From the information on ages and academic qualifications one can readily see that the subjects are mature individuals who have stood the test and weeding-out process of at least three years of university work and who have had time to give serious thought to choice of vocations, to develop sound habits of work, wholesome attitudes and a philosophy of life. Definitely, in those terms, this is a "select" group and cannot be considered representative of all those people who begin their teacher-training studies, nor representative of all the teachers in their second year of teaching. Since about ten times as many teachers graduate from the provincial normal schools as from the Faculty of Education, the implications are so obvious that comparisons must be made guardedly.

2. Sections F and G show that there is a good sampling in terms of "schools" and "subjects taught." As might be expected, however, most of these schools are located in towns and cities where standards, as a general rule, are somewhat high. This is likely to be reflected in ratings of teaching performance. Consequently, a "good" teacher in such environments could easily be better than a "good" teacher in a different environment. This factor no doubt adds to the selectiveness of the sample. Caution must be exercised when speaking of the calibre of teaching.

3. As Table 9 bears out, this sample as a group is superior in some basic abilities when compared with accepted

norms, particularly in mathematics and in critical thinking. No doubt this superiority would even hold when the sample is compared with all teachers in their second year of teaching. This discussion is submitted so that the reader and subsequent research workers in this locality may visualize that it might be fairly difficult to obtain from this sample a so-called "poor" teacher, if this poor teacher is to be typical of the very small percentage of teachers considered poor in the whole teaching profession. Or, to put it differently, the "poor" teacher that will be designated as such in this study need not necessarily be "poor" in terms of a more universal rating.

4. Great care has been taken to present all available information about the sample and about the whole group of fifty-four student-teachers from which the sample was drawn. It is suggested that further research in this field, on a local basis, will take all this into consideration when comparing samples from different studies for the purpose of ultimate generalizations. Important too is the need for showing that the sample shows no significant departure from the population from which it was selected. No such departure is revealed in any of the comparative data submitted. This is fortunate because, otherwise, the soundness of the practice of "throwing in" subjects in order to strengthen the sample, as was done in this study, could indeed have been questionable and extremely precarious.

Thus it might be concluded that this sample, although possibly fairly representative of the graduates of the Faculty of Education, must be recognized as somewhat "selective" when considered in terms of all student-teachers on a provincial level, or all prospective teachers, or, for that matter, all teachers in their second year of teaching. Some difficulty is anticipated in finding a sufficiently large number of "poor" or unsatisfactory teachers in the sample for the purpose of statistical computations which involve discrimination in variables. Likewise, any attempt at larger dichotomies of "good" and "poor" teachers may unfortunately lead to certain complications which yield results of dubious value. This study as well as further studies in this area must proceed with proper attention to these precautions.

## CHAPTER IV

### PROCEDURES AND TECHNIQUES IN DATA-GATHERING

#### A. Securing Measures of Prediction Variables

It has already been indicated that the five sets of measures used as prediction variables in the pilot study were secured from scores on a battery of tests administered to the candidates during the first few days of their enrolment in Education I. The responsibility of administering and scoring the tests was shared by faculty members, who may be considered specialists in this field of work. Other data mentioned in the description of the sample but not used as prediction variables were secured from student application forms, interview records, and individual student files. A newly-designed filing system for the purpose of compiling all relevant data on individual student cards has been introduced. This will be explained later in this chapter and in the closing one with other recommendations for further research and study.

Table 11, on page 75, presents the five sets of measures ( $X_1, X_2, X_3, X_4, X_5$ ) to be used as prediction variables. It has already been pointed out that the "English Usage" measures ( $X_4$ ) were secured by totalling the raw scores

TABLE 11

FIVE SETS OF MEASURES ~~USED~~  
 USED AS PREDICTION VAR-  
 IABLES IN THE PILOT STUDY

	ACE-PSY	COOP-MATH	WAT-GLA	ENGLISH	MMPI					
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>					
1	102	25	59	235	437					
2	128	28	59	266	417					
3	154	37	71	264	489					
4	149	57	82	301	490					
5	93	20	68	233	485					
6	154	51	84	282	476					
7	135	34	63	235	494					
8	133	33	74	278	510					
9	137	46	64	243	466					
10	156	49	67	260	461					
11	126	28	65	241	465					
12	126	29	76	280	460					
13	125	53	72	286	498					
14	137	37	76	242	413					
15	118	21	54	218	432					
16	134	59	71	260	480					
17	127	30	78	255	445					
18	128	21	77	245	456					
19	110	42	64	255	557					
20	123	27	64	229	452					
21	117	19	75	227	500					
22	130	32	70	248	514					
23	103	29	70	231	498					
24	131	53	81	271	499					
25	130	34	64	206	431					
26	146	56	81	256	509					
27	174	62	85	313	504					
28	117	43	73	189	492					
29	125	42	67	262	499					
30	109	42	75	184	411					
31	149	32	74	260	443					
32	168	45	87	269	467					
33	96	23	56	172	500					
34	135	29	67	254	553					
35	83	22	51	179	513					
M	128.7	36.86	70.3	246.79	477.9					
σ	19.8	12.45	8.67	32.66	35.02					

from three tests: Essentials of English Test-Form A, SRA Reading Record, and the Michigan Vocabulary Profile Test. It is assumed here that the total raw score is a fair indication of the student's general ability in all aspects of English practice. An examination of the three tests will show that no important area is omitted. It also assumes that no area is unduly stressed in the total score, thus requiring no weighting of scores on the individual tests.

The MMPI set of measures ( $X_5$ ) was obtained by adding the T-scores of all ten scales except the fifth. This fifth scale ( $M_f$ ) is the masculinity-femininity<sup>in</sup> scale. In totalling the scores in this manner, three major assumptions are made:

- (1) Since all the scores have been transmuted to a single T-scale, it is assumed that an algebraic summation of the scores is statistically permissible.
- (2) Since only positive deviations from the mean in the T-scale are considered as indications of maladjustment, it is assumed that by adding the scores a composite score will present a fairly good estimate of the total picture of the extent to which a student may be generally well adjusted or generally maladjusted. It might be added that it seems somewhat futile to treat the individual scales as variables because a maladjustment in one area does not carry enough strength to have any pronounced effect on the student's capacity to teach nor could it produce statistically any significant separation of "good" teachers from "poor" ones. A general picture of personality - of maladjustment, if any - is the purpose of such scores whether they be used for guidance or for selection.

- (3) The  $M_f$  scale was omitted from the totals because it has been found that male teachers usually show a high  $M_f$ , and that this factor is not generally associated with maladjustment among teachers as far as their teaching is concerned.

Earlier in the thesis it was indicated that the selection of procedures will be based, insofar as applicable, on the findings of the review of previous studies in the field. One of the findings, as given at the end of Chapter II, was:

The current tendency in prediction studies is to use a combination of variables rather than individual ones and to adapt statistical methods to that end; the number of variables used should be governed by the size of the sample employed.

In terms of other studies, a sample of thirty-five subjects used here could scarcely justify the use of more than five variables in a discriminant analysis. The sources of the first three variables have been explained in Chapter III. In future references to them, abbreviations given in Table 10 will be used for the five variables. These abbreviations are now repeated:

- $X_1$  -- ACE Psychological -- total of Q and L scores
- $X_2$  -- Coop Math -- raw scores
- $X_3$  -- Watson-Glaser -- total raw scores
- $X_4$  -- English Usage -- composite score
- $X_5$  -- MMPI -- Composite of scales 1 to 10, less scale No. 5.

B. Securing a Rating (Categorization)  
of Teaching Performance

The procedures adopted for rating, with some modifications appropriate to the local situation, are also based on the findings of previous investigations in the field. The pertinent findings are here repeated:

1. Of the three criteria attempted in the various studies, supervisors' ratings of teaching effectiveness have been found generally most practicable and most productive.
2. The current tendency is toward using a composite rating rather than individual ratings.
3. For prediction purposes, the latest statistical procedures favor categories of teaching effectiveness as against ratings or rankings of teaching success.
4. More emphasis is being placed on a general concept of qualities relating to teaching success rather than a determination of specific factors.
5. A combination of objective and subjective measuring and evaluating devices appears to give better results than either employed separately; a serious attempt should be made to reduce the number of aspects about which information is required.
6. The situation in which a teacher finds herself may have an important bearing on her effectiveness as a teacher.

Many studies emphasize the importance of situational factors in teaching success. Ullman's observations in this regard seem to be typical of the attention that these factors have been receiving:

Rating of teaching success during the first semester of actual teaching is ... too soon. An important criterion of the thoroughness with which the training institution has prepared its graduates lies in their ability to do a successful job of teaching once they have had a reasonable chance to become adapted to the situation in which they are placed.<sup>1</sup>

That is why the subjects in this study were rated during their second year of teaching. Even if the teachers do not remain in the same situation during their second year, they will nevertheless then have a better opportunity and will be better equipped to select a second situation more suitable to their tastes and more conducive to the kind of teaching of which they are capable. With the exception of the "special three", an attempt was made to have all rated during the period from November 15, 1955 to January 15, 1956. No doubt rating the teachers during their third or fourth year of teaching would give an even better control of the situational factor, but, unfortunately, the research worker might then be left without a sample.

It should be pointed out that locating the available subjects in their respective schools posed quite a problem, even with the assistance of the incomplete files of the offices of the Manitoba Teachers' Society and the Administration Branch of the Department of Education. Name changes through marriage and other reasons did not facilitate

---

<sup>1</sup>Ullman, op.cit., pp. 598-608.

the work. It is unfortunate that no cumulative record of individual teachers is kept, from which such information could be more easily obtained.

In order to obtain a composite "expert" rating of the teaching effectiveness of each of the subjects, it was decided to solicit the services of principals, inspectors and supervisors, hoping that at least two separate ratings could be obtained on each subject and from which a composite rating could be deduced. Preparing a mailing list of principals, inspectors and supervisors was done mostly through the offices of the Chief Inspector and the various school boards concerned. In the event of failure to receive two ratings from these sources, or, in the event that the two ratings showed significant discrepancies and were not conducive to a valid composite rating, it was decided that a faculty member would visit the school and secure the second or third rating. Available funds did not permit any wider faculty participation in the work of rating.

In keeping with the findings listed at the beginning of this section, a very simple rating scale was designed for the purpose of categorizing the teaching performance of the subjects as: outstanding, satisfactory, unsatisfactory, or failure. These categories were described in very broad terms. A suggested list of qualities in terms of which teachers could be categorized was submitted as a possible guide. Also provision was made for some basic information

that was used in describing the sample. Copies of the rating scale and of the accompanying letter are supplied in the appendix. It will be noted that four categories were used rather than three or five, because of the tendency of some raters, who are hesitant in reaching a decision, to "dump" doubtful cases into some middle category. Such dumping is not conducive to valid results wherein dichotomies have to be struck. Here the dichotomy, of course, consists of the first two and the last two categories.

Although the scales and the letters were sent out on November 15, 1955 with a plea for the responses to be in the mail by December 15, some of the responses kept dribbling in throughout the month of January, and then only after a reminder by telephone. Some were not received at all, necessitating a rating by the faculty. This was the most discouraging aspect of the whole study. Not only were some ratings not done, or done too late, but some of those that were done had so many qualifications attached to them as to render them valueless. Where a faculty rating was required, a letter from the Dean's office was the instrument used in paving the way for a faculty-member visit to the school. A copy of this letter is supplied in the appendix. The responses even to this "official" appeal were slow and incomplete, unnecessarily delaying the progress of the study. Eventually, consensus of opinion among faculty members had to be resorted to for additional ratings on a few subjects.

Table 12, on page 83, presents the categories as they were received in the ratings from principals (P), inspectors (I), and others (O), either from supervisors or Faculty of Education members. The abbreviations used for categories are as follows: O - outstanding; S - satisfactory; U - unsatisfactory; F - failures, used mainly for categorizing the special three who had left the profession because of their failure as teachers. (S-U indicates either a qualified categorization in one of the two categories involved or the uncalled-for creation of a category midway between the two.) The fourth column (C) gives the composite rating and designates the two halves of the dichotomy, a "+" sign indicating the "favorable" group, and a "-" sign indicating the "unfavorable" group.

### C. Classifying and Recording of Data

Individually, the prediction data and the categorizing data have already been classified sufficiently and are now ready for their combination for statistical analysis. The classification of the former has yielded five prediction variables. The classification of the latter has produced a dichotomy of two groups consisting of 27 (+) or favorable measures, and 8 (-) or unfavorable measures. Table 13, on page 84, now brings these two sets of data together with the favorable (+) opposite numbers 1 to 27, and the un-



TABLE 13

PREDICTION VARIABLES AND BASIC STATISTICAL MEASURES  
IN DICHOTOMY

	DICHOTOMY		PREDICTION VARIABLES				
	+	-	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	+		102	25	59	235	437
2	+		149	57	82	301	490
3	+		154	51	84	282	476
4	+		135	34	63	235	494
5	+		133	33	74	278	510
6	+		137	46	64	243	466
7	+		156	49	67	260	461
8	+		126	28	65	241	465
9	+		126	29	76	280	460
10	+		125	53	72	286	498
11	+		137	37	76	242	413
12	+		118	21	54	218	432
13	+		134	59	71	260	480
14	+		127	30	78	255	445
15	+		128	21	77	245	456
16	+		123	27	64	229	452
17	+		130	32	70	248	514
18	+		103	29	70	231	498
19	+		131	53	81	271	499
20	+		146	56	81	256	509
21	+		174	62	85	313	504
22	+		117	43	73	189	492
23	+		125	42	67	262	499
24	+		149	32	74	260	443
25	+		168	45	87	269	467
26	+		96	23	56	172	500
27	+		135	29	67	254	553
28	-		128	28	59	266	417
29	-		154	37	71	264	489
30	-		93	20	68	233	485
31	-		110	42	64	255	557
32	-		117	19	75	227	500
33	-		130	34	64	206	431
34	-		109	42	75	184	411
35	-		83	22	51	179	513
M <sub>+</sub>			132.74	38.74	71.74	252.4	478.26
σ <sub>+</sub>			17.96	12.47	8.59	29.47	30.36
M <sub>-</sub>			115.5	30.5	65.875	226.75	475.375
σ <sub>-</sub>			20.875	8.94	7.7	32.22	38.07
M <sub>s</sub>			128.7	36.86	70.3	246.79	477.9
σ <sub>s</sub>			19.8	12.45	8.67	32.66	35.02

favorable (-) opposite numbers 28 to 35. The table also includes the means and standard deviations for the whole sample ( $M_s$  and  $\sigma_s$ ), for the favorable group ( $M_+$  and  $\sigma_+$ ), and for the unfavorable group ( $M_-$  and  $\sigma_-$ ), for each set of measures. These statistical measures will be frequently used in the statistical procedures to follow and in subsequent tables.

Table 14, on page 86, presents the intercorrelation coefficients for the five sets of variables. These coefficients will not be used directly in any future statistical computations. They are submitted because it is interesting to note just how much relationship there is amongst the variables whose predictive value is to be tested. Later, it will be interesting to compare these coefficients with the corresponding biserial coefficients which will indicate the "predictive" strength of the variables, and to compare them with their individual contributions to the "separation" strength of various combinations of these variables as revealed by the application of the discriminant function.

The only other instrument of classifying and recording data that perhaps should be mentioned here is the proposed individual filing card referred to earlier. This card will make it possible to record all data about the subjects. Not only is there provision for the type of data that has been used for describing the sample, for selecting variables, for arriving at composite categorization of

TABLE 14

INTERCORRELATION COEFFICIENTS AMONG THE  
FIVE PREDICTION VARIABLES FOR THE  
WHOLE SAMPLE OF 35 SUBJECTS  
(BY THE RAW-SCORE METHOD)

	ACE PSYCH. $X_1$	COOP. MATH. $X_2$	WATSON- GLASER $X_3$	ENGLISH USAGE $X_4$	MMPI 1OT-5 $X_5$
ACE PSYCH. $X_1$	/ / / / /				
COOP. MATH. $X_2$	$r =$ +.681	/ / / / /			
WATSON- GLASER $X_3$	$r =$ +.611	$r =$ +.566	/ / / / /		
ENGLISH USAGE $X_4$	$r =$ +.710	$r =$ +.544	$r =$ +.596	/ / / / /	
MMPI 1OT-5 $X_5$	$r =$ -.048	$r =$ +.181	$r =$ +.043	$r =$ +.168	/ / / / /

teaching effectiveness, but it also makes provision for other data that may possibly be used in further research and in student guidance. A filing system using such individual cards can facilitate a great deal the preliminary work of organizing any project in conjunction with the teacher-training study undertaken by the faculty. As indicated earlier, a copy of this filing card is included in the appendix.

#### D. Important Limitations and Basic Assumptions

Most of the important limitations of the data and the basic assumptions underlying the adopted procedures have already been discussed throughout the foregoing sections of the thesis. For the sake of emphasis a few major ones are repeated here along with additional assumptions that must be clarified at this stage.

- (1) Due to the "selective" character of the sample described in Chapter III, only a somewhat arbitrary but insufficiently "sharp" dichotomizing of subjects into "favorable" and "unfavorable" has been achieved. This lack of sharpness may make it difficult to secure the required amount of separation in the statistical analysis of the data.
- (2) Although great care has been taken to objectify, in so far as possible, the procedure of rating (categorizing) teaching performance, there is no assurance that the final categories are entirely free from the vagaries of subjective judgment.
- (3) As pointed out earlier in this chapter, it is assumed that the algebraic summation of scores employed in arriving at two of the five sets of variables is statistically sound.

- (4) Fearing that the size of the available sample may be reduced drastically had rating been postponed until the third or fourth year of teaching, the "situational" factor in teaching was controlled only to a limited extent with ratings done during the second year.
- (5) It is assumed that five sets of variables are not too many for a sample of thirty-five subjects in the application of the biserial correlation and discriminant function methods of analysis.
- (6) As Table 13 indicates, the dichotomy consists of 27 subjects in the "favorable" group and 8 subjects in the "unfavorable". On a percentage basis this yields a ratio of 77.14:22.86. Wert, in his text on statistical methods stresses the importance of this ratio in the application of biserial correlation which will be used in this study:

Until a formula is developed, either from theoretical or empirical considerations, with more general application, the computation of biserial correlation with less than 5% to 10% in either of the two categories is open to question.<sup>1</sup>

It is here assumed, then, that the categories in this study are sufficiently large for that purpose.

- (7) One of the most important assumptions underlying the use of both the biserial correlation and the discriminant analysis is that a dichotomy is actually a normally distributed variable and that any attempt to indicate the relationship between a variable or variables and a dichotomy does not violate that assumption. Furthermore, Wert makes this additional important observation:

Whenever the foregoing assumption of a normally distributed variable classified into a dichotomy appears tenable, the ad-

---

<sup>1</sup>Wert, Neidt and Ahmann, Statistical Methods in Educational and Psychological Research, New York: Appleton-Century Crofts, Inc., 1954, p. 258.

ditional assumption must be made that the relationship between the numerical variable and the dichotomous variable is linear.<sup>1</sup>

It does not seem unreasonable to assume here, on the basis of tendencies to be "successful" or "unsuccessful", that the dichotomy in this study is a normally distributed variable with the tendency to be "successful" so great that actually 77% are successful and 23% are not. Nor does it seem unreasonable to assume that the relationship of which Wert speaks is linear, in this study.

---

<sup>1</sup>ibid.

## CHAPTER V

### PROCEDURES AND TECHNIQUES IN INTERPRETING DATA

#### A. Biserial Coefficient Technique

The function of biserial correlation is similar to that of the product-moment correlation coefficient, the technique most frequently used in prediction studies. The only important difference is this: whereas the latter indicates the linear relationship between two numerical variables, biserial correlation indicates the relationship between a numerical variable and a variable appearing in a dichotomy. The important assumptions underlying its application have been discussed in the preceding chapter. It now remains to apply it to the data in Table 13. The formula for its computation is:

$$r_{\text{bis}} = \frac{d}{\sigma} \left( \frac{pq}{z} \right)$$

Table 15, on page 91, gives the meanings of the symbols in the formula and presents their numerical values. It should be noted that the percentages of the categories in the dichotomy have been reduced to decimal fractions; the value of "z" has been calculated, by interpolation, from a table supplied in a textbook to which Table 15 refers. The value of  $\frac{z}{\sqrt{pq}}$ , which is required for testing the significance of

TABLE 15

ANALYSIS DATA FOR APPLICATION OF  
BISERIAL CORRELATION FORMULA

(N = 35)

Variable	See Table 13		p %+	q %-	z by interpo- lation <sub>1</sub>	$\frac{z}{\sqrt{pq}}$ by interpo- lation <sub>2</sub>
	d M <sub>+</sub> - M <sub>-</sub>	$\sigma_s$				
X <sub>1</sub>	17.24	19.8	.7714	.2286	.30261	.7206
X <sub>2</sub>	8.24	12.45	.7714	.2286	.30261	.7206
X <sub>3</sub>	5.86	8.67	.7714	.2286	.30261	.7206
X <sub>4</sub>	25.65	32.66	.7714	.2286	.30261	.7206
X <sub>5</sub>	2.88	35.02	.7714	.2286	.30261	.7206

X<sub>1</sub> to X<sub>5</sub> -- prediction variables

d = difference between categories in means of prediction variables

$\sigma_s$  = standard deviation of the prediction variables for the whole sample

p = proportion of cases in (+) dichotomy category

q = " " " " (-) " "

z = height of ordinate dividing the normal curve of unit area into p and q parts

<sup>1</sup>Wert, Neidt, and Ahmann, op.cit., Table III, pp. 406-410.

<sup>2</sup>ibid., Table IV, pp. 411-415.

the biserial correlation coefficients obtained, has also been computed by interpolation from a table supplied in the same volume.

Table 16, on page 93, presents the results of the application of the biserial correlation formula, the value of the tests of significance for each of the coefficients, and the approximate percentage level of confidence at which the coefficients are significant. The value of "t" was computed by the traditional formula:

$$t = \sqrt{\frac{\left[\left(\frac{z}{pq}\right)(r_{bis})\right]^2 (N-2)}{1 - \left[\left(\frac{z}{pq}\right)(r_{bis})\right]^2}}$$

It will be noted that (N - 2) degrees of freedom are used in the formula, which in this study is 33. The levels of confidence were also secured from a table in the same volume. Inferences from these results will be drawn at the end of this chapter, but it is of interest to note at this point that the biserial coefficients are highest for X<sub>1</sub>(ACE) and X<sub>4</sub>(English Usage); as might be expected, the intercorrelation coefficient between these two variables, as reported in Table 14, was also highest. Also, the biserial coefficient for X<sub>5</sub>(MMPI) is not significant, corresponding to the negligible amount of interrelationship between it and the remaining variables as revealed by the intercorrelation coefficients.

TABLE 16  
 BISERIAL CORRELATION COEFFICIENTS  
 AND THEIR SIGNIFICANCE

Vari- able	$r_{bis}$	t	(approximate) % level of significance <sup>1</sup>
X <sub>1</sub>	.507	2.26	3.5
X <sub>2</sub>	.386	1.67	10.0
X <sub>3</sub>	.394	1.70	9.0
X <sub>4</sub>	.458	2.01	5.0
X <sub>5</sub>	.048	0.199	85.0

FORMULAE:

$$r_{bis} = \frac{d}{G} \left( \frac{pq}{z} \right)$$

$$t = \frac{\left[ \left( \frac{z}{\sqrt{pq}} \right) (r_{bis}) \right]^2 (N - 2)}{1 - \left[ \left( \frac{z}{\sqrt{pq}} \right) (r_{bis}) \right]^2}$$

X<sub>1</sub> = ACE

X<sub>2</sub> = COOP MATH

X<sub>3</sub> = WATSON-GLASER

X<sub>4</sub> = ENGLISH USAGE

X<sub>5</sub> = MMPI

<sup>1</sup> ibid., Table VI, p. 418.

### B. Discriminant Analysis Technique

The discriminant equation technique is somewhat similar to the multiple regression equation method. The only important difference is this: whereas the latter yields appropriate weights for utilizing more than one variable in predicting a numerical criterion, the discriminant function is used in ascertaining appropriate weights for more than one variable in obtaining maximum separation between the two categories of a dichotomy. The basic assumptions have already been discussed. It is sufficient to point out that the discriminant function goes one step farther than biserial correlation. Whereas the biserial "r" determines the relationship between one numerical variable and a dichotomized variable, the discriminant function can predict a dichotomy from several numerical variables combined. Thus it should be possible to obtain a combination of variables that will be most potent in predicting a variable dichotomy.

Before the discriminant function could be applied to the data in Table 13, certain preliminary numerical computations from the raw scores had to be made. These are the sums, squares of sums, sums of squares, and sums of cross-products, as presented in Table 17, on page 95. In order to reduce the magnitude of these numerical variables, they had to be transmuted in terms of deviations from the general means. The formulas used and the values obtained are presented in Table

TABLE 17

ANALYSIS DATA FROM RAW SCORES  
 OF PREDICTION VARIABLES  
 FOR THE WHOLE SAMPLE  
 SHOWING  
 SUMS, SQUARES OF SUMS, SUMS OF SQUARES  
 AND SUMS OF CROSS-PRODUCTS

SUMS	SQUARES OF SUMS	SUMS OF SQUARES	SUMS OF CROSS-PRODUCTS	
$\Sigma X_1$ 4,508	$(\Sigma X_1)^2$ 20,322,064	$\Sigma X_1^2$ 594,658	$\Sigma X_1 X_2$ 171,996	$\Sigma X_2 X_4$ 325,559
$\Sigma X_2$ 1,290	$(\Sigma X_2)^2$ 1,664,100	$\Sigma X_2^2$ 51,800	$\Sigma X_1 X_3$ 321,306	$\Sigma X_2 X_5$ 618,835
$\Sigma X_3$ 2,464	$(\Sigma X_3)^2$ 6,071,296	$\Sigma X_3^2$ 176,142	$\Sigma X_1 X_4$ 1,127,467	$\Sigma X_3 X_4$ 613,362
$\Sigma X_4$ 8,629	$(\Sigma X_4)^2$ 74,459,641	$\Sigma X_4^2$ 2,163,829	$\Sigma X_1 X_5$ 2,151,837	$\Sigma X_3 X_5$ 1,177,268
$\Sigma X_5$ 16,716	$(\Sigma X_5)^2$ 279,424,656	$\Sigma X_5^2$ 8,026,990	$\Sigma X_2 X_3$ 92,940	$\Sigma X_4 X_5$ 4,127,884

18, on page 97. These values are later substituted directly into the discriminant equations.

Various combinations from the five variables will be used in this study. Only the first combination of all five variables will be given in detail here. The results of the remaining combinations will be summarized in a table with no further explanations.

The discriminant equation for five variables may be expressed as:

$$v = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5,$$

where  $a_1, a_2, a_3, a_4, a_5$  are the coefficients of variables  $x_1, x_2, x_3, x_4, x_5$ . The values of the coefficients are found by solving a series of simultaneous equations. For five variables they are:

- (1)  $Nzd_1 = a_1 \sum x_1^2 + a_2 \sum x_1x_2 + a_3 \sum x_1x_3 + a_4 \sum x_1x_4 + a_5 \sum x_1x_5$
- (2)  $Nzd_2 = a_1 \sum x_1x_2 + a_2 \sum x_2^2 + a_3 \sum x_2x_3 + a_4 \sum x_2x_4 + a_5 \sum x_2x_5$
- (3)  $Nzd_3 = a_1 \sum x_1x_3 + a_2 \sum x_2x_3 + a_3 \sum x_3^2 + a_4 \sum x_3x_4 + a_5 \sum x_3x_5$
- (4)  $Nzd_4 = a_1 \sum x_1x_4 + a_2 \sum x_2x_4 + a_3 \sum x_3x_4 + a_4 \sum x_4^2 + a_5 \sum x_4x_5$
- (5)  $Nzd_5 = a_1 \sum x_1x_5 + a_2 \sum x_2x_5 + a_3 \sum x_3x_5 + a_4 \sum x_4x_5 + a_5 \sum x_5^2$

where  $N$  is the total number of cases,  $z$  is the height of the ordinate dividing the normal curve into  $p$  and  $q$  parts, and the  $d$ 's are the differences in the means of the numerical variables in the dichotomy as given in Table 15. When values from Table 18 are substituted into these five equations, the

TABLE 18

ANALYSIS DATA FROM DEVIATION SCORES  
OF PREDICTION VARIABLES  
FOR THE WHOLE SAMPLE  
SHOWING  
SUMS OF SQUARES AND SUMS OF CROSS-PRODUCTS  
(N = 35)

SUMS OF SQUARES	SUMS OF CROSS-PRODUCTS	
$\sum x_1^2$ 14,027.6	$\sum x_1x_2$ 5,844	$\sum x_2x_4$ 7,518.7
$\sum x_2^2$ 5,254.3	$\sum x_1x_3$ 3,942.8	$\sum x_2x_5$ 2,731
$\sum x_3^2$ 2,676.4	$\sum x_1x_4$ 16,051.8	$\sum x_3x_4$ 5,880.4
$\sum x_4^2$ 36,410.7	$\sum x_1x_5$ -1,183.8	$\sum x_3x_5$ 461.6
$\sum x_5^2$ 43,428.4	$\sum x_2x_3$ 2,124	$\sum x_4x_5$ 6,673.6
FORMULA: $\sum x^2 = \sum X^2 - \frac{(\sum X)^2}{N}$	FORMULA: $\sum x_1x_2 = \sum X_1X_2 - \frac{(\sum X_1)(\sum X_2)}{N}$	

following set of equations is obtained:

$$(1) 14,027.6a_1 + 5,844a_2 + 3,942.8a_3 + 16,051.8a_4 - 1,183.8a_5 = 182.6$$

$$(2) 5,844a_1 + 5,254.3a_2 + 2,124a_3 + 7,518.7a_4 + 2,731a_5 = 87.3$$

$$(3) 3,942.8a_1 + 2,124a_2 + 2,676.4a_3 + 5,880.4a_4 + 461.6a_5 = 62.1$$

$$(4) 16,051.8a_1 + 7,518.7a_2 + 5,880.4a_3 + 36,410.7a_4 + 6,673.6a_5 = 271.7$$

$$(5) -1,183.8a_1 + 2,731a_2 + 461.6a_3 + 6,673.6a_4 + 43,428.4a_5 = 30.5$$

Solving these equations yields the following values of the coefficients:

$$a_1 = .00774$$

$$a_4 = .00299$$

$$a_2 = .00219$$

$$a_5 = .00028$$

$$a_3 = .00345$$

The original discriminant equation then becomes:

$$v = .00774x_1 + .00219x_2 + .00345x_3 + .00299x_4 + .00028x_5$$

where  $v = \frac{X}{\sigma}$  score in deviation form, and  $x_1, x_2, x_3, x_4, x_5$  are deviation scores for each of the variables. If one were satisfied with the predictive strength of this combination of all five variables, this equation now can be used for predicting the v-score for each individual, by substituting into the equation the individual's deviation scores from each of the five variables.

However, the purpose here is that of arriving at the best possible combination for prediction. Thus other pre-

diction formulas must be deduced and their significance must be tested. Consequently, it is necessary to calculate the multiple biserial correlation coefficient ( $R_{bis}$ ) and then test its significance. To accomplish that, the difference between the predicted means in the dichotomy must first be found by the use of the formula:

$$\Delta = a_1 N z d_1 + a_2 N z d_2 + a_3 N z d_3 + a_4 N z d_4 + a_5 N z d_5,$$

where  $\Delta$  = difference in means of predicted scores in the dichotomy, and  $d_1, d_2, d_3, d_4, d_5$  are the difference in means in the prediction variables as given in Table 15. Substituting the values obtained for  $a_1, a_2, a_3, a_4, a_5$ , the following equation is formed:

$$\begin{aligned} \Delta &= 182.6(.00774) + 87.3(.00219) + 62.1(.00345) + 271.7(.00299) \\ &+ 30.5(.00028), \text{ or } \Delta = 1.413324 + .191187 + .214245 + .812383 \\ &+ .009540, \text{ making } \Delta = 2.639679. \end{aligned}$$

The multiple biserial correlation coefficient is then calculated from the formula:

$$R_{bis} = \frac{pq}{z^2} \sqrt{\frac{\Delta}{N}}$$

where  $\frac{pq}{z^2}$  is found from the reciprocal of  $\frac{z^2}{pq}$  and by interpolation from a prepared table.<sup>1</sup> By substitution,  $R_{bis} = (1.9257) \sqrt{\frac{2.639679}{35}}$ , or  $(1.9257)(.2746)$ , making  $R_{bis} = .529$ .

---

<sup>1</sup>Wert, Neidt and Ahmann, op.cit., Table IV, pp. 411-415.

In order to test the significance of the coefficient so obtained, the following F-formula is used:

$$F_{m, N-m-1} = \frac{(N-m-1)\Delta}{\left(\frac{Nz^2}{pq} - \Delta\right)_m},$$

where  $N = 35$ ,

$m =$  number of variables used (5),

$\frac{z^2}{pq}$ , by interpolation from the table referred to immediately above.

The application of this formula makes  $F = .985$ . This F-value, with 5 and 29 degrees of freedom, according to Table 21, on page 104, which is a reproduction of a standard table,<sup>1</sup> is not significant even at the 20% level.

Often it is desirable to indicate the relative effectiveness of the individual variables used in the combination. Noting the relative size of the coefficients is not sufficient. The method used here is that of comparing the relative contributions of the individual variables to the numerical value of " $\Delta$ ". These contributions are deduced from the equation:

$$\Delta = 1.41332 + .19119 + .21425 + .81238 + .00854, \text{ used previously.}$$

The following are the results:

$X_1$ (ACE).....	53.5%
$X_2$ (Math).....	7.2%
$X_3$ (Watson-Glaser).....	8.2%
$X_4$ (English Usage).....	30.8%
$X_5$ (MMPI).....	.3%

---

<sup>1</sup>See footnote below Table 21, on page 104.

The relative effectiveness of the individual variables applies only when that particular combination is used. Removal of a variable from the combination may change the relative importance of the remaining variables radically. Consequently, the procedure must be repeated for every combination.

As indicated earlier, it will be impossible to give in detail all the steps for further combinations. The results of the foregoing combination and all others that were similarly treated are presented in Tables 19, 20, and 22, on pages 102, 103 and 105 respectively. The symbols used in those tables need no further explanation. Only one point needs clarification here. It will be noted that variable  $X_5$ (MMPI) was used only in the first combination. Both its " $r_{bis}$ " and its contribution in the discriminant equation were so low as to make them insignificant. For that reason it was dropped from all future computations.

A word of explanation pertaining to the use of Table 21 now seems in order. As an illustration, the first combination of all five variables is used here. For 5 and 29 degrees of freedom, the F-value was found to be .985. The table shows that for these degrees of freedom the F-value should be at least 1.57 in order to be significant at the 20% level, at least 2.05 to be significant at the 10% level, at least 2.54 to be significant at the 5% level, and at least 3.73 to be significant at the 1% level. Of course, the F-value found for this combination is too low to be significant

TABLE 19

VALUES OF COEFFICIENTS AS OBTAINED BY SOLVING  
VARIOUS SETS OF DISCRIMINANT EQUATIONS

COMBINATION OF VARIABLES	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>5</sub>
x <sub>1</sub> + x <sub>2</sub> + x <sub>3</sub> + x <sub>4</sub> + x <sub>5</sub>	.00774	.00219	.00345	.00299	.00028
x <sub>1</sub> + x <sub>2</sub> + x <sub>3</sub> + x <sub>4</sub>	.00747	.00248	.00340	.00311	////////
x <sub>1</sub> + x <sub>2</sub> + x <sub>3</sub>	.01011	.00296	.00596	////////	////////
x <sub>1</sub> + x <sub>2</sub> + x <sub>4</sub>	.00795	.00300	////////	.00334	////////
x <sub>1</sub> + x <sub>3</sub> + x <sub>4</sub>	.00824	////////	.00411	.00317	////////
x <sub>2</sub> + x <sub>3</sub> + x <sub>4</sub>	////////	.00670	.00702	.00494	////////
x <sub>1</sub> + x <sub>2</sub>	.01136	.00398	////////	////////	////////
x <sub>1</sub> + x <sub>3</sub>	.01109	////////	.00687	////////	////////
x <sub>1</sub> + x <sub>4</sub>	.00904	////////	////////	.00348	////////
x <sub>2</sub> + x <sub>3</sub>	////////	.01023	.01579	////////	////////
x <sub>2</sub> + x <sub>4</sub>	////////	.00843	////////	.00572	////////
x <sub>3</sub> + x <sub>4</sub>	////////	////////	.01055	.00576	////////

X<sub>1</sub> = ACE PSY.

X<sub>3</sub> = WATSON-GLASER

X<sub>4</sub> = ENGLISH

X<sub>2</sub> = COOP MATH

X<sub>5</sub> = MMPI

TABLE 20

MULTIPLE BISERIAL CORRELATION COEFFICIENTS  
OF DIFFERENT COMBINATIONS OF VARIABLES  
AND THEIR RELATIVE SIGNIFICANCES

COMBINATION OF VARIABLES	$\Delta$	$R_{bis}$	F-VALUE	d.f.	Approx. % significance <sup>1</sup>
$x_1 + x_2 + x_3 + x_4 + x_5$	2.63698	.529	.985	5,29	Not sig. at 20%
$x_1 + x_2 + x_3 + x_4$	2.63665	.528	1.273	4,30	Not sig. at 20%
$x_1 + x_2 + x_3$	2.47461	.512	1.629	3,31	20%
$x_1 + x_2 + x_4$	2.62891	.528	1.747	3,31	20%
$x_1 + x_3 + x_4$	2.63114	.528	1.749	3,31	20%
$x_2 + x_3 + x_4$	2.36305	.501	1.544	3,31	Not sig. at 20%
$x_1 + x_2$	2.42179	.506	2.206	2,32	15%
$x_1 + x_3$	2.45166	.510	2.495	2,32	10%
$x_1 + x_4$	2.59622	.524	2.666	2,32	9%
$x_2 + x_3$	1.87434	.435	1.839	2,32	20%
$x_2 + x_4$	2.29007	.475	2.361	2,32	12%
$x_3 + x_4$	2.22015	.483	2.227	2,32	13%

FORMULAE:

$$\Delta = a_1 N z d_1 + a_2 N z d_2 + a_3 N z d_3 \text{ -----}$$

$$R_{bis} = \frac{pq}{z^2} \sqrt{\frac{\Delta}{N}}$$

$$F_{m, N-m-1} = \frac{\Delta (N-m-1)}{\left( \frac{Nz^2}{pq} - \Delta \right) m}$$

<sup>1</sup> See Table 21, p. 104.

TABLE 21

THE 20, 10, 5 and 1 PERCENT VALUES  
OF "F"<sup>1</sup>

DEGREES OF FREEDOM	% LEVEL	DEGREES OF FREEDOM (FOR GREATER MEAN SQUARE)					
		1	2	3	4	5	6
28	20	1.72	1.71	1.65	1.61	1.57	1.55
	10	2.89	2.50	2.29	2.16	2.06	2.00
	5	4.20	3.34	2.95	2.71	2.56	2.44
	1	7.64	5.45	4.57	4.07	3.76	3.53
29	20	1.72	1.70	1.65	1.60	1.57	1.54
	10	2.89	2.50	2.28	2.15	2.06	1.99
	5	4.18	3.33	2.93	2.70	2.54	2.43
	1	7.60	5.42	4.54	4.04	3.73	3.50
30	20	1.72	1.70	1.64	1.60	1.57	1.54
	10	2.88	2.49	2.28	2.14	2.05	1.98
	5	4.17	3.32	2.92	2.69	2.53	2.42
	1	7.56	5.39	4.51	4.02	3.70	3.47
32	20	1.71	1.69	1.64	1.59	1.56	1.53
	10	2.87	2.48	2.27	2.13	2.05	1.97
	5	4.15	3.30	2.90	2.67	2.51	2.40
	1	7.50	5.34	4.46	3.97	3.66	3.42
40	20	1.70	1.68	1.62	1.57	1.54	1.51
	10	2.84	2.44	2.23	2.09	2.00	1.93
	5	4.08	3.23	2.84	2.61	2.45	2.34
	1	7.31	5.18	4.31	3.83	3.51	3.29

<sup>1</sup>Reproduced from two sources:

A. Wert, Neidt, and Ahmann, op.cit., Table VII, pp. 419-422.

B. Fisher and Yates, Statistical Tables for Biological, Agricultural and Medical Research, Edinburgh: Oliver and Boyd, Tweeddale Court, 1948, Table V, pp. 34-43.

TABLE 22

RELATIVE EFFECTIVENESS OF INDIVIDUAL VARIABLES IN  
THEIR CONTRIBUTION TO THE  
TOTAL PREDICTIVE VALUE

COMBINATION OF VARIABLES	PERCENTAGE CONTRIBUTION				
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
$x_1 + x_2 + x_3 + x_4 + x_5$	53.5	7.2	8.2	30.8	.3
$x_1 + x_2 + x_3 + x_4$	51.7	8.2	8.0	32.1	////
$x_1 + x_2 + x_3$	74.6	10.4	15.0	////	////
$x_1 + x_2 + x_4$	55.2	10.3	////	34.5	////
$x_1 + x_3 + x_4$	57.6	////	9.7	32.7	////
$x_2 + x_3 + x_4$	////	24.7	18.4	56.9	////
$x_1 + x_2$	85.6	14.4	////	////	////
$x_1 + x_3$	82.6	////	17.4	////	////
$x_1 + x_4$	63.6	////	////	36.4	////
$x_2 + x_3$	////	47.7	52.3	////	////
$x_2 + x_4$	////	32.1	////	67.9	////
$x_3 + x_4$	////	////	29.5	70.5	////

$x_1$  = ACE PSY.

$x_3$  = WATSON-GLASER

$x_4$  = ENGLISH

$x_2$  = COOP MATH.

$x_5$  = MMPI

even at the 20% level. Other levels of significance were read off the table in a similar manner or adjusted by interpolation where necessary.

A closer scrutiny and an interpretation of the findings of the application of both the biserial correlation and the discriminant analysis will be presented in the chapter to follow.

## CHAPTER VI

### FINDINGS AND CONCLUSIONS

#### A. Interpretation and Value of Findings

The review of literature, undertaken in Chapter II, showed that very few findings could be designated as "the findings" with any semblance of universal application. This was due to the multiplicity of approaches in the various studies. Nevertheless, it was also shown that many studies did find fair-to-good predictive value in measures of intelligence and in measures of academic achievement, both in high school and in college. Unfortunately, many of these same studies seemed to belittle or even ignore the significance of these two factors and proceeded to seek out more tangible relationships between teaching success and other factors, some of which have no obvious relation to teaching. In this latter group could be included left-handedness, rate of tapping, and height, whose bearing on teaching success could never be more than remote and accidental. It seems that at least some of the research workers accept a philosophy which allows them to admit only reluctantly that teachers must be intelligent and "know their stuff" before they can be successful. Fortunately, however, from the same review, it was possible to arrive at some consensus of opinion regarding general procedures and techniques. It

is in terms of these broad generalities that the findings of this study will be interpreted first, and in this area the following general conclusions seem to be warranted:

1. With a little more cooperation from those enlisted for the task of rating teaching performance, there is no doubt that supervisors' ratings, in composite form, could be quite reliable. In most cases, there was remarkable consistency in the ratings obtained from different "experts" on the same teacher.
2. Judging from the general lack of agreement over the exact nature of the factors relating to teaching success and from the reluctance of raters to rank or numerically evaluate teaching performance, categorization seems to be the only solution to the problem. To avoid "dumping" doubtful cases into some middle category, it is preferable to have an even number of categories rather than an odd number for the purpose of setting up a dichotomy.
3. Raters in this study were very favorably impressed with the practice of using a scale of general concepts of qualities relating to teaching performance rather than a comprehensive list of specific factors. Supervisors like to be allowed a good deal of freedom to temper suggested criteria with their own subjective judgment.

4. A number of raters qualified their categorization of teaching performance on the grounds that the teaching quality could change radically in a different situation. This emphasizes the importance of the situational factor in the evaluation of teaching efficiency. If possible, ratings in future studies should be postponed beyond the second year of teaching.
5. Both the biserial coefficient technique and the discriminant analysis method appear to be powerful weapons in determining the relative values of numerical variables, individually and in combinations, in predicting a variable dichotomy. However, for the discriminant equation method particularly, larger samples must be used. Considering that  $N-m-1$ , where "m" is the number of variables, represents the degrees of freedom in calculating the "F-value", it is not difficult to see how important it is to have a large sample and a correspondingly small number of variables, if the results are to be significant.

Such is the general tone of the findings pertaining to techniques and procedures, not unlike those of the review of previous studies in the field.

In terms of these general observations and within the limitations emphasized throughout the earlier chapters, the more specific findings of the "pilot study" may now be

isolated and discussed.

1. From the intercorrelations between the prediction variables as given in Table 14 on page 86, it is evident that there is a distinct relationship between the ACE measures on the one hand and the Coop Math, Watson-Glaser, and English Usage measures on the other, with the highest positive correlation ( $r = .710$ ) between the ACE and the English Usage. The relationships amongst the last three is not so marked. As might be expected, the relationships between the MMPI and the other four variables is negligible. Bearing in mind that no group intelligence test is "language-free", it is possible that it is the language factor of the ACE that produces the relatively high positive correlation between  $X_1$  and  $X_4$ . These intercorrelations are important in the interpretation of the results of the application of the discriminant analysis. Reference to them will be made again later.
2. The biserial correlation coefficients as given in Table 16 on page 93 show a significant relationship between intelligence, as measured by the ACE test, and teaching success. A biserial coefficient ( $r_{bis}$ ) of .507, significant at the  $3\frac{1}{2}\%$  level, cannot be slighted. Interesting too is the finding that the English Usage measures are second-best in their predictive value,

with a coefficient of .458, significant at the 5% level. Both mathematical ability and critical thinking as measured by the Coop Math and Watson-Glaser tests respectively have little predictive value. Of even greater interest is the finding that the MMPI measures have no predictive value whatsoever. This, of course, does not mean that there is no significant relationship between the teacher's personality and her teaching performance. What it does indicate, within the limitations of this study, is that there is no significant relationship between teaching success and personality as measured by the MMPI. It is generally speculated that the MMPI, not unlike many other personality inventories, evaluates general personality, which need not necessarily be closely related to the so-called "teacher personality". Educators often point to the existence of these two, often radically different, personalities. It is not difficult to recall within one's own experience a teacher who in public, in community life, as a person, exhibits mechanisms of withdrawal, seclusion, self-consciousness, but in the classroom she dominates the situation with confidence and assertion, at the same time maintaining an admirable classroom spirit of friendliness and cooperation. Many other such alleged discrepancies in the two per-

sonalities could easily be cited.

3. Table 20, on page 103, presents the multiple biserial correlation coefficients ( $R_{bis}$ ) for various combinations of variables, based on solutions of discriminant equations involving the combinations. The table also supplies the F-values for the combinations and the approximate level of significance for each F-value. It should be noted that the relationship between the value of  $R_{bis}$  and the level of significance is not linear nor proportional because of the variations in the degrees of freedom. According to the findings here a combination of  $x_1$  (ACE) and  $x_4$  (English Usage) yields maximum separation in the dichotomous variable, although even this is hardly powerful enough for prediction purposes. It will be recalled that individually  $x_1$  and  $x_4$  yielded  $r_{bis}$  coefficients significant at the  $3\frac{1}{2}\%$  and  $5\%$  levels respectively. Here, their combined strength is lower than the weaker of the individual values. Apparently, their combined predictive value is reduced to a large extent by the fact that they are related variables as indicated by their relatively high intercorrelation coefficient of .710, as given in Table 14 on page 86. This is as it should be. One cannot expect to get a more powerful prediction tool by combining two somewhat powerful but closely related tools. On the other hand, one should not expect to obtain a more powerful tool by combining a somewhat

powerful one with one whose predictive value is negligible. For this reason the MMPI( $x_5$ ) variable was dropped from all but the first combination. Apparently, tools must individually be powerful and unrelated if they are to become better weapons in combinations. Proper balance between individual predictive strength and interrelatedness seems to be the key to the search for effective combinations of variables. Herein lies the unique and potential power of the discriminant function in the field of grouped prediction of a dichotomy. However, it should again be emphasized that all of the  $R_{bis}$  coefficients lost considerable predictive value because of the limited sample used. The formula for the F-value bears that out clearly. It is not unlikely that with a larger sample, which might increase their individual predictive value and reduce their interrelatedness, other things being equal, a combination of  $x_1$  and  $x_4$  may be a very useful measure of future teaching success.

4. Table 22, on page 105, presents the relative effectiveness of individual variables in their contributions to the total predictive strength of their combination with other variables. Whatever the combination, both  $x_1$  and  $x_4$  consistently stand out with their relatively high contributions. It is interesting to note how the

emphasis may shift from one variable to another with the removal of other variables from the combinations. Interrelatedness again seems to play an important part in these shifts.

5. As pointed out in the preceding chapter, once one is satisfied with the predictive strength of a certain combination of variables, the next step would be that of setting up the prediction equation. As an illustration one further example will be used here. Assuming that a combination of  $x_1$  and  $x_4$  yields the desired amount of discrimination in the variable dichotomy, the prediction equation would take this form:  $v = a_1x_1 + a_4x_4$  where  $v$  is the  $\frac{x}{\sigma}$  score in deviation form;  $a_1$ ,  $a_4$  are the values of the coefficients obtained by solving a set of two discriminant equations in  $x_1$  and  $x_4$ ; and  $x_1$ ,  $x_4$  are deviate scores for each of the two variables for the individual whose teaching efficiency is being predicted. Substitution of the known values in the right-hand member of the prediction equation will yield a  $v$ -score for the individual. For convenience, however, the foregoing deviation form of the prediction equation may be changed to the raw-score form:  $V - \bar{V} = a_1(X_1 - \bar{X}_1) + a_4(X_4 - \bar{X}_4)$ , where  $X_1$  and  $X_4$  are the individual's raw scores on the ACE and the English Usage tests,  $\bar{X}_1$  and  $\bar{X}_4$  are the means of the scores for the group, and  $\bar{V}$  is the

normal deviate in terms of probability deduced either from the sample in which the discriminant equation was developed or from the percentage of successful teachers, based upon experience over a period of years. By substitution, the value of  $V$ , in sigma units, may be obtained, which upon consulting a table of the normal curve will estimate the probability of success for the particular individual. Other methods of arriving at the value of  $V$  have been used, but enlarging on these is scarcely the purpose of this thesis. Once a sufficiently powerful tool for prediction is determined, no difficulty should be encountered in establishing some procedure for its use in actual prediction. Any recent text in statistics will offer a few possible and equally effective approaches.

6. In order to conclude this section it only remains to point out that the findings of the pilot study are quite consistent with the consensus of findings of previous research in the field. True, some of the procedures and techniques recommended by the review of previous studies have been modified to fit the local situation. Restricted by the size of the sample and by the other limitations indicated throughout the study, one must be on guard when interpreting these findings lest one indulge in generalizations not warranted by this somewhat selec-

tive and limited piece of endeavour.

B. Recommendations for Further Research

Considering the exploratory nature of this thesis and the limited scope of the pilot study, the writer would be more than ambitious if he attempted to formulate any definite plans for research in the field at large. Furthermore, any such general recommendations would be no different from those that can be found appended to any of the many extensive studies which he had reviewed and which have already been highlighted in Chapter II. For these reasons the recommendations given below will pertain more specifically to what can be done here, in the local situation. Some of these recommendations will suggest ways of repeating and expanding this study; others offer suggestions for possible changes in the teacher-training study undertaken here by the faculty.

1. It is unlikely that the size of the annual sample of subjects available for study, similar to the one used in this study, will increase appreciably within the next few years. Therefore, in the next few years any repetition of the study, which is here recommended, will have to be limited to the variables attempted in this one. Such repetition may serve to validate some of the findings, particularly those dealing with procedures and techniques. In view of the relative success with which the proposed techniques were applied in this

study it is recommended that repeat studies adopt the same or at least similar techniques in securing measures of variables, in securing ratings of teaching performance, in classifying and recording data, and in statistically treating the data. Samples, comparable in size to the one used in this study, will be available for at least two repeat studies.

2. In later years, should the size of the sample warrant the inclusion of more variables, it is recommended that as many as possible of the following additional measures be used as prediction variables:

- (a) Scores on the MTAI (Minnesota Teacher Attitude Inventory)
- (b) High school and/or college grade-point averages
- (c) Measures from some personality inventory other than the MMPI.

For purposes of selective retention rather than selective admission, consideration should be given to two additional measures as prediction variables:

- (d) Practice-teaching scores
- (e) Xmas and/or final examination scores in Education I classes.

If practice-teaching scores are to be considered, the grade-scores used presently would have to be quantified into numerical measures. In deriving measures of achievement in Education I classes, averages of standard scores on a uniform scale

could be used.

3. Should the sample remain inadequate for the above purpose, it is recommended that a similar study be conducted at one of the Provincial Normal Schools.
4. Whatever form future studies should take, it is important that the number of variables used be no larger than the size of the sample will warrant. Also, it is important that the smaller of the two categories in the dichotomous variable be at least 10%, preferably 15%, of the total sample for an effective application of both the biserial correlation and the discriminant function.
5. Although it is recommended that some measure of personality other than the MMPI be used as an estimate of personality adjustment for the purpose of prediction, there is no reason why the administering of the MMPI should be discontinued. Its measures will always have considerable value for guidance purposes during institutional training.
6. It will be noted that although Group Rorschach measures are available on the next two samples, it is not recommended that these measures be used as a prediction variable. As a matter of fact, the writer is tempted to recommend the elimination of this test from the battery administered to the newly-enrolled candidates. Most of the recent studies

that have attempted to find some predictive value in this test discount heavily any such possibility. Perhaps typical of such studies is the one conducted by Cooper in 1949.<sup>1</sup> His findings illustrate the inconsistency of measures obtained and show little likelihood that there may be predictive value in the performance on the Group Rorschach as scored by the IR (Monroe) method.

7. In order to facilitate the recording and locating of information about the sample(s) and to maintain a more permanent and more readily accessible record of data it is recommended that the filing system proposed earlier in the thesis be kept up faithfully. A sample copy of the proposed individual filing card is supplied in the appendix. The required data should be inserted as soon as they become available. The card has provision for recording all the types of data to which this study has referred. Additional provisions, as required, could easily be accommodated.

As a concluding note and in general agreement with the remarks of A. S. Barr and other experts like Barr, who have reviewed many studies in the field, the writer leaves this thought. Although no sweeping generalizations are

---

<sup>1</sup>James G. Cooper, "The Inspection Rorschach in the Prediction of College Success," Journal of Educational Research, December 1955, pp. 275-282.

warranted, there is reasonable assurance that it is possible to find several factors, individually and in combinations, that are closely enough related to teaching to make prediction of teaching success possible. Further research will no doubt uphold that contention. For similar reasons it may be said that the teacher-training study undertaken by the faculty here is both worthwhile and feasible, and further research based on this preliminary survey should be encouraged and extended. This present study is only one of hundreds of similar modest attempts to add some knowledge to our understanding of the effective teacher. The educational process cannot be improved through the efforts of educators, psychologists and philosophers working alone. This study may be an insignificant and intangible contribution for the common good, but it is hoped that this research effort may be as helpful to others working in the area of teaching effectiveness, particularly in the local situation, as the many studies examined have been helpful to the writer.

## BIBLIOGRAPHY

### Articles

- Bach, Olav Jacob. "Practice Teaching Success in Relation to Other Measures of Teaching Ability," Journal of Experimental Education, Sept. 1952, pp. 57-80.
- Baker, F. E. "Selective Admission and Selective Promotion in Teacher Education Institutions," Yearbook XXIII, National Society of College Teachers of Education, University of Chicago, 1935, pp. 16-71.
- Barr, A. S. "Measurement of Teaching Ability," and "Prediction of Teaching Success," Review of Educational Research, 10 (1940), 182-190.
- Barr, A. S. "Measurement and Prediction of Teaching Efficiency," and "Recruitment, Institutional Selection and Guidance of Teachers," Review of Educational Research, 13 (1943), 218-227.
- Barr, A. S. "Summary and Comments," Journal of Experimental Education, 15 (Sept. 1946), 99-100.
- Barr, A. S. "The Measurement and Prediction of Teaching Efficiency," Review of Educational Research, 16 (1946), 203-208.
- Barr, A. S. "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigations," Journal of Experimental Education, 16 (June 1948), 195-226.
- Barr, A. S. "Measurement and Prediction of Teaching Efficiency," Review of Educational Research, 25 (June 1955), 261-269.
- Bretsch, H. S. and Jacobson, G. S. "Recruitment, Guidance and Screening of Prospective Teachers," Review of Educational Research, June 1955, p. 204.
- Butsch, R. L. C. "Teacher Rating," Review of Educational Research, 1 (1931), 99-107, 156-157.
- Carlile, Amos B. "Predicting Performance in the Teaching Profession," Journal of Educational Research, 47 (May 1954), 641-668.

- Cooper, James G. "The Inspection Rorschach in the Prediction of College Success," Journal of Educational Research, December 1955, pp. 275-282.
- Eliasson, R. H. and Martin, R. L. "Pretraining Selection of Teachers During 1940-43," Journal of Educational Research, 38 (1945), 666-677.
- Fulkerson, Glen. "A Resumé of Current Teacher Personnel Research," Journal of Educational Research, May, 1954, pp. 669-671.
- Gough, G. H. and Pemberton, W. H. "Personality Characteristics Related to Success in Practice Teaching," Journal of Applied Psychology, 36 (Oct. 1952), 307-309.
- Hale, Peter P. "Isolating Objective Factors for the Teaching Profession," Journal of Educational Research, March 1955, pp. 497-507.
- Jones, Ronald D. "The Prediction of Teaching Efficiency from Objective Measures," Journal of Experimental Education, 15 (Sept. 1946), 85-99.
- Knoell, Dorothy M. "The Prediction of Teaching Success from Word Fluency Data," Journal of Educational Research, 46 (May 1953), 673-683.
- Kriner, H. L. "Five-Year Study of Teacher College Admissions," Educational Administration and Supervision, 23 (1937), 192-199.
- La Duke, C. V. "The Measurement of Teaching Ability," Journal of Experimental Education, 14 (1945), 75-100.
- Lamke, Tom Arthur, "Personality and Teaching Success," Journal of Experimental Education, Dec. 1951, p. 253.
- Leavitt, Jerome E. "Personal Data and Prediction of Success of Student Teachers," Journal of Teacher Education, 4 (Sept. 1953), 194-197.
- Lins, Leo Joseph. "The Prediction of Teaching Efficiency," Journal of Experimental Education, 15 (Sept. 1946), 59.
- MacLean, Gowan and Gowan. "A Teacher Selection and Counseling Service," Journal of Educational Research, May 1955, pp. 669-677.

- Magee, Robert M. "Selection of Candidates for Teacher Education," Journal of Teacher Education, 3 (Sept. 1952), 168-172.
- Mead, A. R. "Questions of Merit in Good and Poor Teachers," Journal of Educational Research, 20 (Nov. 1929), 257-258.
- Michaelis, John V. "The Prediction of Success in Student Teaching from Personality and Attitude Inventories," University of California Publications in Education, 11 (1954), 415-481.
- Montross, Harold W. "Temperament and Teaching Success," Journal of Experimental Education, 23 (Sept. 1954), 73-97.
- Odenweller, A. L. "Predicting the Quality of Teaching," Contributions to Education, No. 676, Teachers' College, 1936, p. 158.
- Reed, Harold J. "An Investigation of the Relationship Between Teaching Effectiveness and the Teacher's Attitude of Acceptance," Journal of Experimental Education, June 1953, pp. 277-325.
- Ringness, T. A. "Relationships Between Certain Attitudes Towards Teaching and Teaching Success," Journal of Experimental Education, 21 (Sept. 1952), 1-55.
- Rolfe, J. F. "The Measurement of Teaching Ability," Journal of Experimental Education, 14 (1945), 52-74.
- Rostker, Leon E. "The Measurement and Prediction of Teaching Ability," School and Society, 51 (Jan. 1940), 30-32.
- Sanford, C. W. and Trump, J. L. "Teacher Education - Pre-service Selection," Encyclopedia of Educational Research, Revised edition, New York: The Macmillan Co., 1952, p. 1392.
- Schmid, John Jr. "Factor Analysis of Prospective Teachers' Differences," Journal of Experimental Education, 18 (June 1950), 295.
- Symonds, Percival M. "Characteristics of the Effective Teacher Based on Pupil Evaluation," Journal of Experimental Education, 23 (June 1955), 289-310.

- Torgerson, T. L. "The Measurement and Prediction of Teaching Ability," Review of Educational Research, 4 (June 1934) 261-266, 329-330; 7 (1937) 242-246, 319-320.
- Troyer, M. E. and Pace, C. R. "Evaluation in Teacher Education," American Council on Teacher Education, 1944, p. 368.
- Ullman, Roy R. "The Prediction of Teaching Success," Educational Administration and Supervision, 16 (1930), 598-608.
- Von Haden, Herbert I. "An Evaluation of Certain Types of Personal Data Employed in the Prediction of Teaching Efficiency," Journal of Experimental Education, 15 (Sept. 1946), 83-84.
- White, Miss Verna. "Selection of Prospective Teachers at Syracuse University," Journal of Teacher Education, 1 (March 1950) 24-31.
- Witty, Paul. "An Analysis of the Personality Traits of the Effective Teacher," Journal of Educational Research, 40 (1947) 662-671.
- Yaukey, J. V. and Anderson, P. L. "A Review of the Literature on the Factors Conditioning Teacher Success," Educational Administration and Supervision, 19 (1933), 511-520.

#### Books

- Barr, Burton, and Brueckner. Supervision, New York: Appleton-Century-Crofts Inc. 1947.
- Bowers, Henry. Research in the Training of Teachers, J. M. Dent & Sons (Canada) Limited, The Macmillan Co. of Canada Ltd., 1952, pp. 1-167.
- Wert, Neidt and Ahmann. Statistical Methods in Educational and Psychological Research, New York: Appleton-Century-Crofts Inc., 1954.
- Whitney, Frederick Lamson. The Elements of Research, New York: Prentice-Hall Inc., (3rd ed.), pp. 106-107.

Bulletins

- Betts, G. L. "The Education of Teachers Evaluated Through Measurement of Teaching Ability," Special Survey Studies, U. S. Office of Education, Bulletin No. 10, 1933, National Survey of Education of Teachers, Vol. 5, Part 2, pp. 87-153.
- Brammell, P. R. "Articulation of High School and College," U. S. Office of Education, Bulletin No. 17, 1932, National Survey of Secondary Education, Monograph No. 10, pp. 96.
- Sandiford, Peter and Others. "Forecasting Teaching Ability," University of Toronto, Department of Educational Research, Bulletin No. 8, 1937, pp. 93.
- Stein, H. L. "Aptitudes and Achievements of Faculty Students," Research Bulletin No. 17, Faculty of Education, University of Manitoba, Dec. 1953.

Pamphlet

- National Education Association, Teacher Rating, Discussion Pamphlet No. 10, Washington: National Education Association, 1950.

Unpublished Material

- Gray, R. H. A. "Teacher Evaluation," Unpublished Master's Thesis, 1954, on file in the University of Manitoba, Faculty of Education, Library.

A P P E N D I X

## THE MANITOBA TEACHERS' SOCIETY

The following statement of Factors in Teaching Competence represents the co-operative work of many educators and laymen throughout California, working under the leadership of the California Council on Teacher Education. For this reason, the statement is commonly referred to as the "California Statement of Teaching Competence." It is presented to you with the compliments of The Manitoba Teachers' Society.

### FACTORS IN TEACHING COMPETENCE

#### 1. A Director of learning

- 1.1 Adapts psychological principles of learning to individuals and groups in providing:
  - 1.11 Effective and continuing motivation.
    - 1.111 Recognizes and utilizes pupil interests, abilities, and needs
    - 1.112 Utilizes pupil experiences and life situations
    - 1.113 Draws on interest appeal of the subject matter
    - 1.114 Uses intrinsic motivation primarily.
  - 1.12 A variety of learning experiences.
  - 1.13 Effective practice procedures.
  - 1.14 Effective review procedures.
  - 1.15 Experiences which possess meaning and structure for the pupil.
  - 1.16 For effective transfer.
  - 1.17 For individual capacities of pupils.
- 1.2 Adapts principles of child growth and development to planning of learning activities.
  - 1.21 Adapts learning experiences to needs and capacities, recognizing
    - 1.211 A developmental sequence of growth
    - 1.212 Social dynamics in children's groups
    - 1.213 Home influences
    - 1.214 Community influences
  - 1.22 Provides for differentiated activities and assignments
  - 1.23 Observes principles of mental hygiene in classroom activities
  - 1.24 Demonstrates intelligence in problems of pupil health and hygiene
- 1.3 Maintains an effective balance of freedom and security in classroom
  - 1.31 Demonstrates ability to plan co-operatively with pupils
  - 1.32 Develops increasingly pupil leadership and responsibility
  - 1.33 Provides democratic classroom organization and procedure
    - 1.331 Large and small group activities
    - 1.332 Opportunities for leadership and co-operation
  - 1.34 Provides opportunity for independent, critical thinking
    - 1.341 Emphasis on freedom of expression, open-mindedness
  - 1.35 Provides for wide participation, at various levels of ability
  - 1.36 Provides opportunity to develop attitudes deemed socially, psychologically, biologically desirable.
- 1.4 Demonstrates effective instructional procedures
  - 1.41 Skill in making assignments
  - 1.42 Presentation techniques
  - 1.43 Discussion techniques
  - 1.44 Effective development of pupil participation
  - 1.45 Stimulation of individual interests and creative activity
  - 1.46 Skillful use of supervised study
    - 1.461 Providing facility for study helps
    - 1.462 How-to-study techniques
  - 1.47 Developing self-evaluation procedures in pupils
- 1.5 Plans learning experiences effectively
  - 1.51 Leads pupils to define acceptable objectives
    - 1.511 For the longer units
    - 1.512 For daily classwork
    - 1.513 For special class activities
  - 1.52 Integrates subject material and teaching procedure with pupil need through selection and organization of:
    - 1.521 Subject matter content
    - 1.522 Learning experiences
    - 1.523 Materials of instruction

- 1.53 Selects, makes, and utilizes multisensory learning aids
  - 1.531 Blackboard, bulletin board, flat pictures
  - 1.532 Models, specimens, exhibits
  - 1.533 Radios, recordings
  - 1.534 Films, slides, sound
  - 1.535 Collateral literature
    - 1.5351 Teaching use of library
    - 1.5352 Utilization of references and supplementary material
    - 1.5353 Utilization of periodicals and newspapers
  - 1.536 Resource personnel
  - 1.537 Community resources
    - 1.5371 Life and home applications
    - 1.5372 Use of Community surveys
    - 1.5373 How to study the community informally
- 1.6 Is effective in classroom management
  - 1.61 Routinizes activities where appropriate
  - 1.62 Demonstrates careful planning in handling materials, equipment, supplies
  - 1.63 Uses time efficiently
  - 1.64 Controls physical aspects - heat, light, ventilation, etc.
  - 1.65 Is sensitive and competent in dealing with health and hygiene
- 1.7 Uses diagnostic and remedial procedures effectively
  - 1.71 Is familiar with common diagnostic tests in his own and related fields
  - 1.72 Can construct a simple diagnostic test
  - 1.73 Can administer and interpret a diagnostic test
  - 1.74 Can use other appropriate diagnostic procedures
- 1.8 Utilizes adequate evaluation procedures
  - 1.81 Informal evaluation procedures (anecdotal record, interview, questionnaire)
    - 1.811 Can collect needed information
    - 1.812 Can interpret data collected
  - 1.82 Standard achievement tests
    - 1.821 Is familiar with the more common ones in her field
    - 1.822 Is able to select the ones appropriate for the occasion
    - 1.823 Can administer them
    - 1.824 Can interpret their results
    - 1.825 Can use test results effectively
  - 1.83 Teacher made tests
    - 1.831 Is skillful in constructing appropriate tests
    - 1.832 Can appraise the test as to its principal characteristics
    - 1.833 Can interpret test results
    - 1.834 Can utilize test results
  - 1.84 Accurate and adequate records
    - 1.841 Case studies
    - 1.842 Cumulative records
  - 1.85 Grading and reporting

## 2. A counselor and guidance worker

- 2.1 Adapts principles of child development and mental hygiene to individual and group guidance
  - 2.11 Knows each pupil as an individual
  - 2.12 Recognizes range of emotional maladjustments
  - 2.13 Co-operates with specialists in remedial programs
  - 2.14 Provides experiences through which pupils gain insight into vocational and avocational needs
  - 2.15 Provides opportunity for success experiences for all pupils
- 2.2 Is competent to collect and utilize significant counseling data
  - 2.21 Can administer aptitude interest and intelligence tests
  - 2.22 Can interpret results of such tests
  - 2.23 Keeps records suitable for personal guidance
  - 2.24 Maintains effective relationship with homes

- 2.3 Can utilize suitable counseling techniques
  - 2.31 Is familiar with individual and group techniques
  - 2.32 Differentiates between directive and non-directive counseling
  - 2.33 Employs adequate informational and diagnostic procedures
  - 2.34 Recognizes own abilities in counseling
    - 2.341 Understands symptoms of maladjustment
    - 2.342 Refers severe cases to specialist
  - 2.35 Understands basic principles of effective counseling

### 3. A mediator of the culture

- 3.1 Can direct individuals and groups to significant life applications of classroom learning
  - 3.11 Can utilize his field of subject matter to develop an understanding of social and economic problems
  - 3.12 Can develop an understanding of the wide significance of his own field of subject-matter
- 3.2 Draws on an experiential background to enrich cultural growth of groups and individuals
- 3.3 Develops an understanding and appreciation of current social problems
  - 3.31 Is able to identify key problems, describe their inter-relationships and define the issues
  - 3.32 Develops in pupils a desire to find democratic solutions to current social problems
- 3.4 Can direct pupils in learning to use those materials from which they will continue to learn after leaving school
  - 3.41 Knows source of current information
  - 3.42 Can teach effective procedures for using current materials as sources of information
  - 3.43 Can interest pupils in using current materials
  - 3.44 Can plan learning materials which necessitate the use of current materials
  - 3.45 Can evaluate the progress of the current materials program
- 3.5 Can develop pupil attitudes necessary for democratic participation in society
  - 3.51 Through pre-planning of teaching unit
    - 3.511 Knows an effective unit structure for the problem approach
    - 3.512 Is familiar with many types of classroom activities
    - 3.513 Has a resource of related projects suitable to individual differences so that the problem becomes personalized
    - 3.514 Can list outcomes useful as the basis for teacher evaluation
  - 3.52 Through teacher-pupil planning
    - 3.521 Develops the problem in terms of pupil needs and interests
    - 3.522 Plans and directs research so that each pupil may make a contribution to the group
    - 3.523 Plans a culminating unit that will necessitate full use of available resources and talents
    - 3.524 Provides for evaluation that will effectively indicate progress
- 3.6 Can develop pupil skills necessary for effective participation in a democratic society
  - 3.61 By developing effective discussion practices
    - 3.611 Knows discussion techniques
      - 3.6111 Types of discussion
      - 3.6112 Preparation for discussion
      - 3.6113 Effective discussion
      - 3.6114 Evaluation of discussion
    - 3.612 Has the skills for carrying on discussion
      - 3.6121 Stimulating discussion
      - 3.6122 Following through steps in thinking
      - 3.6123 Conducting self-evaluation
      - 3.6124 Securing effective participation
    - 3.613 Has the ability to develop these skills in the pupils
      - 3.6131 Recognition of the steps in thinking
      - 3.6132 Effective participation procedures
      - 3.6133 Effective self-evaluation
    - 3.614 Can develop pupil discussion leaders
      - 3.6141 Knows the role of leader
      - 3.6142 Understands responsibilities of each at different steps in the discussion

- 3.6143 Aids pupils to appreciate these factors
- 3.6144 Stimulates pupil growth in role
- 3.6145 Directs self-evaluation for role
- 3.62 By developing intelligent committee participation
  - 3.621 Uses the committee as a learning experience
    - 3.6211 Develops techniques of committee organization and participation
    - 3.6212 Aids pupils to progress from individual to group responsibilities
    - 3.6213 Recognizes and meets the needs of individual pupils
    - 3.6214 Prevents exploitation of members
    - 3.6215 Stimulates, holds, directs pupil interest
    - 3.6216 Evaluates effectiveness of organization and participation
  - 3.622 Uses the committee as a contribution to classroom activities
    - 3.6221 Pre-plans in terms of pupil needs and available resources
    - 3.6222 Correlates each job to the total program
    - 3.6223 Provides for individual and group contributions
- 3.63 By developing effective community participation
  - 3.631 Knows the community, its resources, and key people
  - 3.632 Knows the broad implications of subject matter and how it ties into the community
  - 3.633 Has a sensitivity for planning, for follow-up, for adapting jobs to the individual
  - 3.634 Is sensitive to those community resources which are for purposes of information only and to those which are fields for active pupil co-operation
  - 3.635 Can evaluate the effectiveness and the quality of participation and of pupil growth
  - 3.636 Is sensitive to possible exploitation of pupils by community agencies.

#### 4. A member of the school community

- 4.1 Is able to plan co-operatively on education and administrative objectives
  - 4.11 Is competent in curricular planning
  - 4.12 Is competent in evaluation projects
    - 4.121 Definition of objectives
    - 4.122 Collection of data
    - 4.123 Making diagnosis
    - 4.124 Improving the program
- 4.2 Reveals sense of responsibility for his share in the overall effectiveness of the school
  - 4.21 Shares willingly in administrative responsibility
  - 4.22 Participates in planning and administering extra-curricular activities
  - 4.23 Is willing to start with schools as they are and work for improvement

#### 5. A liaison between school and community

- 5.1 Utilizes available education resources of the community in classroom procedures
- 5.2 Secures co-operation of parents in school activities
- 5.3 Interprets the school to the community
- 5.4 Demonstrates ability to assist lay groups in developing understanding of modern education
- 5.5 Participates in definition and solution of community problems
  - 5.51 Draws on available and appropriate school resources
  - 5.52 Assists in defining and developing awareness in pupils and parents

#### 6. A member of the profession

- 6.1 Demonstrates an appreciation of the social importance of the profession
  - 6.11 To parents
  - 6.12 To pupils

- 6.13 To other members of the profession
- 6.14 To the community at large
- 6.2 Develops and adheres to a professional code of ethics
- 6.3 Contributes to the profession through its organizations
  - 6.31 Belongs to the professional organization
  - 6.32 Works effectively in the activities of the organization
    - 6.321 Works within the organization through democratic processes
      - 6.3211 Demonstrates skill in group processes
      - 6.3212 Develops co-operation among workers in the group
      - 6.3213 Forms policies in co-operation with other workers
      - 6.3214 Meets individual responsibilities for accomplishing goals of the organization
    - 6.322 Secures public co-operation and understanding in audience situations
      - 6.3221 Speaks effectively to formal and informal audiences
      - 6.3222 Explains topics in his special field of knowledge
      - 6.3223 Presents controversial topics effectively so as to increase understanding
      - 6.3224 Participates effectively as a member of a panel or discussion group
    - 6.323 Maintains working relationship with lay groups and individuals
      - 6.3231 Finds appropriate occasions to discuss educational issues with lay individuals
        - 6.32311 With lay associations interested in the study of school problems
        - 6.32312 With lay associations in general
      - 6.3232 Maintains contacts based on understanding of power structure in the community
    - 6.324 Recognizes and identifies sources of community concern on school problems
      - 6.3241 Enters into other-than-school problems
      - 6.3242 Participates in other-than-school meetings
      - 6.3243 Participates with lay organizations in analyzing and interpreting information about schools
    - 6.325 Contributes to the development of an effective organization
      - 6.3251 Contributes to activities designed to strengthen the organization for meeting its responsibilities
      - 6.3252 Communicates effectively across subject-matter and grade-group lines
      - 6.3253 Makes special talents available to the services of the organization
        - 6.32531 Writes articles for organizational purposes
        - 6.32532 Participates in arranging for formal meetings and convocations
        - 6.32533 Edits and reviews materials for organizational publications
        - 6.32534 Participates in plans of group projects and researches
        - 6.32535 Participates in organizational research program
        - 6.32536 Serves effectively as moderator, chairman, or in other leadership roles.

October 18, 1954.

GG

PROPOSED INDIVIDUAL FILING CARD  
FOR TEACHER-TRAINING STUDY

(see next page for reverse side)

TEACHER - TRAINING STUDY  
FACULTY OF EDUCATION  
UNIVERSITY OF MANITOBA

\_\_\_\_\_  
(Code)

\_\_\_\_\_  
(Number)

Name \_\_\_\_\_  
in Full (Surname) (Given Names)

Present School \_\_\_\_\_  
(Name) (Address)

\_\_\_\_\_  
(Type of School) (Grade(s) Taught) (Major Subjects)

\_\_\_\_\_  
(Age-Yrs.) (Sex: M or F) (Marital Status) (Year of Ed. 1) (Yrs. T. Exp.)

\_\_\_\_\_  
(Racial Origin) (Father's Occupation) (Religion)

\_\_\_\_\_  
H. S. Attended) H. S. Awards, Ho nors, Scholarships)

-----  
\_\_\_\_\_  
(University Ed. - Degrees etc.) (Major Subjects)

\_\_\_\_\_  
(Previous T. Exp. - School, type, grades, major subjects.)

\_\_\_\_\_  
(Extra - curricular proficiency - drama, oratory, music, etc.)

\_\_\_\_\_  
(General Appearance - dress, vitality, etc.)

\_\_\_\_\_  
(Estimate of Suitability for Teaching - from Interview)

GENERAL REMARKS:





Faculty of Education,  
University of Manitoba,  
Fort Garry, Manitoba,  
November 15, 1955.

.....,  
.....,  
.....

Dear Sir or Madam:

For my Master's thesis, upon the suggestion and under the sponsorship of the Faculty of Education of the University of Manitoba, I have undertaken a study in the "Prediction of Teaching Efficiency". The purpose is that of determining what predictive values can be attached to certain tested and measured variables. Our subjects are the teachers who were enrolled in Education I at this university during the 1953-54 term, most of whom are now in their second year of teaching. The teaching efficiency of these subjects is to be evaluated by competent observers. We hope that you will consent to act in that capacity.

Unfortunately, the size of our sample has been reduced so considerably during the last two years that it is imperative that we obtain a rating on all those who are still teaching in the province this year. Will you please assist us? Also, to add to the validity of our techniques, it is important that the performance of the subjects be evaluated at approximately the same time. Will you please have yours done by the date stipulated below.

We are asking you for an evaluation of the teaching performance of our teacher-subjects in your division or school. The name of the teacher appears at the top of the attached form. In order to conceal the identity of the teacher and the person rating, the responses, upon their receipt here in the office, will be coded, and this top portion of the form will be destroyed. No names will be used thereafter in our procedures, and no teacher or rater will suffer in any way from this assessment.

To complete the attached form, please fill in the information required at the top of the form, and indicate the efficiency category for the teacher rated by drawing an "X" across the frame (box) opposite one of the categories of your choice. Below the category scale you will find a list of qualities in terms of which teachers may be categorized. This list is tendered as a possible guide only; it represents a consensus of opinion of investigators who have conducted research work in this field. Whether or not you use this guide is left to your discretion.

Please have the completed form in the mail by December 15, next. A stamped addressed envelope is enclosed for your convenience. From the faculty and myself, may I thank you for your kind assistance and cooperation.

Yours respectfully,

J. M. Kochan,  
Research Assistant.

THE UNIVERSITY OF MANITOBA

DEAN OF EDUCATION

WINNIPEG, CANADA

January 16th, 1956.

Dear

You will probably know that it is our custom to try and keep in touch with our former students by a follow-up of their work in schools. We find that it is sometimes very profitable for us to visit our former students in their teaching posts in order that we may find out by direct contact how they are doing and how we may improve our own offerings here in light of their subsequent experience. You have now been out of the Faculty long enough to have formed definite conclusions about our offerings and we would, therefore, like to visit you sometime in the near future.

I will be glad to know that you are willing to cooperate with us in this way. As soon as we hear your agreement we will make arrangements to come out and see you in the real practical situation of teaching.

Kindest regards.

Yours sincerely,

N. V. Scarfe,  
Dean.

NVS/sc