



1960

## An experimental study for the selection of a test battery for use in the screening interview

Richard Manning Hawes  
*University of the Pacific*

Follow this and additional works at: [https://scholarlycommons.pacific.edu/uop\\_etds](https://scholarlycommons.pacific.edu/uop_etds)



Part of the [Psychology Commons](#)

---

### Recommended Citation

Hawes, Richard Manning. (1960). *An experimental study for the selection of a test battery for use in the screening interview*. University of the Pacific, Thesis. [https://scholarlycommons.pacific.edu/uop\\_etds/1457](https://scholarlycommons.pacific.edu/uop_etds/1457)

This Thesis is brought to you for free and open access by the Graduate School at Scholarly Commons. It has been accepted for inclusion in University of the Pacific Theses and Dissertations by an authorized administrator of Scholarly Commons. For more information, please contact [mgibney@pacific.edu](mailto:mgibney@pacific.edu).

AN EXPERIMENTAL STUDY FOR THE SELECTION OF A TEST  
BATTERY FOR USE IN THE SCREENING INTERVIEW

---

A Thesis  
Presented to  
the Faculty of the Department of Psychology  
College of Pacific

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

---

by  
Richard Manning Hawes  
June 1960

## TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM AND DEFINITIONS OF TERMS USED . . . .	1
The Problem . . . . .	1
Statement of the problem . . . . .	1
Need for such an investigation . . . . .	1
Statement of the hypothesis . . . . .	3
Brief statement of the experimental procedure . . . . .	3
Definitions of Terms Used . . . . .	4
STEP . . . . .	4
SCAT . . . . .	4
MVI . . . . .	4
LAT . . . . .	5
MTAI . . . . .	5
AI . . . . .	5
SSHA . . . . .	5
CTA . . . . .	5
ACE . . . . .	6
II. REVIEW OF THE LITERATURE . . . . .	7
Literature on the Purposes and Criteria in Student Teacher Selection and Guidance . . .	7
Literature on Experiments in the Field . . . .	9
Early experiments . . . . .	9

	111
CHAPTER	PAGE
Recent experiments and trends . . . . .	12
Summary . . . . .	19
III. ORGANIZATION OF THE INVESTIGATION . . . . .	20
Definition of the Sample . . . . .	20
Description of the Instruments Used . . . . .	20
Sequential Test of Educational Progress . . . . .	21
School and College Abilities Test . . . . .	21
Michigan Vocabulary Profile Test . . . . .	22
Language Abilities Test . . . . .	22
Minnesota Teacher Attitude Inventory . . . . .	23
Adjustment Inventory . . . . .	23
Survey of Study Habits and Attitudes . . . . .	24
Critical Thinking Appraisal . . . . .	25
Methods of Procedure . . . . .	25
IV. THE FINDINGS . . . . .	28
Experimental Results . . . . .	28
Summary . . . . .	31
V. FINAL INTERPRETATION AND CONCLUSIONS . . . . .	33
Summary . . . . .	33
Conclusions . . . . .	34
Values of the selected battery . . . . .	35
BIBLIOGRAPHY . . . . .	38
APPENDIX . . . . .	43

## LIST OF TABLES

TABLE	PAGE
I. The Correlation Coefficients of Each Test With Every Other Test Included in the Study and Their Significance at the .01 and .05 Levels of Confidence . . . . .	29
II. The Number of Times Each Test was Significant at the .05 and .01 Levels of Confidence . . . .	32

## CHAPTER I

### INTRODUCTION

This study is based on the assumption that there is a need for identifying, selecting, counseling, and developing teacher competencies and the recognition of the need for further studies in defining the good teacher. Such studies should contribute to more efficient methods of selecting and preparing teachers. The problem of selecting persons who are apt to make good teachers, and who have a reasonably high potential is of primary importance. By statistically selecting a brief battery of tests from a larger battery, this study hopes to present the counselors at the College of the Pacific with a practical and usable battery of tests which will aid in their selection, counseling, and guidance of prospective teachers.

### THE PROBLEM AND DEFINITIONS OF TERMS

#### I. THE PROBLEM

Statement of the problem. The problem of this study is as follows: Is it possible to select from a large battery of tests a minimum selection of tests which will offer relatively the same information offered by the original battery?

Need for such an investigation. Presently and for

years to come there exists a need for thousands of effective teachers. Poor teachers and their attitudes induce an unpleasant atmosphere within the profession and often discourage capable men and women who would otherwise be attracted to teaching. These poor influences cause the teaching profession severe hardships for improving working conditions. More and better teachers will be developed by selective recruitment based on personality factors as well as on scholastic performance.

It is impossible for the counselor, by subjective means alone, to screen and evaluate effectively all the teacher candidates. Therefore, the School of Education at the College of the Pacific selected thirteen tests considered potentially useful in screening candidates for teaching. As the test results were used in the counseling situation, it was felt that the many test results duplicated information, caused confusion in the counseling sessions, and were not economical relative to time or expense. If it could be shown statistically that several of the tests provided duplicate information, a minimum battery of tests could be selected offering relatively the same information, reducing the time element, and avoiding the confusion caused by the multiplicity of scores resulting from the original battery.

In the absence of an available criterion of later success in student teaching, on the part of the college

students involved in this study, no attempt at an actual prediction of such success is to be inferred from any or all of the measures used. The sole objective of this study is to aid in the selection of a limited battery of tests by which the screening committees will be in possession of data to be used by the committee members in alerting them to characteristics which, in the experience of such committee members may be significant for the purpose of the committees. A very extensive mass of test data can be more confusing than helpful. Hence it is necessary to eliminate those instruments which can be shown by the correlation technique, to overlap in identification of student characteristics.

Statement of the hypothesis. The null hypothesis of this study was that the thirteen tests of the original battery are sufficiently different in what they measure to justify the use of the total battery.

A brief statement of the experimental procedure. Each student in a beginning class in education was administered several tests considered potentially useful in the selection of teacher candidates. The tests were selected by the School of Education at the College of the Pacific to offer the counselor information in two general areas:

- (1) Academic aptitudes and
- (2) Attitudes, habits, and



adjustment. Each of the thirteen tests in the original battery was paired with every other test to determine their correlation coefficients. If a significant relationship was found between two tests, it was inferred that these two tests probably measured relatively the same factors; therefore, one of the tests could be eliminated. If the relationship between the paired tests measured somewhat different factors and both tests would be retained in the final battery.

## II. DEFINITIONS OF TERMS USED

The following are the definitions of terms which will be used frequently in this study:

STEP: This will be used to identify the Sequential Tests of Educational Progress developed by the Cooperative Test Division, Educational Testing Service, Princeton, New Jersey and Los Angeles, California.

SCAT: This is used to identify the School and College Ability Tests developed by the Cooperative Test Division, Educational Testing Service, Princeton, New Jersey and Los Angeles, California.

MVI: This will be used to identify the Michigan Vocabulary Profile Test which was prepared under the direction of Edward B. Green from the University of Michigan,

Ann Arbor, Michigan, and published by the World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois.

LAT: This will be used to identify the Language Abilities Test developed by Harry A. Green, University of Iowa, and Helen I. Stapp, Decatur High School, Decatur, Illinois, and published by the World Book Company, Yonkers-on-Hudson, New York, and Chicago, Illinois.

MTAI: This will identify the Minnesota Teacher Attitude Inventory which was developed by Walter Cook from the University of Minnesota, Carroll H. Leeds from Furman University, and Robert Callis from the University of Missouri for the Psychological Corporation, New York, New York.

AI: This will identify the Adjustment Inventory constructed by Hugh M. Bell, Chico State College, California, and published by the Stanford University Press, Stanford, California.

SSHA: This is used to identify the Survey of Study Habits and Attitudes developed by William F. Brown and Wayne H. Holtzman for the Psychological Corporation, New York, New York.

CTA: This will identify the Watson - Glaser Critical

Thinking Appraisal published by the World Book Company,  
Yonkers-on-Hudson, New York, and Chicago, Illinois.

ACE: This is used to identify the American Council  
on Education Psychological Examination for College Freshmen  
as edited by L. L. Thurston and Thelma Gwinn and published  
by the Cooperative Test Division, Educational Testing  
Service, Princeton, New Jersey and Los Angeles, California.

## CHAPTER II

### REVIEW OF THE LITERATURE

#### I. LITERATURE ON THE PURPOSES AND CRITERIA IN STUDENT TEACHER SELECTION AND GUIDANCE

The unreliability of subjective observation by itself is well accepted by educators today. Many universities are using various instruments as part of their entrance, selection, and guidance programs<sup>1</sup>. Subjective evaluation alone results in admitting hundreds of teacher candidates each year who will have discouraging and unsuccessful teaching experiences in the following years.

One of the primary concerns of most educators is to select those students who will eventually finish school and remain in the teaching profession. Berdie<sup>2</sup> emphasizes this concept as he comments on the purpose of selection in higher education. Although he talks of college students in general his ideas are basic to any further selection of teacher candidates:

---

<sup>1</sup> John E. Sands, "Off Campus Student Teaching Practices in 112 Institutions," Ed., 73:636-44, June, 1953.

<sup>2</sup> Ralph F. Berdie, "Improving Evaluation of Student Recruitment and Selection," The Personnel and Guidance Journal, 34:481, April, 1956.

The purpose of selection in higher education is to select from the applicants the person who in any given school will benefit the most personally and at the same time develop skills, attitudes, interests, and knowledge that will make them of maximum social value.

Recently, research has indicated that many prospective teachers do not graduate and enter teaching. One university reports that only sixty per cent of the freshman class of students in general enter the sophomore year<sup>3</sup>. Other reports indicate that up to sixty per cent of students fail or become drop outs before graduation<sup>4</sup>. This is typical of many similar reports.

Although numerous studies have been developed in an attempt to identify factors in preservice performance which will provide a reliable basis for predicting the efficiency or ability of prospective teachers in student teaching, all the reports indicate that it is difficult to select a complete set of criteria which will adequately describe the successful teacher. Several investigators have continued to seek single factors which will relate to subsequent

---

<sup>3</sup> Leonard O. Andrews, "Admissions and Selective Retention in a University College of Education," Teacher Education: The Decade Ahead, De Kalb Conference Report, Washington, D. C.; N. E. A., pp. 252-258. 1955.

<sup>4</sup> Joe Alton Apple and J. R. Shunert, "Toward a Better Secondary Teacher Education Program," Ed. Adm. and Sup. 41:129-144, March, 1955.

performance, and many studies have produced inconclusive and conflicting findings. However, as variables are combined for prognostic purposes, more accurate and consistent predictions seem to have been achieved. Because adequate and reliable criteria to describe the successful teacher continue to be a consistent problem in education Bretsch and Jacobsen<sup>5</sup> suggest:

. . . Since the identification of teaching success is such a complex matter, it may be easier to rule out candidates on the basis of characteristics which are rather clearly causes of failure, than to select candidates who are likely to be notably successful.

## II. LITERATURE ON EXPERIMENTS IN THE FIELD

Early Experiments. It has been mentioned that one of the most difficult problems related to teacher selection and guidance is choosing criteria which will describe successful teaching. Another equally difficult problem in this field is developing instruments which will be sensitive to these criteria. The following are representative experiments, primarily concerned with student teacher evaluation, which have been reported.

---

<sup>5</sup> Howard S. Bretsch and Gene S. Jacobsen, "Recruitment, Guidance, and Screening of Prospective Elementary and Secondary School Teachers," R. Ed. Res. 25:204-212, June, 1958.

As early as 1916 one of the first attempts was made to predict success in student teaching on the basis of preservice performance of the subjects. Mead and Holly<sup>6</sup> used 39 student teachers in a study which found student teaching success as indicated by a supervisor's rating scale correlated with general scholarship .24, scholarship in the major field of study .19, and a general methods course mark .57. In conclusion, they recommended combining the three measures for predicting student teacher performance.

Since this time many experiments have been concerned with the predictive value of tests in student teaching. In 1948 Barr<sup>7</sup> published a summary of 153 studies concerned with the measurement and prediction of teaching efficiency. Later, Domas and Tideman<sup>8</sup> reported a bibliography of 1,006 items dealing with teacher evaluation. Only a few outstanding studies are here reported.

---

<sup>6</sup> A. R. Mead and C. E. Holly, "Forecasting Success in Practice Teaching," J. of Ed. Psych. 7:495-97, October, 1916.

<sup>7</sup> Arvil Sylvester Barr, "The Measurement and Prediction of Teaching Efficiency: A Summary of Investigation," J. of Exp. Educ. 26:203-283, June, 1948.

<sup>8</sup> Simeon J. Domas and David Tideman, "Teacher Competence: An Annotated Bibliography," J. of Exp. Ed. 29:101-218, December, 1950.

In a study of ninety student teachers, Dodd<sup>9</sup> used ratings by a supervisor as a criterion of teaching ability. He found that the criterion correlated .35 with scholarship and .39 with weighted aggregate scores on the Cox-Orleans Prognostic Test for Teaching Ability. At the University of Minnesota R. K. Bent,<sup>10</sup> using 1,084 prospective teachers, found that the highest correlation of student teaching performance and the factors studied were: hours of credit, .46; scholarship, .46; scholarship in the major, .45; and credits in education, .30.

Walter W. Cook<sup>11</sup> reported a correlation of .60 between scores on an educational information and application test, and graduate course work. Three hundred and twenty-one students were used in a study by Glenn W. Durflinger<sup>12</sup>. The students were administered tests of intelligence, achievement, english, music, personality, and personal

---

<sup>9</sup> M. R. Dodd, "A Study of Teaching Aptitude," J. of Ed. Res. 26:517-21, March, 1933.

<sup>10</sup> Rudyard K. Bent, "Relationships Between Qualified Examinations, Various Other Factors, and Student Teacher Performance at the University of Minnesota," J. of Exp. Ed. 5:251-55, March, 1937.

<sup>11</sup> Walter W. Cook, "Predicting Success of College Graduate Students in Teacher Education," School and Society, 46:192-95, September, 1942.

<sup>12</sup> Glenn W. Durflinger, "Scholastic Prediction in a Teachers College," J. of Exp. Ed. 11:257-67, June, 1943.



fitness. He found a multiple correlation of .543 for men and .562 for women between the first semester grade point average and the combined test scores.

During this period two studies were reported from England. In using two groups of forty-five students, Lawton<sup>13</sup> found correlation coefficients of .70 for one and .82 for the other between a combination of pre-college examination and admission interviews, and the general in service teaching marks. Later, Tudhope<sup>14</sup> used the "final teaching mark" as a predictor of teaching success. The "final teaching mark" was a joint assessment of the student's teaching made by four observers. A correlation of .81 was found between this mark and the rating made by a royal inspector after the student had been teaching for three years.

Recent experiments and trends. Evaluating the collected information concerning student teacher candidates

---

<sup>13</sup> J. A. Lawton, "A Study of Factors Useful in Choosing Candidates for the Teaching Profession," British J. of Ed. Psych. 9:131-44, June, 1939.

<sup>14</sup> William B. Tudhope, "A Study of the Training College Final Teaching Mark as a Criterion of Future Success in the Teaching Profession, Part I," British J. of Ed. Psych. 12:167-71, November, 1942.

is becoming more important as the years pass. Breuckner<sup>15</sup> suggests how the collected information should be used as he describes educational diagnosis:

Educational Diagnosis relates to the techniques by which one discovers and evaluates both strengths and weaknesses of the individual as a basis for more effective guidance. Diagnosis is a logical process based on a consideration of all the available data concerning a particular individual or group of individuals. The analysis of these data and their interpretation in the light of knowledge gained from past experiences enables the diagnostician to suggest necessary developmental or remedial measures.

The question is usually asked, "Can the information obtained from the use of tests actually identify those students who are apt to succeed or fail in teaching?". Some reports suggest that it is doubtful<sup>16</sup>. However, recent studies are more optimistic.

According to Ohlsen and Shultz<sup>17</sup>, sociometric techniques show some degree of efficiency in identifying potentially good and poor student teachers; and they suggest

---

<sup>15</sup> Leo J. Breuckner, "The Need of Educational Diagnosis," 34th Yearbook of the National Society for the Study of Education, (Bloomington, Ill.: Public School Publishing Co.), 1935.

<sup>16</sup> Arvil Sylvester Barr, "The Measurement and Prediction of Teaching Efficiency," J. of Exp. Ed. 3:193-196, September, 1952.

<sup>17</sup> Merle Marvel Ohlsen and R. E. Shultz, "Study of Variables for Use in Selection," J. of Teacher Ed. 5:279-282, December, 1954.

further study with the Thematic Aperception Test and possibly other informal projective procedures.

Downie and Bell<sup>18</sup> reported a significant relationship between over all grade point average and scores on the Minnesota Teacher Attitude Inventory. The authors subjectively judged that students who scored higher on the M.T.A.I. usually had a broader background of experiences with youngsters and expressed a greater interest in teaching. Their instructors also tended to judge them as better teaching prospects.

Inlow<sup>19</sup> reported a study in which he had the prospective student teachers interviewed by two staff members. The applicants were then listed by the staff members, on the basis of observed personality factors, according to anticipated success. When the student teaching assignments were finished a twenty-five item evaluative instrument was completed by the master teachers. This instrument was used as the criterion for successful student teaching. Correlations from .12 to .83 were found between the listings of anticipated

---

<sup>18</sup> Norville M. Downie and C. R. Bell, "Minnesota Teacher Attitude Inventory as an Aid in the Selection of Teachers," J. of Ed. Res. 46:699-704, May, 1953.

<sup>19</sup> Gail M. Inlow, "Evaluating Student Teachers Experiences," J. Ed. Res. 45:705-14, May, 1952.

success and of student teacher success.

In another study Leavitt<sup>20</sup> reported that students who scored above the sixtieth percentile on the American Council of Educational Psychology Examination were more successful in student teaching than those who scored below. Leavitt also noted that neither the Ohio State University Psychological Test nor the Northwestern University Analogies Test were able to predict between those who were successful in student teaching and those who were not.

Some studies have been concerned with the use of certain tests or the way of presenting results in counseling students.

The interview as a screening device was studied by Shaw<sup>21</sup> as he attempted to find a relationship between successful student teaching and objective data available at admission time to college. He found no significant relationship. However, when he evaluated the data supplemented with an interview, he found a significant relationship between successful student teaching and his predictions.

---

<sup>20</sup> Jerome E. Leavitt, "Predicting Teaching Success," Phi Delta Kappa, 35:281-2, April, 1954; J. Teacher Ed. 4:194-7, May, 1953.

<sup>21</sup> Jack Shaw, "Function of the Interview in Determining Fitness for Teacher Training," J. of Ed. Res. 45:667-81, May, 1952.

The idea of examinee selection of tests and counselee participation in interpretation of the results has been discussed in several articles. There are relatively few research attacks on these issues. Dressel and Matteson<sup>22</sup>, regarding client participation in test interpretation concluded that the main need is for more research in this area. Strange<sup>23</sup> did a study at the preparatory-school level which raised some questions which should be studied further and Bordin<sup>24</sup> set forth basic purposes for client participation: (1) A source of information for the client. (2) Help the counselor better understand the client. (3) Stimulate the client to examine himself from a new point of view. (4) Clarify the nature of the counseling process to the client.

Another general question about how tests are interpreted in counseling concerns the stability of scores or patterns of scores. Since investigation of this type

---

<sup>22</sup> Paul Leroy Dressel and R. W. Matteson, "The Effect of Client Participation in Test Interpretation," Educ. Psych. Meas. 10:693-706, Winter, 1950.

<sup>23</sup> Frank B. Strange, "Student Self Selection of Group Tests," Pers. and Guid. J. 32:30-33, September, 1953.

<sup>24</sup> Edward S. Bordin, "Four Uses for Psychological Tests in Counseling," Educ. Psych. Meas. 11:779-81, Winter, 1951.

usually requires longitudinal studies, experiments in this area are scarce. Some studies have offered conflicting results. Rosenberg<sup>25</sup> studied stability of interests in high school and noted that use of interest patterns in counseling in the ninth grade was not a waste of time. However, Mallinson and Crumerine<sup>26</sup> concluded that students should not be vocationally counseled at the ninth grade level on the assumption that interest patterns will remain through high school.

Although most guidance people do not advocate the use of test results without attention to other evidences, the layman who reads testing research may arrive at that conclusion because of the emphasis on scores. As one counsels students, the necessity for considering other information gained from interviews and anecdotal records should be realized. Sacks and Lewin<sup>27</sup> suggested that serious errors may result when diagnosis is not supplemented by available

---

<sup>25</sup> Nathan Rosenberg, "Stability and Maturation of Kuder Interest Patterns During School," Educ. Psych. Meas. 13:449-58, Autumn, 1953.

<sup>26</sup> George Greisen Mallinson and W. M. Crumerine, "An Investigation of the Stability of Interests of High School Students," J. of Ed. Res. 45:369-83, January, 1952.

<sup>27</sup> Joseph M. Sacks and Herbert S. Lewin, "Limitations of the Rorschach as Sole Diagnostic Instruments," J. of Couns. Psych. 14:479-81, December, 1950.

clinical approaches.

The use of local validity studies should make prediction results more useful to the counselor<sup>28</sup>. The need for "localization" of prediction was brought out in two studies at the college level. Frederiksen and Shrader<sup>29</sup> reported the results of a study of twelve colleges in which they used the A. C. E. tests and high school standing as predictors of college success. Once more the range of coefficients from school to school, or from group to group within schools was notably large. In another study Wallace<sup>30</sup> used several tests in a freshman testing program to predict grades in specific college courses. Some of his conclusions apply not only to his studies but also as rather general conclusions applicable to most prediction studies. He noted that all the correlations are so low that considerable caution must be used against placing too much weight on an individual's test scores in selection, guidance, or counseling.

---

<sup>28</sup> Norman Oliver Frederiksen, "Making Test Scores More Useful for Prediction," Educ. Psych. Meas. 11:783-87, Winter, 1951.

<sup>29</sup> Norman Oliver Frederiksen and W. B. Shrader, "The A. C. E. Psychological Examination in High School Standing of Predictors of College Success," J. of Ap. Psych. 36:261-65, August, 1952.

<sup>30</sup> Wimburn Leroy Wallace, "The Prediction of Grades in Specific College Courses," J. of Educ. Res. 44:587-97, April, 1951.

He also urged that each type of institution determine its validities locally so that test scores may have meaning as applied to its curriculum and students.

Summary. Research concerned with teaching prediction and evaluation has produced a large quantity of varied data. Many of the results are conflicting and inconclusive. Since there presently is no adequate definition of teaching efficiency, each research project should be considered relative to the validity of the criterion used.

Some significant trends are apparent in the research. There is an increasing respect for the complexity of the teaching function and a willingness to consider all of the variables presented by a teaching situation. There seems to be serious and consistent attempts to develop more precisely defined criteria of teaching ability. There is a trend to improve the functional predictive value and quality of tests and a realization that test results supplemented by various counseling techniques and pertinent information can be more effective than test scores by themselves.



## CHAPTER III

### ORGANIZATION OF THE INVESTIGATION

This chapter defines the sample and its selection. It describes the instruments used in the investigation and discusses the procedures of the investigation.

#### I. DEFINITION OF THE SAMPLE

The subjects used in this investigation were thirty-five sophomore and junior college men and women enrolled in a beginning education class at the College of the Pacific in the spring of 1958. The group was comprised of those men and women who successfully fulfilled the following conditions:

1. Took the tests during their course in Education 11.

2. Completed all the tests in the battery given.

In the beginning, forty-nine students, which was the entire class, participated in the testing program. The number of students who fulfilled the previously numbered conditions was reduced to thirty-five.

#### II. DESCRIPTION OF INSTRUMENTS USED

All the instruments included in this study were part of a testing program used in an introductory education

course. The tests were selected by the School of Education of the College of the Pacific, to cover various academic, attitudinal, and adjustment elements.

Sequential Tests of Educational Progress. The STEP was developed by the Cooperative Test Division, Educational Testing Service and copyrighted in 1957. The STEP is a series of achievement tests which measure applied skills in seven fields of school and college instruction.

Four levels of tests are offered in each of the fields. The four levels extend from the fourth grade to the sophomore year of college.

In this experiment three of the STEP tests were used at the Freshman and Sophomore level: (1) Mathematics, (2) Science, (3) Social Studies.

School and College Abilities Test. The SCAT was developed by the Education Testing Service to measure two kinds of school related abilities: Verbal and Quantitative.

In the SCAT series each booklet contains four subtests. Parts I and III measure developed verbal ability; Parts II and IV measure developed ability in basic quantitative areas. Since all the test items are multiple choice, the student chooses the best answer from among five presented choices.

Three scores are obtained: Verbal, Quantitative and

total.

The SCAT series covers five levels of difficulty from elementary grade 4 to college freshman and sophomore years with two forms at each level. The freshman and sophomore level were used in this study.

Michigan Vocabulary Profile Test. The MVI was published by the World Book Company first in 1939 and revised in 1949 and is designed to test an individual's vocabulary in eight different fields. There is no time limit. Only one subtest, "Human Relations", was used from this battery.

Language Abilities Test. The LAT was designed by Green and Stapp to provide a comprehensive measure of the proficiency of high school student and college freshmen in the use of the English language. The test is divided into five parts. Two of these parts, Sentence Structure and Usage, were used in this investigation. These two particular sections also measure the understanding of the grammatical principles underlying correct practices in these two areas.

The following are descriptions of the two tests which were used in this study:

Test 3: Sentence Structure and applied Grammar consists of thirty-eight items, each containing a sentence which may be correct or incorrect. The incorrect sentences

may be incomplete statements or they may contain double negative, unnecessary words, misplaced modifiers, the wrong verb form, or other errors. The student is asked whether or not a sentence is correct; if it is incorrect, he is to choose the statement which tells what should be done to improve it. The time limit for this test is 20 minutes.

Test 5: Usage. The student is asked to tell whether the word underlined in each of fifty sentences is used correctly. If the word is incorrect, he is then to choose the statement which tells why it is wrong. The time limit for this test is 25 minutes.

Minnesota Teacher Attitude Inventory. The MTAI was constructed by Cook, Leeds, and Callis in an effort to measure attitudes of a teacher which predict how he will get along with pupils in interpersonal relationships. It also indirectly predicts how satisfied he may be with the teaching profession.

Form A of the MTAI, consists of 150 attitude statements which are not answered "right" or "wrong", but rather by agreement or disagreement. The test is practically self administering with no time limit, but it usually requires 20 to 30 minutes to complete.

Adjustment Inventory. The AI, designed by Bell of Chico State College, attempts to provide four separate

measures of personal and social adjustment: (1) Home adjustment, (2) Health adjustment, (3) Social adjustment, and (4) Emotional adjustment. The four types of adjustment help the guidance person to more easily locate areas of poor adjustment. The general adjustment status may be indicated by the total score.

The AI is a self administering instrument which has been successfully used with both males and females of high school and college ages. There is no time limit, although most students can complete the test within 25 minutes.

Survey of Study Habits and Attitudes. The SSHA was composed by Brown and Holtzman as a seventy-five item, one form, paper and pencil test which has no time limit but usually requires 20 minutes to complete. The subjects rate themselves, on the male or female scoring keys, according to whether the statements are true of them almost always, generally, frequently, sometimes, or rarely.

The SSHA was developed for the following purposes:

1. To identify those students whose habits and attitudes are different from those of students who earn higher grades.
2. To aid in understanding students with academic difficulties.
3. To provide a basis for helping students to

improve their study habits and attitudes and thus more fully realize their best potential.

Critical Thinking Appraisal: The CTA was designed by Watson and Glaser in an attempt to measure some of the major factors involved in critical thinking and as an aid in developing that ability.

The test consists of five subtests: (1) Inference, (2) Recognition of assumption, (3) Deduction, (4) Interpretation, and (5) Evaluation of arguments. The test can be completed in 40 minutes by most college students, although there is no specific time limit.

### III. METHODS OF PROCEDURE

During one of the first class meetings the students were asked if they would be willing to participate in this study by taking the tests which had been selected by the School of Education at the College of the Pacific. They were informed participation in the project was not required and would have no effect on their grades for the course and that all test results would be treated confidentially. All the class members expressed a desire to participate in the project.

The group form of each test was administered to each subject during a class period. Absentees were administered

the tests by appointments at the testing office of the College of the Pacific. Because class and appointment times were limited, some tests required several days to be completed.

After administering and scoring the instruments, the results were recorded as shown on Table III on page 44. Since the MTAI presented some negative results, 30 points were added to the subjects' individual score on this instrument in order to have all positive scores. Following the recording of the raw scores, the statistical procedures ensued.

The Pearson product-moment method of correlation was used in this study. In order to facilitate the statistical process, an electronic computer was employed which necessitated the use of the following formula in obtaining the correlation coefficients. The formula for the correlation coefficient is:

$$r = \frac{\sum XY - \frac{\sum X \cdot \sum Y}{N}}{\sqrt{\left( \sum X^2 - \frac{(\sum X)^2}{N} \right) \left( \sum Y^2 - \frac{(\sum Y)^2}{N} \right)}}$$

in which

$r$  = correlation coefficient.

$X$  = summation of raw scores on one test.

$Y$  = summation of raw scores on the other test.

$n$  = number of cases.

The  $r$ 's were then tested for their significance at the .05 and .01 levels. The results are shown on Table I.



## CHAPTER IV

### THE FINDINGS

#### I. EXPERIMENTAL RESULTS

The first step in the treatment of the test results was to determine the relationship between each test to every other test. The relationship was determined by the use of the Pearson product-moment correlation method described by Lacey<sup>31</sup>. The results are shown on Table I, page 29.

The second step was to determine the level of confidence of the obtained correlation coefficients by entering Table E<sup>32</sup> which involves the significance of  $r$ . at the .05 and .01 levels. The results are shown on Table I, page 29.

Before analyzing the statistical data each test was considered for administrative, scoring, and interpretive ease; administrative time relative to the class period time of fifty minutes, and the possibility that similar or related instruments were being administered to the students upon entering the College of the Pacific so that such scores could be used.

---

<sup>31</sup> Oliver L. Lacey, Statistical Methods in Experimentation, (New York: The Macmillan Company), 1953, pp. 155-77.

<sup>32</sup> Ibid., p. 245.

TABLE I

THE CORRELATION COEFFICIENTS OF EACH TEST WITH EVERY OTHER TEST INCLUDED IN THE STUDY AND THEIR SIGNIFICANCE AT THE .05 AND .01 LEVELS OF CONFIDENCE

A	B	C	D	E	F	G	H	I	J	K	L	M
A	.51*	.37+	.22	.74*	.29	.39+	.55*	.47*	.45*	-.0014	.081	.50*
B		.39+	.30	.37+	.19	.15	.97*	.38+	.41+	-.097	.067	.55*
C			.41+	.44+	.23	.47*	.49*	.37+	.35+	.038	-.026	.35+
D				.68*	.31	.55*	.56*	.35+	.24	-.077	.24	.41+
E					.37+	.22	.56*	.50*	.46*	-.057	.17	.38+
F						.19	.30+	.22	.26	-.031	.093	.21
G							.46*	.45*	.36+	.13	.15	.37+
H								.71*	.52*	.22	.16	.56*
I									.45*	.17	.19	.47*
J										.025	.091	.43+
K											-.38	-.092
L												.26
M												---

\* Indicates significance at the .01 level.

+ Indicates significance at the .05 level.

#### Key

- A - STEP (MATHEMATICS)
- B - STEP (SCIENCE)
- C - STEP (SOCIAL STUDIES)
- D - SCAT (VERBAL)
- E - SCAT (QUANTITATIVE)
- F - SCAT (TOTAL SCORE)
- G - MVI (HUMAN RELATIONS)
- H - LAT (SENTENCE STRUCTURE)
- I - LAT (WORD USAGE)
- J - MTAI
- K - AI (TOTAL SCORE)
- L - SSHA
- M - CTA

The third step was concerned with the selection of the recommended battery. In this study the correlation coefficient between two instruments was required to be significant at the .01 level of confidence; indicating that a relationship such as this would not happen by chance more than one time in a hundred.

If a significant relationship at the .01 level was found between two tests, it was inferred that these two tests probably measured relatively the same factors.

If no significant relationship was found between two tests, it was inferred that these tests probably measured different factors.

It can be observed from Table II that several relationships between tests are significant at the .01 level for thirty-three degrees of freedom. This seems to reject the null hypothesis of this study, i.e., that the thirteen tests of the original battery are sufficiently different in what they measure to justify the use of the total battery.

As can be seen on Table II, Sentence Structure of the Language Abilities Test had the highest frequency of significance at the .01 level. It was found to be significant at the .01 level with nine other instruments and significant at the .05 level with the other test. On the basis of these results, LAT (H) was selected to represent the area of academic aptitudes, in lieu of STEP (A), STEP (B), STEP (C),

SCAT (D), SCAT (E), MVI (G), LAT (I), MTAI (J), and CTA (M).

The AI (K) and SSHA (L) were found to have no significant relationship at either the .05 or .01 levels with any of the tests with which it was compared. It is also noted that the correlation coefficient between the AI (K) and the SSHA (L) was found to be  $-.38$ . It should be noted that a negative correlation may be fully as predictive as a positive correlation of the same size or value. However, it may also be inferred that a negative correlation will suggest that the instruments involved probably measure different factors.

Because of these findings, the AI (K) and the SSHA (L) were selected for the area of attitudes, habits and adjustment.

Summary. Three instruments were selected as a result of the statistical treatment. Sentence Structure of the Language Abilities Test was selected to measure in the area of scholastic aptitude. Survey of Study Skills and Attitudes and the Adjustment Inventory were selected to measure in the general area of attitudes, habits and adjustment.

TABLE II

THE NUMBER OF TIMES EACH TEST WAS SIGNIFICANT AT THE  
.05 AND .01 LEVELS OF CONFIDENCE

TEST	FREQUENCY OF SIGNIFICANCE AT THE .05 LEVEL	FREQUENCY OF SIGNIFICANCE AT THE .01 LEVEL
STEP (MATHEMATICS)	6	2
STEP (SCIENCE)	3	4
STEP (SOCIAL STUDIES)	7	2
SCAT (VERBAL)	3	3
SCAT (QUANTITATIVE)	4	5
SCAT (TOTAL)	2	0
MVI (HUMAN RELATIONS)	3	4
LAT (SENTENCE STRUCTURE)	1	9
LAT (WORD USAGE)	3	6
MTAI	4	4
AI (TOTAL SCORE)	0	0
SSHA	0	0
CTA	5	4

## CHAPTER V

### FINAL INTERPRETATION AND CONCLUSIONS

In this study the writer has reported his investigation of the null hypothesis that the thirteen tests of the original battery are sufficiently different in what they measure to justify the use of the total battery. If the null hypothesis is rejected it would then be possible to select a minimum battery of tests which would offer relatively the same information, reduce the time element, and avoid confusion caused by the multiplicity of scores resulting from the original battery.

Summary. Students in a beginning course in education were administered several tests selected by the School of Education at the College of the Pacific considered necessary in the selection, counseling, and guidance of prospective teachers.

Each of the thirteen tests were paired with every other test to determine their correlation coefficients. The  $r$ 's were then tested for their significance at the .05 and .01 levels. If a significant relationship at the .01 was found between two tests, it was inferred that these two tests probably measured relatively the same factors. If no significant relationship was found between two tests, it was

inferred that these tests probably measured different factors.

The statistical results indicated that a significantly high relationship existed between nine tests involved in this study with Sentence Structure of the Language Ability Test. The correlation of these tests was significant at the .01 level. The results also indicated that the Adjustment Inventory and the Study Skills and Habits Inventory had no significant relationships with any of the other instruments to which they were compared.

Conclusions. If the investigator follows the statistical inference regarding the null hypothesis, he must conclude that the significant relationship at the .01 level between two tests would not occur by chance more than one time in one hundred. However, he may not conclusively assume that the high significance of relationship between the tests means that the information obtained from one instrument will be specifically the same information as that offered by the other.

It should also be noted that although objective instruments are needed in the selection, counseling, and guidance of prospective teachers, predictive instruments have been the subject of controversy. Each instrument is limited because they usually are designed to measure only

certain areas of teaching ability. Their effectiveness is determined by their validity and use. When selection is based solely on the results of tests, they are obviously misused. However, they may serve a useful function, if they are perceived as offering valid and reliable information about a specific element of teaching ability.

The investigator concludes that several instruments in the original battery measure relatively the same factors and therefore it is possible to select a minimum battery which will present relatively the same information presented by the original battery of thirteen tests.

The final battery includes the following:

1. LAT - Sentence Structure.
2. Adjustment Inventory.
3. Study Skills and Habits Inventory.

Values of the selected battery:

The selected battery obviously requires much less time to administer, and measures relatively the same factors as the original battery. The time needed to complete the selected battery is approximately sixty-five minutes as compared to almost six hours required for the original battery. The results are represented by three scores in the minimum battery while the original set of tests resulted in thirteen scores. The mere decrease in number should



facilitate the individual counseling sessions.

It is becoming increasingly clear that "good teaching" is not only difficult to define but is also complex. This study had hoped to suggest to the counselor a battery of instruments which would give him information in three areas: (1) scholastic aptitude, (2) general adjustment, and (3) study habits and aptitudes. Because of the consistently significant correlations of Sentence Structure with other scholastic aptitude instruments, it was selected to measure this factor in the teaching complex. In this study, SSHA and AI had very little correlation with the other instruments of scholastic aptitude and it was assumed that the predictive power of these two instruments rests on their measurement of traits largely untouched by such measures of mental ability. The traits measured by the two instruments are general adjustment and study skills and attitudes.

It is further indicated that none of these tests are commonly called intelligence or achievement tests. Usually there will be available considerable material information in this area from past records accumulated throughout the students educational experiences. The results of the A.C.E. test, which is administered to incoming freshmen, could be used to supplement the selected battery resulting from this study.

Other factors of personality and interests should be

given considerable attention in the selection, counseling, and guidance of prospective student teachers. Some of these factors can be measured by objective tests, but others cannot. A valid evaluation program should always include the observations of the instructors, health information, general cumulative and anecdotal material, and in some instances evaluations and observations by specialists.

---

## BIBLIOGRAPHY

## BIBLIOGRAPHY

### A. BOOKS

Buros, Oscar K. The Fourth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1953.

Edwards, Allen R. Experimental Design in Psychological Research. New York: Rinehart and Company, Inc., 1936.

Guidance in Teacher Education. Thirty-sixth Yearbook, The Association for Student Teaching, Iowa State Teachers College, Cedar Falls, Iowa. Ann Arbor, Michigan: Edwards Brothers, Inc., 1957.

Lacey, Oliver L. Statistical Methods in Experimentation. New York: The Macmillan Company, 1953. 249ppp.

Teacher Education: The Decade Ahead. De Kalb Conference Report. Washington, D. C.: National Education Association, 1955.

Whitney, Frederick L. The Elements of Research. New York: Prentice-Hall Inc., 1950.

### B. PERIODICALS

Andrews, Leonard O. "Admissions and Selective Retention in a University College of Education," Teacher Education: The Decade Ahead, De Kalb Conference Report, Washington, D. C.: National Education Association, 1955. pp. 252-258.

Apple, Joe A. and J. R. Shunert. "Toward a Better Secondary Teacher Education Program," Educational Administration and Supervision, 41:129-144, March, 1955.

Bancroft, Roger W. "Recruitment for Elementary School Teaching," Journal of Teacher Education, 3:193-196, September, 1952.

Barr, Arvil Sylvester. "The Measurement and Prediction of Teaching Efficiency," Journal of Experimental Education, 3:193-196, September, 1952.

- Barr, Arvil Sylvester. "The Measurement and Prediction of Teaching Efficiency: Summary of Investigations," Journal of Experimental Education, 16:203-283, June, 1948.
- Bent, Rudyard K. "Relationships Between Qualified Examinations, Various Other Factors, and Student Teacher Performance at the University of Minnesota," Journal of Experimental Education, 5:251-55, March, 1937.
- Berdie, Ralph F. "Improving Evaluation of Student Recruitment and Selection," The Personnel and Guidance Journal, 34:481, April, 1956.
- Bordin, Edward S. "Four Uses for Psychological Tests in Counseling," Educational and Psychological Measurement, 11:779-81, Winter, 1951.
- Bretsch, Howard S. and Gene S. Jacobsen. "Recruitment, Guidance, and Screening of Prospective Elementary and Secondary School Teachers," Review of Educational Research, 25:204-212, June, 1958.
- Breuckner, Leo J. "The Need of Educational Diagnosis," 34th Yearbook of the National Society for the Study of Education, Bloomington, Ill.: Public School Publishing Co., 1935. pp. 1-14.
- Cook, Walter W. "Predicting Success of College Graduate Students in Teacher Education," School and Society, 46:192-95, September, 1942.
- Dodd, M. R. "A Study of Teaching Aptitude," Journal of Educational Research, 26:517-21, March, 1933.
- Domas, Simeon J. and David Tideman. "Teacher Competence: An Annotated Bibliography," Journal of Experimental Education, 29:101-218, December, 1950.
- Downie, Norville M. and C. R. Bell. "Minnesota Teacher Attitude Inventory as an Aid in the Selection of Teachers," Journal of Educational Research, 46:699-704, May, 1953.
- Dressel, Paul Leroy and R. W. Matteson. "The Effect of Client Participation in Test Interpretation," Educational and Psychological Measurement, 10:693-706, Winter, 1950.
- Durflinger, Glenn W. "Scholastic Prediction in a Teachers College," Journal of Experimental Education, 11:257-67, June, 1943.

- Frederiksen, Norman Oliver. "Making Test Scores More Useful for Prediction," Educational and Psychological Measurement, 11:783-87, Winter, 1951.
- Frederiksen, Norman Oliver and W. B. Shrader. "The A. C. E. Psychological Examination in High School Standing of Predictors of College Success," Journal of Applied Psychology, 36:261-65, August, 1952.
- Gough, H. G. and W. H. Pemberton. "Personality Characteristics Related to Success in Practice Teaching," Journal of Applied Psychology, 36:307-309, October, 1952.
- Grim, Paul R. and W. H. Edison. "Continuous Selection in a Program of Professional Laboratory Experience," Education, 72:325-331, January, 1952.
- Hanna, Joseph V. "Estimating Intelligence by Interviews," Educational and Psychological Measurement, 10:420-30, Autumn, 1950.
- Inlow, Gail M. "Evaluating Student Teachers Experiences," Journal of Educational Research, 45:705-14, May, 1952.
- Lawton, J. A. "A Study of Factors Useful in Choosing Candidates for the Teaching Profession," British Journal of Educational Psychology, 9:131-44, June, 1939.
- Leavitt, Jerome E. "Predicting Teaching Success," Phi Delta Kappan, 35:281-2, April, 1954; Journal of Teacher Education, 4:194-7, May, 1953.
- Mallinson, George Greisen and W. M. Crumerine. "An Investigation of the Stability of Interests of High School Students," Journal of Educational Research, 45:369-83, January, 1952.
- Mead, A. R. and C. E. Holly. "Forecasting Success in Practice Teaching," Journal of Educational Psychology, 7:495-97, October, 1916.
- Ohlsen, Merle Marvel and R. E. Shultz. "Study of Variables for Use in Selection," Journal of Teacher Education, 5:279-282, December, 1954.
- Page, M. H. and R. M. W. Travers. "Relationship Between Rorschach Performance and Student Teaching," Journal of Educational Psychology, 44:31-40, January, 1953.

- Ringness, T. A. "Relationships Between Certain Attitudes Towards Teaching and Teaching Success," Journal of Experimental Education, 21:1-55, September, 1952.
- Rosenberg, Nathan. "Stability and Maturation of Kuder Interest Patterns During School," Educational and Psychological Measurement, 13:449-58, Autumn, 1953.
- Sacks, Joseph M. and Herbert S. Lewin. "Limitations of the Rorschach as Sole Diagnostic Instruments," Journal of Counseling Psychology, 14:479-81, December, 1950.
- Sands, John E. "Off Campus Student Teaching Practices in 112 Institutions," Education, 73:636-44, June, 1953.
- Shaw, Jack. "Function of the Interview in Determining Fitness for Teacher Training," Journal of Educational Research, 45:667-81, May, 1952.
- Strange, Frank B. "Student Self Selection of Group Tests," Personnel and Guidance Journal, 32:30-33, September, 1953.
- Tudhope, William B. "A Study of the Training College Final Teaching Mark as a Criterion of Future Success in the Teaching Profession, Part I," British Journal of Educational Psychology, 12:167-71, November, 1942.
- Wallace, Wimburn Leroy. "The Prediction of Grades in Specific College Courses," Journal of Educational Research, 44:587-97, April, 1951.
- White, V. "Selection of Prospective Teachers at Syracuse University," Journal of Teacher Education, 1:24-31, March, 1950.

## APPENDIX



TABLE III

THE SUBJECT'S RAW SCORES ATTAINED ON THE VARIOUS  
INDIVIDUAL INSTRUMENTS

No. +	STEP			SCAT			MVI	LAT		MTAI	AI	SSHA	CTA
	A	B	C	D	E	F		H	I				
(1)	44	28	58	49	44	46	20	26	26	42	29	48	55
*(2)	36	24	48	35	48	40	15	11	14	45	31	36	
(3)	21	29	42	30	25	25	7	10	16	3	28	57	42
(4)	46	50	64	45	40	43	16	22	29	71	55	24	59
(5)	42	40	57	49	57	63	18	29	29	45	37	46	68
(6)	46	52	67	77	57	69	24	32	48	79	32	66	75
(7)	37	35	55	55	42	48	14	24	27	73	30	53	59
*(8)	37	38	57	25	40	27	13	21	24	28	58		
(9)	33	46	56	58	37	46	21	27	22	112	16	48	76
(10)	36	40	60	51	38	45	16	23	28	50	6	43	63
*(11)	22	26	53	33	32	31				22	28	53	57
*(12)	31	40	51							4	25		
(13)	28	36	58	57	29	41	23	18	28	34	42	36	59
(14)	35	37	56	64	46	55	22	26	39	23	20	54	66
(15)	36	36	63	72	38	54	21	31	42	70	34	56	74
*(16)	41	41	60	48	56	57				17	15	42	70

+ The numbers are in lieu of the subject's names.

\* Indicates subjects who did not complete the full battery and therefore were not included in the study. Blank spaces indicate no available data.

TABLE III (continued)

No. +	STEP			SOAT			MVI	LAT		MTAI	AI	SSHA	CTA
	A	B	C	D	E	F		H	I				
(17)	35	44	58	51	40	44	17	25	40	20	50	39	58
(18)	40	40	59	46	40	41	19	29	34	62	52	45	69
*(19)	28	28	63					22	24	48	25	54	
(20)	36	49	36	53	37	44	16	27	38	72	34	58	50
(21)	40	42	53	61	44	52	21	30	49	95	23	45	62
(22)	41	36	56	36	43	31	25	26	40	61	38	53	57
(23)	47	46	68	62	70	67	22	35	49	129	40	66	83
(24)	36	41	65	51	29	37	19	23	30	25	19	60	65
*(25)	35	40	54					29	25	20	14	30	59
(26)	40	49	61	48	51	48	20	26	42	100	23	59	72
(27)	42	51	55	41	54	48	16	17	21	49	7	48	70
(28)	39	41	63	71	50	62	20	29	23	45	18	58	52
(29)	44	47	62	57	47	52	21	28	35	45	45	43	78
(30)	47	53	62	51	50	51	20	25	40	55	13	66	68
(31)	25	38	47	51	20	31	16	26	26	11	25	50	59
(32)	42	38	56	52	69	63	15	30	41	54	21	45	60
(33)	48	55	65	60	54	60	20	29	41	65	4	46	75
(34)	42	33	54	42	49	44	14	24	29	77	12	51	63
*(35)	27	33		40	41	38		26	38	22		58	
(36)	37	43	58	57	39	46	18	26	31	37	5	60	70
(37)	30	42	64	54	35	43	17	28	34	82	18	44	60

TABLE III (continued)

No.+	STEP			SCAT			MVI	LAT		MTAI	AI	SSHA	CTA
	A	B	C	D	E	F		H	I				
*(38)	34	41	51				16	30	36	39	5		67
(39)	41	40	54	61	39	49	15	27	31	15	18	58	81
*(40)	41	47	63	66	57	62		29	32	38	42	41	70
(41)	36	41	60	39	56	45	12	30	42	46	30	49	68
(42)	39	43	61	60	50	55	17	30	37	58	31	34	66
*(43)	44	42	62				17	16	16	5	16	43	
(44)	44	51	65	67	62	66	20	30	35	73	29	57	64
(45)	47	41	60	48	55	50	20	32	39	71	25	50	69
*(46)		44	54	51	36	41	18			26	30	44	66
*(47)	32	40	48	35	44	37		22	26	36	33	62	45
*(48)	41	43	58	44	40	40		22	24	35	15	42	56
(49)	46	52	68	55	45	49	17	31	40	82	25	48	71