

THE RELATIONSHIP OF A BATTERY OF SELECTED TESTS
TO SUCCESS IN THE FIRST YEAR PROGRAM IN THE
COLLEGE OF EDUCATION
OKLAHOMA STATE UNIVERSITY

By

JO STRIBLING DORRIS

Bachelor of Arts

Oklahoma College for Women

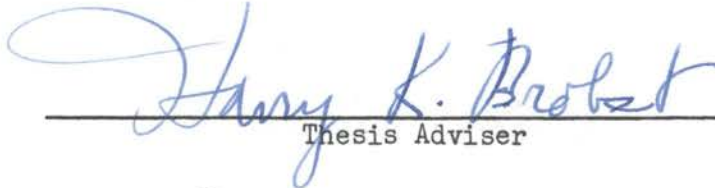
Chickasha, Oklahoma

1954

Submitted to the faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1960

THE RELATIONSHIP OF A BATTERY OF SELECTED TESTS
TO SUCCESS IN THE FIRST YEAR PROGRAM IN THE
COLLEGE OF EDUCATION AT OKLAHOMA STATE
UNIVERSITY

Thesis Approved:



Thesis Adviser







Dean of the Graduate School

SEP 1 1960

ACKNOWLEDGMENT

The author is especially grateful to Dr. Harry K. Brobst, Chairman of her Advisory Committee, for his many hours of encouragement, direction, and guidance in this research. Sincere appreciation is also extended to the other members of the Advisory Committee, Dr. Roy Gladstone and Dr. William Rambo, for their valuable suggestions and advice.

For services in statistical computation by I. B. M., the author is indebted to Professor William Granet and Mrs. Cassie Spencer of the Computation Center. For cooperation in making the necessary information available from student personnel records, the author is indebted to the staff of the Registrar's Office, and the staff of the Bureau of Tests and Measurements.

To those mentioned above and to many others who offered encouragement, suggestions and time, the author extends her deepest appreciation.

452703

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	1
Purpose of the Study and Statement of the Problem.	5
Limitations of the Study	6
II. PROCEDURE	8
Subjects	8
The Test Battery	9
The Criterion.	11
Statistics Involved	12
III. TEST BATTERY	13
Cooperative School and College Ability Tests	13
Nelson-Denny Reading Test	17
Essential High School Content Battery.	20
Minnesota Teacher Attitude Inventory	25
Summary	29
IV. REVIEW OF THE LITERATURE.	31
SCAT and Its Predecessor, ACE.	31
Nelson-Denny Reading Test.	39
Essential High School Content Battery.	43
Minnesota Teacher Attitude Inventory	45
V. PRESENTATION AND ANALYSIS OF RESULTS	48
Test Intercorrelation Matrices	48
Correlation of Tests and Over-All Grade Average at the End of Freshman Year	52
VI. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	58
Review of Objectives	58
Summary of Results	59
Conclusions	60
Recommendations.	62
BIBLIOGRAPHY	63
VITA	66

LIST OF TABLES

Table	Page
I. Test Numbers Arbitrarily Assigned to Scores Obtained on Each Test and/or Subtest: Parts of Test Combined to Obtain the Scores; and Types of Items Utilized in Each Part	10
II. Reliabilities and Standard Errors of Measurement for Verbal, Quantitative, and Total Raw Scores for Forms 1A and 2A, SCAT	15
III. Split-Half and Alternate-Form Reliability Coefficients, By Separate Grade Levels for Essential High School Content Battery	22
IV. Alternate-Form Reliability Coefficients for Parts A-H of English Test of Essential High School Content Battery	23
V. Data Concerning Significant Characteristics Which Permit Appraisal of EHSCB Normative Group	24
VI. Normative Group Samples of Minnesota Teacher Attitude Inventory: Number of Cases per Sample, the Mean Scores and Standard Deviations of Each Sample	28
VII. Spearman-Brown Reliability Estimates for School and College Abilities Test, Form 1A, and American Council on Education Psychological Examination, 1948 College Freshman Edition, Based on Results of 1955 Fall Testing Program	33
VIII. Intercorrelations of the Scores of the SCAT, Form 1A, and the ACE, 1948 College Freshman Edition, Based on Results of 1955 Fall Testing Program	34
IX. Correlations of Scores on the SCAT, Form 1A, and the ACE, 1948 College Freshman Edition, with IQ's on the Wechsler-Bellevue Intelligence Scale for Twenty-Three Ninth Grade Pupils in One Independent School	36
X. Comparison of High School Rank and Grade Point Average at the End of the Freshman Year in College	37

Table	Page
XI. Relationship of All Variables with College Grade Average	39
XII. Relationship Between Linguistic Skills and Grade Average with Intelligence Ruled Constant.	40
XIII. Mean Scores of Academically Higher and Lower Students on the Scholastic Ability Test and Each Measure of Reading Ability for Combined Courses (History, Government, English, Economics) and Fisher's t Values of Difference Between Means	41
XIV. The Mean Score for Three Groups in Relation to the Nelson-Denny Reading Scores and Corresponding Grades in College Freshman English	42
XV. Correlation Coefficients Estimating Relationship Between the Four Subtests of the EHSCB and Grade Point Average in Universities X, Y, and Z.	44
XVI. Correlation Coefficients Estimating Relationship Between the Terman-McNemar Test of Mental Abilities IQ's and Grade Point Average	44
XVII. Intercorrelations Among the Four Subtests of the EHSCB and the Terman-McNemar Test of Mental Abilities	44
XVIII. Intercorrelations of Eleven Tests (Group I): Verbal, Quantitative, Total (SCAT); Vocabulary, Paragraph Meaning, Total (Nelson-Denny Reading Test); Mathematics, Science, Social Studies, English (Essential High School Content Battery); Attitude (Minnesota Teacher Attitude Inventory)	50
XIX. Intercorrelations of Eleven Tests (Group II): Verbal, Quantitative, Total (SCAT); Vocabulary, Paragraph Meaning, Total (Nelson-Denny Reading Test); Mathematics, Science, Social Studies, English (Essential High School Content Battery); Attitude (Minnesota Teacher Attitude Inventory).	51

Table	Page
XX. Correlations of Score on Eleven Tests with Criterion, Grade Point Average at the End of the Freshman Year in College	53
XXI. Means and Standard Deviations on Eleven Tests and the Criterion for Group I (Male) and Group II (Female)	57

CHAPTER I

INTRODUCTION

"He, therefore, that is about children, should well study their natures and aptitudes, and see, by often trials, what turns they easily take; and what becomes them; observe what their native stock is, how it may be improved, and what it is for: He should consider what they want; whether they be capable of having it wrought into them by industry and incorporated there by practice; and whether it be worthwhile to endeavor it. For in many cases, all that we can do, or should aim at, is to make the best of what nature has given; to prevent the vices and faults to which such a constitution is most inclined, and give it all the advantages it is capable of. Everyone's natural genius should be carried as far as it could, but to attempt the putting another upon him, will be but labor in vain; and what is so plaister'd on, will at best fit untowardly, and have always hanging to it the ungracefulness of constraint and affectation."¹

As early as the seventeenth century John Locke recognized man's individual differences, his limitations, and his potentialities, together with the need of varying degrees and channels of education to so develop them.

It was not until the twentieth century, however, that the science of mental measurements began to take significant strides. The years from 1900 to 1915 can best be regarded as pioneer years characterized by experimentation and exploration. During this time we see the emergence

¹ Quotation by John Locke in Arthur Traxler, Margaret Selover, and Agatha Townsend, Introduction to Testing and the Use of Test Results in Public Schools, (New York, 1953), p. 1.

of the Binet Intelligence Scale and the beginning of standardized testing. Subsequent to 1915 we see great development in standardized testing in all school areas: the appearance of achievement batteries, group intelligence tests, and personality questionnaires. In the 1930's the Cooperative Test Service of the American Council on Education published achievement tests in most of the high school subjects; in 1936 the American Council on Education established the Committee on Measurement and Guidance, which has done much to enhance the effectiveness of the testing program; and in 1938 Oscar K. Buros began his publications of test reviews in the form of the First Mental Measurements Yearbook. Only recently has the State of Oklahoma established a division for certification of public school counselors. So with each passing year and its ever-increasing enrollments in our schools and universities one becomes more acutely aware of the fields of opportunity lying before him and the advantages afforded him by increased knowledge and skill in the use of the varied forms of mental measurements.

It would seem reasonable to assume that there are certain basic characteristics in terms of achievement, varied aptitudes, and interest patterns which would be necessary for success in a program designed to train students to become teachers. It further would seem reasonable that, if these characteristics could be quantified and measured, they could be used as indices capable of predicting success insofar as the academic program is concerned.

One often speaks in terms of "aptitudes," "achievement," and "attitude" without the realization of the full impact of these terms. For a better point of reference, let us pause and consider more closely what we mean when speaking of these intangibles.

According to Traxler², aptitude is "a present condition which is indicative of the individual's potentialities for the future." Underlying this definition he assumes three basic postulates:

1. Few, if any, individuals have equally strong aptitudes in all directions.
2. Individuals differ from one another in every aptitude they possess.
3. Differences among individuals and within individuals tend to persist within limits.³

As Crawford and Burnham⁴ point out, educational aptitudes are concerned "initially with individual's abilities to acquire by whatever means, knowledge and skills demanded for specific curricula of schools and colleges." For purposes of this study the latter point of reference is adopted.

Difficulty arises when an attempt is made to separate aptitude from achievement by means of testing. It is the contention of some that the commonality of elements is so great and the correlation between aptitude and achievement so high that it is impossible to measure each separately.

The tests designed to measure aptitude and those intended to measure achievement truly do bear relationship to each other. Both are concerned, at least indirectly, with more specific lines of endeavor and both are validated by correlation of their scores with other measures of subsequent attainment.

² Arthur E. Traxler, Techniques of Guidance (New York, 1945), p. 42.

³ Ibid., p. 43.

⁴ Albert B. Crawford and Paul S. Burnham, Forecasting College Achievement (New Haven, 1946), p. 4.

As used in this study, their proposed intentions differ in that an achievement test is a review, so to speak, designed to measure how effectively a student has learned a particular subject and, of course, is administered upon completion of such course of study. On the other hand, the aptitude test is more on the order of a preview and is intended to give an indication of how well a student is likely to do even along lines yet untried.

Many questions arise when entering the realm of attitudes: Do they influence a person's choice of vocation? Do they influence his success or failure in a vocation? Will they determine his personality? Chave has defined attitude as "a complex of feelings, desires, fears, convictions, prejudices or other tendencies that have given a set, or readiness to act, to a person because of varied experiences."⁵

Downie⁶ defines attitude in much the same way, speaking of it as "a change of potential that the individual has," including in it the aspects of personality, such as interests, appreciations, mores, and morals.

Allport and Thurston have conducted extensive investigation in the general area of attitude and have devised attitude scales and summarized the methods of acquisition. Various techniques have been employed for the measurement of attitude, including error-choice, free-response, paired comparisons, and opinion polls.

⁵ Traxler, Techniques of Guidance, p. 105.

⁶ N. M. Downie, Fundamentals of Measurement, (New York, 1958), p. 333.

Purpose of the Study and Statement of the Problem

This study is designed to assist the counselors in the College of Education in improving their effectiveness in working with entering freshmen by giving these counselors some idea of what to expect from a student on the basis of his test scores. If the counselor has these test data at hand, on the basis of the results of this study, by substituting the weights, he can make some prediction at the beginning of the program as to how the individual will do by the end of the year.

Although it is true that no test is an absolute indicator of a student's ability to succeed in any given program, the fact remains that any degree of success in guidance is to be desired over and above a chance guess. Therefore, by continued investigations such as this one, the degree of success in counseling should be improved and result in a more adequate counseling program.

The specific problem may be stated as follows: To what extent can success in the first year of training in the College of Education be predicted by the following battery of measuring instruments:

1. Cooperative School and College Ability Tests (SCAT), Form 1A.
2. Nelson-Denny Reading Test, Form A.
3. Essential High School Content Battery (EHSCB), Form AM.
4. Minnesota Teacher Attitude Inventory (MTAI), Form A.

On the basis of these four groups of tests, what is the multiple correlation and corresponding weights for each test for predicting the criterion, i.e., the cumulative grade point average, at the end of the first year. More specifically:

1. What is the relationship between performance on the SCAT and the criterion?
2. What is the relationship between performance on the Nelson-Denny Reading Test and the criterion?
3. What is the relationship between performance on the Essential High School Content Battery and the criterion?
4. What is the relationship between performance on the Minnesota Teacher Attitude Inventory and the criterion?
5. To what extent does the best combination of these tests predict the criterion at the end of the first year program?

Limitations of the Study

Although it is true that the study is undertaken for the purpose of obtaining information for use in counseling entering freshmen in the College of Education, the research is far from complete and should be used only as a beginning of a more thorough study of the situation. The study is limited in its scope because it deals with the freshmen entering Oklahoma State University during only one semester. Also, the process of subject selection is limited - the subjects were selected due to the fact that complete test data were available in their files. There were also many other students who enrolled during this semester from whom we did not have complete data. Therefore, it is entirely possible that the sample is not representative of the entire class of 1958.

A further limitation is inherent in the size of the groups included in the study, more particularly Group I (males), which included only thirty-one subjects. This is hardly a large enough sample to allow

generalization. As will be seen in a subsequent chapter, the size of Group I limited the selection of the tests to only one.

It goes without saying that there is no claim for universality, for the study is confined to one particular college on the Oklahoma State University campus. It probably would not be even suitable for use in another college on the campus since there is no evidence that this sample of students is representative of entering freshmen as a whole.

However, despite the limitations inherent in the study, it does offer possibilities by laying a foundation for the accumulation of normative data which could be used for more accurate prediction of success in the first year of training in the College of Education.

CHAPTER II

PROCEDURE

This chapter is devoted to the presentation of the conditions under which the present study was conducted and the procedures employed. The subject groups and the tests involved will be described, the test administration conditions discussed, and the criterion indicated and defined.

Subjects

Subjects included in this study were selected from the total number of freshmen students who enrolled in the College of Education at the Oklahoma State University in the fall of 1958 and who remained in school both semesters. From this group subjects were selected whose individual records contained the following information:

1. Complete test data on all tests included in this study.
2. An end-of-the-year grade point average.

The total group was then divided according to sex. Group I consists of thirty-one male subjects, and Group II consists of seventy-five female subjects. It is realized that Group I constitutes a small sample and that it would be inadvisable to rely alone on the results reflected by this study. However, these results may be used as a foundation for the accumulation of future normative data. Group II is a more adequate representation of the freshmen women in the College of Education, and results obtained from these data provide a more reliable possibility

for prediction. It is believed that more valid information is obtained through this division than would otherwise result from merely one group consisting of both male and female subjects. Counselors in the College of Education are interested in any existing differences in the academic success of freshmen men and freshmen women and the possible reasons for these differences. Although this study is not designed to indicate reasons for these differences, the division by sexes makes it possible for the study to lend itself as supporting evidence for further study.

The Test Battery

The tests involved in this study were administered to the students at the beginning of their freshman year by the College of Education. These tests were scored by the Bureau of Tests and Measurements at the Oklahoma State University, and aid was given by the bureau in the interpretation of test results.

Four groups of tests were selected to constitute the test battery. The Cooperative School and College Ability Tests (SCAT) yields verbal, quantitative, and total scores; the Nelson-Denny Reading Test yields vocabulary, paragraph meaning, and total scores; the Essential High School Content Battery yields mathematics, science, social studies, and English scores; and the Minnesota Teacher Attitude Inventory yields a single score.

As stated above, each test yields from one to four separate scores. Each separate score has been assigned a number and will hereinafter be treated as an individual measure. Table I reflects the number arbitrarily assigned to each score and indicates which parts of the tests were combined

TABLE I

TEST NUMBERS ARBITRARILY ASSIGNED TO SCORES OBTAINED ON EACH TEST AND/OR SUBTEST; PARTS OF TEST COMBINED TO OBTAIN THE SCORE; AND TYPES OF ITEMS UTILIZED IN EACH PART

Assigned Test Number	Corresponding Test or Part Name	Description of Test
Test 1 (Verbal)	Part I (SCAT) Part III (SCAT)	Sentence Understanding Word Meanings
Test 2 (Quantitative)	Part II (SCAT) Part IV (SCAT)	Numerical Computation Numerical Problem Solving
Test 3 (Total)	Parts I, II, III, and IV (SCAT)	
Test 4 (Vocabulary)	Part I (Nelson-Denny Reading Test)	Vocabulary
Test 5 (Paragraph Meaning)	Part II (Nelson-Denny Reading Test)	Reading and Understanding Paragraphs
Test 6 (Total)	Parts I and II (Nelson-Denny Reading Test)	
Test 7 (Mathematics)	Test 1 (Essential High School Content Battery)	Mathematics
Test 8 (Science)	Test 2 (Essential High School Content Battery)	Science
Test 9 (Social Studies)	Test 3 (Essential High School Content Battery)	Social Studies
Test 10 (English)	Test 4 (Essential High School Content Battery)	English
Test 11 (Attitude)	Minnesota Teacher Attitude Inventory	Factors Related to Teachers' Attitudes Toward Pupils

to obtain this score in addition to describing the types of items utilized in each part.

A brief description of the tests is presented in Chapter III, including expectations, opinions of test experts regarding the tests, information concerning reliability and validity of the tests, and methods by which normative data were obtained and norms established.

The Criterion

Although there has been much discussion regarding the accuracy of grades as a basis for predicting a student's ability to succeed in further schooling, this still remains the means by which success or failure in college is determined. Therefore, the over-all grade point average at the end of the freshman year is the criterion used for this study. This is based on the assumption that, whereas one or even two quite low grades might not be a true representation of a student's ability, the summation and average of grades obtained in all courses will counter-balance the discrepancies which might occur as a function of teacher-student personality factors, a student's lack of interest and his effort exerted in a particular course.

The official files of the registrar at the Oklahoma State University were consulted in gathering these data. All averages were computed from the records made by the students of the two groups and based on the following weights:

A = 4 points

B = 3 points

C = 2 points

D = 1 point

F = 0 points

All courses which used the accepted five-point grade system were included in computing the over-all grade point averages.

Statistics Involved

The Pearson-Product Moment Method was used to compute the zero-order correlation coefficients between the criterion and each test and between each test and every other test. In Chapter V these results are presented in a correlation matrix for each sex.

The Wherry-Doolittle Method of Test Selection was used to select the best combination of tests for counseling purposes and for purposes of prediction of a student's probable success in the academic program as outlined by the College of Education. By this method the multiple correlations and weights were computed for the groups and the tests of the battery were analytically selected and added one at a time until a maximum R was obtained. From this a multiple regression equation was developed from which the criterion can be predicted with the greatest possible accuracy provided by the optimum combination of tests.

The results and analysis in addition to a more detailed discussion of the method used is presented in Chapter V.

CHAPTER III

TEST BATTERY

This chapter is devoted to the test battery used in this investigation. Discussions have been included concerning the reliability, validity, standardization, and norms of the various tests. Also included is a summary of the opinions of prominent test experts as found in Buros' Mental Measurements Yearbooks.

Cooperative School and College Ability Tests (SCAT)

The principal aim of the SCAT is to aid in the estimation of a student's capacity to undertake additional schooling. The test consists of four parts designed to measure the following skills:

Part I - getting the meaning of isolated sentences.

Part II - performing numerical computations rapidly.

Part III - associating meanings of isolated words.

Part IV - solving arithmetic problems.

Parts I and III are combined to obtain a verbal (V) score and Parts II and IV to obtain a quantitative (Q) score. The addition of all four parts yields a total (T) score. The time required for administration is approximately one hour and fifteen minutes.

The SCAT Manual for Interpreting Scores¹ suggests the following ways to use the test:

¹ SCAT Manual for Interpreting Scores, (Princeton, 1957), p. 6.

1. Identifying a student who is especially advanced or retarded in abilities needed for the next higher level of schooling.

2. Suggesting the level of instruction appropriate to the needs of the individual.

3. Comparing a student's abilities as measured by SCAT with his achievement in school courses, in order to identify and to help the over or under achiever.

4. Grouping students of similar levels of ability for purposes of instruction.

5. Guiding a student toward appropriate academic goals and courses by encouraging him to consider his present developed abilities in making future academic plans.

6. Giving parents a realistic and straightforward estimate of a student's capacity for the next level of academic work.

7. Studying the development of a student's ability over a period of years.

At the college freshman level the test is available in Forms 1A and 1B. These tests are constructed in the same manner and purported to be equal in difficulty.

The reliability coefficients listed in the SCAT Technical Report are computed by internal analysis. No test-retest correlations or alternate form correlations have been obtained. Table II reflects the reliability coefficients for Forms 1A and 2A designed for use at the college freshman and high school senior levels, respectively. The Kuder-Richardson formula was used to estimate all reliabilities and the standard error of measurement. Reference to Table II will show that in every instance the

TABLE II²

RELIABILITIES AND STANDARD ERRORS OF MEASUREMENT FOR
VERBAL, QUANTITATIVE, AND TOTAL RAW SCORES FOR
FORMS 1A AND 2A.

Form	Grade	Number of Examinees	Score	Number of Items	Reliability	Standard Error of Measurement
1A	13	612	V	60	.92	3.26
			Q	50	.93	2.80
			T	110	.95	4.29
2A	11	2,292	V	60	.92	3.27
			Q	50	.90	2.86
			T	110	.95	4.34

² SCAT Technical Report, (Princeton, 1957) p. 11.

reliability coefficient is greater than .90 with the total score in both instances .95.

Unfortunately, there are no completed validity studies reported in the manual. The validity can only be inferred by the items appearing in the test and their relevance to the test. However, validity studies are presently being conducted and are slated to be published upon completion in a supplemental report. Hanford M. Fowler³ reports that a preview of these studies indicates favorable results: the reported validity coefficients are as high or higher than similar coefficients reported for other tests. It is his feeling that the SCAT shows promise of being an efficient measure of future scholastic success.

The standardization program as presented in the manual reflects great care on the part of the authors to obtain a truly representative national sample. Norms are presented by grade for the interpretation of scores of the individual student and school mean norms for interpretation of mean scores of a group of students for grades four through twelve. For the college norms the sampling program drew twelve students from grade 13 and twelve from grade 14 in 120 different colleges over the United States.

Stratification of the sample was accomplished through classification by regions (North, South, and far West) and classification of types of colleges (two-year and four-year). This was done in an effort to partially compensate for the difference in sizes of the colleges.

The norm tables present percentile bands for each score interval, the higher band corresponding to a score about one standard error of measurement

³ Oscar K. Buros, Fifth Mental Measurements Yearbook (New Jersey, 1959), p. 451.

above the midpoint of the interval and the lower corresponding to a score about one standard error of measurement below the midpoint. This is to impress the user with the fallibility of such an instrument and the caution necessary in the interpretation thereof.

Fowler⁴ and Frederick P. Davis⁵ present their opinions of the SCAT in the Fifth Mental Measurements Yearbook. In essence they are in agreement as to the strong features of SCAT, citing such things as the attractive brochure which presents adequate directions for administration, the straightforward scoring of the SCAT, the use of percentile bands in order to call attention to the error inherent in the instrument, discussion relating to the meaningful interpretation, and the high reliability coefficients.

Both men feel that the most serious limitation lies in the lack of validation. However, a supplemental report which should at least partially remedy the situation should be appearing in the near future. Further questions were raised regarding the high intercorrelation coefficients between quantitative and verbal scores, and the amount of time required to administer the SCAT. Davis feels that the SCAT will be moderately useful for many educational purposes but probably not as useful for one particular purpose. For example, Davis is of the opinion that grades are a better index for prediction than a SCAT score.

Nelson-Denny Reading Test

This test consists of two parts, a test of vocabulary and a test of paragraph meaning, and yields these two scores plus a total score. The

⁴ Ibid., p. 451-453.

⁵ Ibid., p. 450-451.

vocabulary test is a 100-item multiple choice test, and the comprehension test is a 36-item test consisting of nine paragraphs, each followed by four comprehensive questions.

According to the manual of directions, the test serves a three-fold purpose:

1. To predict probable success in college.
2. To section incoming college or high school classes.
3. To aid in the diagnosis of students' difficulties.

The test is available in Forms A and B, which are constructed in the same manner and purported to be equal in difficulty. The entire test requires thirty minutes for administration: ten minutes for the vocabulary section and twenty minutes for the paragraph test.

The reliability coefficient was determined by the alternate form method. Form A, followed by Form B, was administered to 171 college freshmen. Standard deviations of scores were 21.7 and 22.2 for Forms A and B, respectively. The correlation between the two forms of the test is .91 plus or minus the standard error of .01.

Other than content validity, the manual has little to say regarding the validity of the test. The authors list their sources of vocabulary as:

1. Thorndike Word List.
2. Horn's Basic Writing Vocabulary.
3. Various Tests.

Four hundred words were selected from the Thorndike List and Horn's Basic Writing Vocabulary, and 200 additional words were taken from various tests not enumerated by the author. Five definitions, one of which was correct, were assigned to these 600 words, and the words were then divided into three groups of 200 items each. All of them were administered to

approximately 390 students representing all of the four college classes. A number of high school students were also included in this group. The frequency of error for each item was determined, and from the initial 600-word list 200 were chosen, 100 of which appear in each form of the test. The items appearing in Form A of the test are purported to be of a difficulty equal to those in Form B. The items are arranged in order of increasing difficulty.

The procedure for the selections appearing in the paragraph test was similar. Twenty-seven 200-word selections which supposedly represent various types of reading matter were chosen with the following ideas in mind:

1. The reading would not require acquaintance with a technical vocabulary.
2. The general character of the reading is that which a student is likely to meet in college work.

These selections were administered to 450 students, and eighteen selections were finally chosen, nine of which appear in each test form.

The only claim for congruent validity that appears in the manual is the statement that one form of the test and the scores of an objective child psychology test correlated $.70$ plus or minus $.04$. Thus the author claims the test is capable of predicting college success.

The manual of instructions is quite limited with regard to standardization and norms. Two tables reflect the norms: one for senior high school students and one for college and university students. There is virtually no information about how these norms were established, i.e., the standardization procedures, the size of the normative groups, the

characteristics of the group. This seriously limits the user since there is no way to determine the applicability of the norms to the group concerned. If the test were to be used to any great extent, the only alternative would be to establish norms applicable to the population involved.

Ivan A. Booker⁶ and Hans C. Gordon⁷ are favorably impressed with the Nelson-Denny Reading Test, citing such advantages as careful item construction, selection and difficulty gradation, equated alternate forms, and reliability coefficients of approximately .90.

The considered limitations include no correction for wrong answers, too restrictive time limits, and the possibility of having encountered the reading material in previous course work.

Essential High School Content Battery (EHSCB)

The Essential High School Content Battery is designed to be a comprehensive battery of high school achievement tests, including four basic areas: mathematics, science, social studies, and English. The test appears in alternate forms, AM and BM, and requires five class periods for administration.

The primary intention of the battery is based on the premise that "there is a common body of knowledges and skills which it is reasonable to expect a high school graduate to possess . . . to provide continuing measures of the student's growth and development with respect to these knowledges and skills."⁸

⁶ Oscar K. Buros, Fourth Mental Measurements Yearbook, (New Jersey, 1953), pp. 544 - 545.

⁷ Oscar K. Buros, The Nineteen Forty Mental Measurements Yearbook, (New Jersey, 1941), pp. 357-358.

⁸ David P. Harry and Walter N. Durost, Essential High School Content Battery Manual of Directions, (New York, 1951), p. 1.

The reliability information contained in the manual is quite extensive. It presents the reliability coefficients, using the split-half and alternate-form methods. Tables III and IV, taken from the manual, best reflect the coefficients. Table III shows that in only one instance was the reliability coefficient .90 or above by using the alternate-form method. Table III also reflects the standard error of measurement, which is a measure of reliability to be expected of an individual score. The chances are about two out of three that an individual's score will not vary from his true score more than the amount indicated by the standard error of measurement. Table IV is a breakdown of the eight parts of the English Test and their reliability coefficients. The highest coefficient reflected by Table IV is Part H in Grade 10 (.78), which is hardly high enough to be relied upon for individual counseling.

The validity presented in the manual deals with curricular validity. An analysis was conducted of the instructional materials and authoritative pronouncements in the various fields. The justification of the items included were based on the frequency of inclusion in commonly used textbooks and expert judgment as to importance.

The authors¹¹ point out what they consider to be limitations of this validity:

1. The impossibility to measure all of the important outcomes in any area.
2. The restrictions placed on a machine-scored test.
3. The fact that tests designed to discriminate well even among the top high school seniors will appear unduly difficult for ninth graders.

¹¹ Ibid., p. 2.

TABLE III⁹SPLIT-HALF (r_{11}) AND ALTERNATE-FORM (r_{AB}) RELIABILITY COEFFICIENTS, BY SEPARATE GRADE LEVELS

TEST	GRADE 10				GRADE 11				GRADE 12			
	N	r_{AB}	r_{11}	SE Meas.	N	r_{AB}	r_{11}	SE Meas.	N	r_{AB}	r_{11}	SE Meas.
Mathematics	114	.85	.87	4.2	101	.92	.93	4.2	96	.87	.95	3.9
Science	297	.67	.76	5.1	268	.78	.89	3.8	252	.81	.89	4.5
Social Studies	176	.83	.88	4.7	151	.87	.90	4.1	145	.84	.88	4.1
English	215	.86	.90	4.1	181	.87	.90	3.0	119	.85	.91	3.6
Total Battery	157	-	.95	2.6	113	-	.95	2.7	155	-	.96	2.4

⁹ Ibid., p. 15.

TABLE IV¹⁰

ALTERNATE-FORM RELIABILITY COEFFICIENTS FOR PARTS A-H OF ENGLISH TEST

P A R T	Reliability Coefficients		
	Grade 10 N = 215	Grade 11 N = 181	Grade 12 N = 120
A. Reading for Information	.53	.49	.38
B. Vocabulary	.63	.73	.70
C. Business Definitions	.36	.46	.36
D. Use of References	.47	.48	.58
E. Literature Acquaintance	.52	.62	.58
F. Language Usage	.55	.62	.53
G. Capitalization and Punctuation	.61	.60	.62
H. Spelling	.78	.72	.76

10 Ibid.

The norms are presented in percentiles. At the end of the school year 1950 the test was administered to 46,000 students in 288 high schools from 35 states. Data were determined which made it possible to decide whether the group was actually representative of the total population of which it is a sample. These data are reflected in Table V.

TABLE V¹²

DATA CONCERNING SIGNIFICANT CHARACTERISTICS WHICH PERMIT
APPRAISAL OF EHSCB NORMATIVE GROUP

Grade	Number of Cases	Median Terman-McNemar IQ	Median Chronological Age
9	10,700	100.64	15-1
10	11,870	102.77	15-11
11	10,850	103.64	16-11
12	9,960	105.77	17-11

The normative data were computed in terms of various geographical divisions of the country and weighted so that they are based on a group whose distribution corresponds closely to the national high school population. Normative data by grade and test are presented for total school population, students taking academic and scientific courses, and students taking commercial and general courses.

Herbert S. Conrad¹³ and J. Thomas Hastings¹⁴ commend the test authors on the excellent presentation of data, the meaningful interpretation, and

¹² Ibid., p. 10.

¹³ Oscar K. Buros, Fourth Mental Measurements Yearbook, pp. 18-24.

¹⁴ Ibid., pp. 24-25.

the careful item selection. Several questions were voiced relating to the absence of intercorrelations between tests, standard deviations of difference scores, and the relevance of some of the items.

Harry A. Passow¹⁵ further points out the fact that there is no correction made for guessing. He is concerned, too, with the somewhat low reliability coefficients, more particularly those reflected by the English Test. Passow is of the opinion that the English Test could have been improved with more stress on communication and less on formal grammar.

Minnesota Teacher Attitude Inventory

This inventory consists of 150 items designed to reflect the attitudes of teachers toward pupil-teacher relations and includes five basic areas:

1. Moral status of children in the opinion of adults.
2. Discipline and problems of conduct in the classroom and elsewhere and methods employed in dealing with such problems.
3. Principles of child development and behavior related to ability, achievement, learning, motivation, and personality.
4. Principles of education related to philosophy, curriculum, and administration.
5. Personal reactions of teachers, likes and dislikes, sources of irritation.

The manual suggests that the inventory may be used for the selection of students for teacher preparation, and the selection of teachers for both teaching and student counseling positions.

¹⁵ Ibid., pp. 25-26.

No investigations are mentioned which deal with the validity of the test in predicting academic success. All validation studies cited in the manual deal with the prediction of the type of social atmosphere a teacher will maintain in the classroom. The test is included in this battery to determine whether there is a relationship between attitudes expressed and the student's success in the first year program of the College of Education.

A great deal of time and effort went into the development of the inventory. The manual describes in detail the experimental tryouts and presents reliability coefficients for the experimental forms based on three criteria: principals' ratings, pupils' ratings, and experts' ratings. Upon completion of Form A, three methods of scoring were devised, from which the one yielding the highest split-half reliability coefficient of .93 was selected.

The development of the test included two experimental forms, X-164 and X-239, from which resulted the published Form A. The items appearing in Form A were selected according to:

1. Their power to discriminate between successful and unsuccessful teachers.
2. The extent to which item responses are influenced by educational courses.
3. The extent to which each item response is influenced by experience.
4. The extent to which the content of one item duplicated another.
5. Clarity of the statement.

6. Consistency of response patterns of superior and inferior teachers.¹⁶

Of the 150 items appearing in Form A, 129 were taken from experimental Form X-164 and 21 from Form X-239.

Two studies, the South Carolina Study and the Missouri Study, were conducted to further check the validity of Form A. The South Carolina Study yielded an R of .63 when the three criteria were combined, i.e., principals' ratings, pupils' ratings, and experts' ratings. The correlation coefficient obtained from the composite criterion in the Missouri Study is .46, somewhat lower than that of the Carolina Study.

Norms are presented for high school seniors, university freshmen, graduate students in education, and experienced teachers. Table VI reflects the various norm group samples, the number of cases per sample, and the mean scores and standard deviations of the samples.

When the norm groups for experienced teachers was determined, a stratified random sample was selected from the teacher population in Minnesota, and analysis of variance and co-variance techniques were applied to determine the significance of such factors as experience, amount of post-high school education, size of the community in which the teacher is employed, grade level taught, and subjects taught.

Two reviews of the Minnesota Teacher Attitude Inventory appear in the Fourth Mental Measurements Yearbook. Both Dwight Arnold¹⁷ and Lee J. Cronbach¹⁸ agree that careful and extensive study went into the preparation

¹⁶ Walter W. Cook, Carroll H. Leeds, and Robert Callis, Minnesota Teacher Attitude Inventory Manual, (New York, 1951), p. 13.

¹⁷ Buros, Fourth Mental Measurements Yearbook, pp. 797-798.

¹⁸ Ibid., pp. 798-799.

TABLE VI¹⁹

NORMATIVE GROUP SAMPLES OF MINNESOTA TEACHER ATTITUDE INVENTORY: NUMBER OF CASES PER SAMPLE, THE MEAN SCORES AND STANDARD DEVIATION OF EACH SAMPLE.

Norm Group	Number	Mean	Standard Deviation
High School Seniors	122	12.8	31.6
University Freshmen	384	4.8	28.9
Beginning Education Juniors			
Early Childhood	134	65.9	29.8
Elementary	228	59.5	26.3
Secondary			
Academic	136	48.3	29.3
Nonacademic	238	44.1	27.1
Graduating Education Seniors			
Early Childhood	108	80.4	22.6
Elementary	150	77.4	24.7
Secondary			
Academic	237	67.8	24.3
Nonacademic	185	63.3	25.4
Graduate Students in Education	200	64.0	33.3
Experienced Teachers			
Elementary Teachers (Systems with more than 21 Teachers)			
2 years training	249	40.1	37.2
4 years training	247	55.1	36.7
Secondary Teachers (Academic)			
4 years training	264	24.7	40.6
5 years training	218	40.8	39.5
Secondary Teachers (Nonacademic)			
4 years training	98	9.7	42.7
5 years training	70	28.9	36.5

¹⁹ Harry and Durost, pp. 8 and 9. (Taken from information in Tables 1 and 2.)

and development of the inventory. However, both reviewers caution against placing much faith in the inventory until considerably more investigation of its validity can be completed.

Arnold emphasizes a very evident limitation: none of the data presented in the manual involves studies of scores made by students in or before teacher training in relation to success later in teaching.

A further critical issue arises when examining the norm tables - see Table VI. The mean score of students is about 30 to 45 points higher than the mean score of experienced teachers. This raises several questions:

1. Which attitudes are most desirable: those of students or those of experienced teachers?

2. Will these discrepancies appear in other groups?

3. If they do, then either the inventory is of no value, the teacher education programs are unrealistic, or experienced teachers have undesirable attitudes.

Summary

For purposes of this study the SCAT is used as a measure of aptitude, the Essential High School Content Battery and the Nelson-Denny Reading Test as measures of achievement in the content areas mentioned above, and the Minnesota Teacher Attitude Inventory as a measure of attitudes toward teacher-pupil relationships. With the exception of the Minnesota Teacher Attitude Inventory, the tests of this battery are designed to aid in the prediction of a student's academic ability in the various content area. The Minnesota Teacher Attitude Inventory is designed to predict the type of social atmosphere the teacher will maintain in the classroom. Since

this study is based on students who enrolled in a program of teacher education, it is possible that attitudes expressed might also bear a relationship to academic success in the program. For this reason the Minnesota Teacher Attitude Inventory is included in the study.

CHAPTER IV

REVIEW OF THE LITERATURE

A great deal of research has been conducted along the line of prediction. Almost every imaginable factor bearing even a slight relationship with academic success has become the focal point of investigation. The present study is concerned with the factors of aptitude, achievement, and attitude as indices of academic success through the freshman year in the College of Education at the Oklahoma State University. The specific tests involved are the SCAT, the Nelson-Denny Reading Test, the Essential High School Content Battery, and the Minnesota Teacher Attitude Inventory. This chapter will be devoted to previous studies related to these and similar tests.

SCAT and Its Predecessor, the ACE

The Cooperative School and College Ability Tests (SCAT) has recently superseded the American Council on Educational Psychological Examination (ACE), and as yet very few validity studies are available on the SCAT. On the basis of their similarity in content and purpose, studies utilizing the ACE will also be included in the review of literature.

The ACE is basically considered a scholastic aptitude test. However, it contains parts which are considered essentially achievement: the completion test and the arithmetic reasoning test. The ACE has been found to correlate between total test scores and freshman grade point average about

.45. Traxler¹ has conducted a rather extensive study comparing the ACE and the SCAT as to difficulty and value for predicting school marks.

The SCAT is intended to aid in estimating the capacities of students to undertake additional schooling. Frederick Davis² is of the opinion that SCAT will be less useful for prediction of academic success than would marks in previous grades of schooling. Traxler³ states that the "SCAT is what is known in Educational Testing Service parlance as a test of developed ability."

According to Traxler's study there is no true marked difference between the SCAT and the ACE with regard to difficulty: SCAT quantitative is a little easier than the ACE quantitative, SCAT verbal and ACE linguistic are comparably difficult with SCAT being perhaps a bit more difficult, and the total scores of each test are comparable.

As to their relative predictive value, Traxler ran 108 correlations: eighteen sets of correlations were obtained with English marks, eighteen sets with marks in mathematics, and eighteen sets with average marks. It is interesting to note the existing relationships. The SCAT verbal and ACE linguistic were found to be equally effective in predicting English marks, each yielding r's of .53. SCAT quantitative seems to hold a slight margin over ACE quantitative in predicting mathematics grades, yielding a median r of .45 as compared to .35 for the ACE. SCAT total again holds a

¹ Arthur E. Traxler, "Should SCAT Scat ACE?" Educational Records Bulletin, LXVII (1956), 51-63.

² Buros, Fifth Mental Measurements Yearbook, p. 450.

³ Traxler, "Should SCAT Scat ACE?" p. 53.

slight margin over ACE total in predicting average marks, with a median r of .52 as compared to .48 for ACE. Based on the above results, which were obtained from 587 students in five schools, Traxler concluded that the ACE and SCAT are equally effective in predicting school success.

North⁴ also conducted a study comparing the ACE and SCAT with special emphasis on reliabilities, intercorrelations and correlations with diagnostic tests. In his study North used test results for 600 schools which had administered both the SCAT, Form 1A, and the ACE, 1948 College Freshman Edition, to a total group of 184 tenth grade pupils and 153 twelfth grade pupils in the fall of 1955. Table VII reflects the reliability estimates North obtained when using the Spearman-Brown formula for estimating reliability.

TABLE VII⁵

SPEARMAN-BROWN RELIABILITY ESTIMATES FOR SCHOOL AND COLLEGE ABILITY TESTS, FORM 1A, AND THE AMERICAN COUNCIL ON EDUCATION PSYCHOLOGICAL EXAMINATION, 1948 COLLEGE FRESHMAN EDITION, BASED ON THE RESULTS OF THE 1955 FALL TESTING PROGRAM

Test Score	Grade 10 (184 Pupils)		Grade 12 (153 Pupils)	
	SCAT	ACE	SCAT	ACE
Verbal (SCAT) or Linguistic (ACE)	.89	.93	.92	.96
Quantitative	.88	.90	.93	.92
Total	.93	.94	.94	.97

⁴ Robert D. North, "Comparison of SCAT and ACE," Educational Records Bulletin, LXVII (1956), 65-72.

⁵ Ibid., p. 67.

One will note that these estimates indicate that the ACE test scores tend to be slightly more reliable except in the quantitative area at the twelfth grade level.

Table VIII reflects the intercorrelations of the SCAT scores and the ACE scores based on the results from the 1955 independent school testing program.

TABLE VIII⁶

INTERCORRELATIONS OF THE SCORES OF THE SCAT, FORM 1A, AND THE ACE, 1948 COLLEGE FRESHMAN EDITION, BASED ON RESULTS OF 1955 FALL TESTING PROGRAM.

Test Score	Grade 10 (184 Pupils)		Grade 12 (153 Pupils)	
	SCAT	ACE	SCAT	ACE
Verbal or Linguistic vs. Quantitative	.53	.54	.55	.56
Verbal or Linguistic vs. Total	.88	.92	.90	.94
Quantitative vs. Total	.87	.83	.86	.80

One may notice that the intercorrelations are rather high with the highest ones being between verbal or linguistic and total score.

In order to better determine the efficiency of differential prediction, North used the results reflected in Tables VII and VIII and computed the estimated reliability of difference between the verbal and quantitative SCAT scores and linguistic and quantitative ACE scores. The estimated reliability of difference between verbal and quantitative SCAT is .76 at the tenth grade level and .83 at the twelfth grade level,

⁶ Ibid., p. 69

between the ACE linguistic and quantitative scores it is .82 at the tenth grade level and .86 at the twelfth grade level. It, therefore, appears that the ACE has a slight margin over the SCAT at both levels with respect to reliability of difference between scores on the two tests.

According to North's study, the SCAT verbal scores correlate with ACE linguistic scores .78 at the tenth grade level and .83 at the twelfth grade level, the SCAT quantitative correlates with ACE quantitative .65 at the tenth grade level and .75 at the twelfth grade level, and the SCAT total correlates with the ACE total .81 at the tenth grade level and .85 at the twelfth grade level.

When correlated with the Diagnostic Reading Test, Survey Section, Form F, the SCAT verbal correlated .74 at the tenth grade level and .75 at the twelfth grade level, whereas the ACE linguistic correlated .79 and .80 at the tenth and twelfth grade levels, respectively.

A further investigation was conducted by North⁷ wherein he compared the 1955 and 1956 fall testing programs. In addition he compared the correlations of the SCAT and ACE with the Wechsler-Bellevue IQ's for a small group of pupils in one independent school.

Even though the groups tested in 1956 were somewhat smaller than the groups tested in 1955, the medians for the two programs are quite similar: total score median for 1955 group, grade nine, is 289.0 as compared to 293.9 for the 1956 group; at the tenth grade level, 295.2 for the 1955 group as compared with 297.0 for 1956; at the twelfth grade level the median total score for the 1955 group was 304.6 as compared to

⁷ Robert D. North, "Further Report on SCAT," Educational Records Bulletin, LXIX (1957), 60-62.

302.8 for the 1956 group. This is evidence that independent school norms, when based on groups of several hundred or more pupils, tend to be quite stable.

Table IX reflects the correlations of scores on the SCAT and the ACE with IQ's on the Wechsler-Bellevue Intelligence Scale. One may note that the highest correlation with the Wechsler-Bellevue full scale IQ is .78 yielded by the ACE total score. The SCAT total scores have only a moderate correlation of .56 with the Wechsler-Bellevue Full Scale IQ.

TABLE IX⁸

CORRELATIONS OF SCORES ON THE SCAT, FORM 1A, AND THE ACE, 1948 COLLEGE FRESHMAN EDITION, WITH IQ'S ON THE WECHSLER-BELLEVUE INTELLIGENCE SCALE FOR TWENTY-THREE NINTH GRADE PUPILS IN ONE INDEPENDENT SCHOOL

Wechsler-Bellevue Test	SCAT Verbal	ACE Linguistic	SCAT Quant.	ACE Quant.	SCAT Total	ACE Total
Verbal IQ	.41	.49	.71	.65	.67	.76
Performance IQ	.13	.06	.36	.62	.30	.52
Full Scale IQ	.37	.33	.56	.75	.56	.78

Manual⁹ conducted an investigation at the University of Texas regarding aptitude tests for college admissions. In the fall of 1955 some 1,500 freshmen were administered the SCAT, the Test of Word-Number Ability, and the ETS Cooperative English Test A.

Intercorrelations were computed between the tests and zero-order correlations were determined between the tests and grade point average at the

⁸ Ibid., p. 61

⁹ Herschel T. Manual, "Aptitude Tests for College Admissions," Fourteenth Yearbook: National Council on Measurements Used in Education, (New Jersey, 1957), 20-27.

end of the freshman year. The Word-Number Total reflected the highest correlation coefficient of .61 with grade point average and the Word-Number verbal correlated .57. The SCAT verbal, quantitative, and total correlated .54, .50, and .58, respectively. When the various tests were averaged and the averages correlated with grade point, the average of the total SCAT and total Word-Number Test yielded the highest correlation coefficient of .62.

A further predictive comparison was offered in the form of high school rank and grade point average at the end of the freshman year. The results, presented in Table X, indicate that high school rank was probably the best predictor.

TABLE X¹⁰

COMPARISON OF HIGH SCHOOL RANK AND GRADE POINT AVERAGE AT THE END OF THE FRESHMAN YEAR IN COLLEGE

High School Rank According to Quartiles	Grade Point Average Freshman Year in College
First Quartile Students	1.65
Second Quartile Students	1.06
Third Quartile Students	.75
Fourth Quartile Students	.62

(A equals 3 points; B equals 2 points; C equals 1 point; D and F equal 0 points)

Frick's study¹¹ yielded a correlation of .48 between grades and the ACE. It is interesting to note that with the addition of the Minnesota

¹⁰ Ibid., p. 23.

¹¹ J. W. Frick, "Improving the Prediction of Academic Use of the Minnesota Multiphasic Personality Inventory," Journal of Applied Psychology, XXXIX (1955), 49-52.

Multiphasic Personality Inventory, he obtained a multiple R of .64.

In their study of male veterans and non-veterans, Frederiksen and Schrader¹² found that high school marks correlated best with the first year college grades yielding an r of .57; ACE fell slightly lower in predictive value with an r of .47. When combined, these measures produced a multiple R of .60 for veterans and .68 for non-veterans.

Stone¹³, in 1954, reported the results of a predictive study at Brigham Young University and again found high school grades to be a better predictor than the ACE or the Cooperative General Culture Test.

Berdie¹⁴ reported a correlation coefficient of .56 between high school percentile rank and freshman honor point ratio for engineering students and a correlation of .21 between ACE total score and honor point ratio, again concluding that high school rank was the better single predictor.

Evenson and Smith¹⁵ obtained a correlation coefficient of only .29 between total SCAT and first year university average at the University of Alberta, Canada. They concluded that their study gave no evidence that

12 N. Frederiksen and W. B. Schrader, "American Council on Education Psychological Examination and High School Standing as Predictors of College Success," Journal of Applied Psychology, XXXVI (1952), 261-65.

13 Joics B. Stone, "Differential Prediction of Academic Success at Brigham Young University," Journal of Applied Psychology, XXXVIII (1954), 109-110.

14 Ralph F. Berdie, "Predicting of College Achievement and Satisfaction," Journal of Applied Psychology, XXVIII (1944), 239-45.

15 A. B. Evenson and D. E. Smith, "Study of Matriculation in Alberta," Alberta Journal of Educational Research, IV (1958), 67-83.

the SCAT would present a better predictive index than would high school rank.

The Nelson-Denny Reading Test

The Nelson-Denny Reading Test is an achievement test designed to predict probable success in college, to section incoming classes, and to aid in the diagnosis of students' difficulties.

Vineyard and Massey¹⁶, using the Nelson-Denny Reading Test as a measure of vocabulary and speed of paragraph comprehension, the Differential Aptitude Test as a measure of spelling ability, and the American Council on Education Psychological Examination as a measure of intelligence, found these measures related to college grade average as reflected in Table XI.

TABLE XI¹⁷

RELATIONSHIP OF ALL VARIABLES WITH COLLEGE GRADE AVERAGE

Variable	r
Vocabulary	.51
Speed of Paragraph Comprehension	.43
Spelling	.56
Intelligence	.52

They hypothesized that the linguistic skills were thoroughly saturated with the common factor of intelligence. They, therefore,

¹⁶ Edwin E. Vineyard and Harold W. Massey, "Interrelationships of Certain Linguistic Skills and Their Relationship with Scholastic Achievement with Intelligence Ruled Constant," Journal of Educational Psychology, XLVIII (1957), 279-286.

¹⁷ Ibid., p. 283.

attempted to rule intelligence constant and computed a partial r between linguistic skills and grade average. See Table XII.

TABLE XII¹⁸

RELATIONSHIP BETWEEN LINGUISTIC SKILLS AND GRADE AVERAGE WITH INTELLIGENCE RULED CONSTANT

Linguistic Skill	Partial r
Vocabulary	.25
Speed of Paragraph Comprehension	.14
Spelling	.37

According to the results reflected in Table XII, there still remains a significant relationship between grade average and the linguistic skills of vocabulary and spelling. It is their opinion that emphasis on these skills offer definite possibilities for improvement in scholarship. However, a different picture is presented by speed of comprehension. When intelligence is ruled constant, the relationship between grade average and this skill is reduced so greatly that it is no longer significant. They, therefore, concluded that emphasis on speed of paragraph comprehension would do little to improve scholarship.

Anderson and Dearborn¹⁹ conducted an investigation to determine whether a significant relationship existed between reading ability and

¹⁸ Ibid.

¹⁹ Irving H. Anderson and Walter F. Dearborn, "Reading Ability as Related to College Achievement," Journal of Psychology, XI (1941), 387-396.

achievement with intelligence ruled constant. They used the Scholastic Aptitude Test for a measurement of intelligence and course marks for a measurement of scholarship. For measurements of reading they used a battery of reading tests composed of the Nelson-Denny Reading Test, the Iowa Silent Reading Test, and the Whipple Reading Test. They tested 68 pairs of Harvard University freshmen, matching students who were equal in intelligence but differed in scholarship.

Table XIII reflects the mean scores of the academically higher and lower students for combined course work. The t values of difference between the means are also shown.

TABLE XIII²⁰

MEAN SCORES OF THE ACADEMICALLY HIGHER AND LOWER STUDENTS ON THE SCHOLASTIC ABILITY TEST AND EACH MEASURE OF READING ABILITY FOR COMBINED COURSES (HISTORY, GOVERNMENT, ENGLISH, ECONOMICS) AND FISHER'S t VALUES OF DIFFERENCE BETWEEN THE MEANS OF THESE TWO GROUPS

Measure	Number of Pairs	Mean of Academically Higher Students	Mean of Academically Lower Students	Fisher's t Value
Scholastic Aptitude Test	68	588.1	590.0	3.97 *
Nelson-Denny Raw Score	66	57.6	53.5	3.79 *
Nelson-Denny Accuracy	66	85.9	82.2	2.80 *
Iowa Raw Score	66	55.2	52.1	1.93
Iowa Accuracy	66	90.5	89.2	1.23
Iowa Rate of Reading	66	35.9	34.0	1.38
Whipple Raw Score	37	13.9	13.2	0.82

* Significant at the one per cent level.

²⁰ Ibid., p. 391.

According to the results of this investigation the authors concluded that the Nelson-Denny differentiates more effectively than do the other reading tests. They further concluded that the better students cover more material and read more accurately.

Traxler²¹ found a reliability coefficient of .88 between different forms of the Nelson-Denny Reading Test when administered a year apart. The Nelson-Denny correlated .78 with the Iowa Advanced Reading Test, Form B. Held²² found the Nelson-Denny to discriminate fairly well between good, fair, and inadequate preparation for the study of freshman college English. Table XIV reflects the mean scores of three groups in relation to Nelson-Denny Reading scores and corresponding grades in college freshman English.

TABLE XIV²³

THE MEAN SCORE FOR THREE GROUPS IN RELATION TO THE NELSON-DENNY READING SCORES AND CORRESPONDING GRADES IN COLLEGE FRESHMAN ENGLISH

Group	College Grade	Number	Mean Score
Group I	A and B	33	111.60
Group II	C	52	95.85
Group III	D and F	26	84.05

²¹ Arthur E. Traxler, "One Reading Test Serves the Purpose," Clearing House, XIV (1940), 419-21.

²² Omar C. Held, "The Nelson-Denny Reading Test as an English Placement Test," School and Society, XLIX (1939), p. 64.

²³ Ibid.

According to Garrett's survey²⁴, the Nelson-Denny Reading Test gives a good indication of probable college success.

The Essential High School Content Battery (EHSCB)

The Essential High School Content Battery is intended to be a comprehensive battery of high school achievement tests, including the areas of mathematics, science, social studies, and English. In reviewing the literature, only one investigation could be located which included the Essential High School Content Battery. Consequently, until other studies are available, caution should be exercised in the use of the battery for counseling purposes.

Dolansky²⁵ conducted an investigation using the EHSCB and the Terman-McNemar Test of Mental Abilities as predictors of college success. Grade points were obtained at the end of the freshman year in college for 240 students from three colleges. All of these students had taken the EHSCB during their senior year in high school. The correlation coefficients obtained are presented in Tables XV and XVI.

In order to compute a multiple R, intercorrelations were determined among the four subtests of the EHSCB and the Terman-McNemar Test. Table XVII reflects these intercorrelations. The obtained multiple R's were .60 for University X, .63 for University Y, and .57 for University Z.

Dolansky reminds the reader that the multiple R of .60 gives only about twenty per cent improvement over pure chance prediction. Consequently, there is room for much greater accuracy of prediction.

²⁴ Harley F. Garrett, "Review and Interpretation of Investigations of Factors Related to Scholastic Success in College of Arts and Science," Journal of Experimental Education, XVIII (1949), 91-131.

²⁵ Marie P. Dolansky, "Essential High School Content Battery as a Predictor of College Success," Journal of Educational Psychology, XLIV (1953), 361-365.

TABLE XV²⁶

CORRELATION COEFFICIENTS ESTIMATING RELATIONSHIP BETWEEN THE FOUR SUBTESTS OF THE EHSCB AND GRADE POINT AVERAGE IN UNIVERSITIES X, Y, AND Z

Variable	UNIVERSITIES		
	X (N = 43)	Y (N = 73)	Z (N = 118)
Mathematics	.18	.51	.45
Science	.31	.57	.51
Social Studies	.40	.50	.38
English	.57	.44	.43

TABLE XVI²⁷

CORRELATION COEFFICIENTS ESTIMATING RELATIONSHIP BETWEEN THE TERMAN-McNEMAR TEST OF MENTAL ABILITIES IQ'S AND GRADE POINT AVERAGE

University	Coefficient	Mean IQ	S.D. of IQ
University X	.44	116.36	12.56
University Y	.45	116.60	11.06
University Z	.40	114.12	9.88

TABLE XVII²⁸

INTERCORRELATIONS AMONG THE FOUR SUBTESTS OF THE EHSCB AND THE TERMAN-McNEMAR TEST OF MENTAL ABILITIES

	Mathematics	Science	Social Studies	English	Terman- McNemar IQ
Mathematics		.58	.48	.43	.55
Science			.65	.52	.65
Social Studies				.54	.64
English					.63
Terman-McNemar IQ					

²⁶ Ibid., p. 363.

²⁷ Ibid., p. 364.

²⁸ Ibid.

The Minnesota Teacher Attitude Inventory (MTAI)

The basic approach of this inventory seems similar to the Strong Vocational Interest Blank. Its aim is to measure the attitudes of teachers toward pupil-teacher relationships. The Minnesota Teacher Attitude Inventory assumes that teachers' attitudes are the result of interaction of many factors, such as academic and social intelligence, general knowledge and abilities, social skills, and personality traits. It further assumes that these attitudes can provide an index of prediction to the type of social climate a teacher will maintain in the classroom.

Several studies have been conducted using this inventory, a few of which will be cited subsequently. These studies have to do with the prediction of what type of social climate a teacher will maintain in the classroom. As has been previously mentioned, no existing studies were found dealing with this inventory as a predictor of academic success. However, the following studies will give some insight into the type of work conducted on the inventory.

When correlated with three criteria (pupils' ratings of teachers, experts' ratings of teachers, and principals' ratings of teachers), the validity coefficients have been between .46 and .60.

In Fuller's experiment²⁹ with 74 senior women in the University of Minnesota College of Education, he found that scores of the MTAI show no significant relationship to self-ratings, supervised ratings, honor points,

²⁹ Elizabeth M. Fuller, "Use of Teacher Pupil Attitudes, Self-Ratings, and Measurements of Ability in the Preservice Selection of Nursery School-Kindergarten-Primary Teachers," Journal of Educational Research, XLIV (1951), 675-686.

intelligence test scores, or rank order lists of student teachers as measured by the Spearman Rank Difference correlation method.

Rogers and Smith³⁰ of Syracuse University administered the MTAI to 64 students at the beginning of an intensive teacher training program and again during the last days of the six weeks program. They obtained a mean score of 70.13 for the second administration as compared to 43.84 for the first. This would seem to indicate that professed attitude can be modified during an intensive training period. However, there is no evidence of the permanence of this modification, or, for that matter, the desirability of such modification.

This brings to mind the caution emphasized by Fishman³¹ and Rogers³² that the test is still in the experimental stage and should be used as such. It is interesting to note that students tend to score higher on a second administration after more intensive training, that students working on secondary education tend to score lower than those working in early childhood and elementary education, and that experienced teachers tend to score lower than students.

Perhaps, subsequent validity studies will support the predictive value of the inventory as to the social climate a teacher will maintain. However,

³⁰ Vincent R. Rogers and James A. Smith, "Professional Attitudes of Students in an Intensive Teacher Training Program," Elementary School Journal, LVII (1956), 100-101.

³¹ Joshua A. Fishman, "Minnesota Teacher Attitude Inventory in an American Minority Group Setting," Journal of Educational Psychology, XLVIII (1957), 41-51.

³² Rogers and Smith, p. 101.

based on the available evidence, there seems to be little likelihood that an inventory such as this one would lend any value to a battery designed to predict probable success in the academic program outlined for students who are training for the teaching fields.

CHAPTER V

PRESENTATION AND ANALYSIS OF RESULTS

This chapter is devoted to the presentation and analysis of the results obtained in the present study. The test intercorrelation matrices will be presented. These include the correlation coefficients obtained between scores on each test of the SCAT, Nelson-Denny Reading Test, the Essential High School Content Battery, and the attitude score of the Minnesota Teacher Attitude Inventory, in addition to the correlation between each test and the over-all college grade point average at the end of the freshman year in college.

By use of the Wherry-Doolittle Test Selection Method, the multiple correlation between the optimum combination of tests and the over-all grade point average at the end of the freshman year will be noted, and the regression equations for prediction of this criterion from known test scores will be given.

Test Intercorrelation Matrices

The International Business Machine 650 computer was utilized to compute the test intercorrelation matrices. Consequently, after test scores and over-all grade point averages at the end of the freshman year were obtained for all members of Group I (male) and II (female), the data were transferred to IBM cards and processed through the I.B.M. 650 computer. Computations included the means and standard deviations of each test and of the criterion, the summations, summations of squares, summations of cross-products for

each test of the battery, the zero-order correlations between all tests of the battery, and the zero-order correlations between each test and the grade point average criterion. Intercorrelation data for Groups I and II are presented in Tables XVIII and XIX, respectively.

The highest intercorrelations for Group I (males) appear to be the subtests within each test. Verbal score of the SCAT correlates with quantitative SCAT score and total SCAT .96 and .99, respectively, and quantitative SCAT score correlates with total SCAT score .99. The Nelson-Denny vocabulary score correlates with paragraph meaning and total scores of the Nelson-Denny .78 and .94, respectively, and paragraph meaning with total Nelson-Denny score .94. The English test of the EHSCB correlates with the science test and social studies test of the same battery .78 and .73, respectively. This would seem to indicate a high degree of overlap within the tests, i.e., the tests appear to be measuring the same or similar abilities.

The lowest correlations for Group I are as follows: SCAT verbal with Nelson-Denny vocabulary, EHSCB mathematics, science, social studies, and the Minnesota Attitude score .27, .21, .24, .16, and .23, respectively; quantitative SCAT score with Nelson-Denny vocabulary, total score, EHSCB science, social studies, and the attitude score .17, .29, .24, .16, and .23; SCAT total score with Nelson-Denny vocabulary, EHSCB mathematics, science, social studies, and the Minnesota Attitude score .25, .29, .26, .18, and .23; and the Minnesota Attitude score with Nelson-Denny vocabulary, paragraph meaning and total, EHSCB mathematics, science, social studies, and English .09, .16, .16, .09, .03, .04, and .01, respectively. The size of these coefficients indicates little relationship between abilities tested.

TABLE XVIII

INTERCORRELATIONS OF ELEVEN TESTS (GROUP I): VERBAL, QUANTITATIVE, TOTAL (SCAT); VOCABULARY, PARAGRAPH MEANING, TOTAL (NELSON-DENNY READING TEST); MATHEMATICS, SCIENCE, SOCIAL STUDIES, ENGLISH (ESSENTIAL HIGH SCHOOL CONTENT BATTERY); ATTITUDE (MINNESOTA TEACHER ATTITUDE INVENTORY)
N = 31 males

Tests	1	2	3	4	5	6	7	8	9	10	11
1. Verbal		.96	.99	.27	.46	.39	.21	.24	.16	.33	.23
2. Quantitative			.99	.17	.36	.29	.31	.24	.16	.34	.23
3. Total (SCAT)				.25	.44	.37	.29	.26	.18	.37	.23
4. Vocabulary					.78	.94	.59	.65	.65	.69	.09
5. Paragraph Meaning						.94	.54	.66	.55	.68	.16
6. Total (Nelson-Denny)							.61	.67	.63	.71	.16
7. Mathematics								.64	.59	.68	.09
8. Science									.69	.78	.03
9. Social Studies										.73	.04
10. English (EHSCB)											.01
11. Attitude (MTAI)											

TABLE XIX

INTERCORRELATIONS OF ELEVEN TESTS (GROUP II): VERBAL, QUANTITATIVE, TOTAL (SCAT); VOCABULARY, PARAGRAPH MEANING, TOTAL (NELSON-DENNY READING TEST); MATHEMATICS, SCIENCE, SOCIAL STUDIES, ENGLISH (ESSENTIAL HIGH SCHOOL CONTENT BATTERY); ATTITUDE (MINNESOTA TEACHER ATTITUDE INVENTORY)

N = 75 females

Tests	1	2	3	4	5	6	7	8	9	10	11
1. Verbal		.52	.72	.73	.69	.74	.48	.73	.67	.68	.29
2. Quantitative			.56	.37	.33	.35	.55	.46	.49	.56	.13
3. Total (SCAT)				.55	.56	.58	.55	.54	.65	.61	.18
4. Vocabulary					.68	.91	.54	.66	.52	.56	.27
5. Paragraph Meaning						.89	.49	.69	.57	.58	.39
6. Total (Nelson-Denny)							.52	.69	.56	.59	.36
7. Mathematics								.47	.47	.54	.00
8. Science									.74	.69	.31
9. Social Studies										.62	.29
10. English (EHSCB)											.26
11. Attitude (MTAI)											

The remainder of the intercorrelations fall within the range of .40 to .70 and indicate a moderate but not too high relationship.

The highest intercorrelations for Group II (females) appear to be between SCAT verbal and SCAT total, Nelson-Denny vocabulary and total, EHSCB science test, .72, .73, .74, and .73, respectively; total Nelson-Denny with vocabulary and paragraph meaning of the same test .91 and .89, respectively, and science and social studies of the EHSCB .74.

The lowest intercorrelations in Group II are found between the Minnesota Teacher Attitude score and quantitative and total SCAT, Nelson-Denny vocabulary, mathematics, science, social studies, and English of EHSCB .13, .18, .27, .002, .31, .29, and .26. This would seem to indicate that the attitude score was measuring some trait almost negligibly related to the other abilities represented in the battery.

It is interesting to note that the highest intercorrelations in Group II appear to be between those tests purporting to measure some type of verbal or language ability, i.e., Tests 1, 3, 4, 5, and 6. High intercorrelations appear between only two other sets, Tests 1 and 8, and Tests 8 and 9. It is reasonable to assume here that verbal ability would greatly influence scores and be related to those in social studies and science courses.

Correlation of Tests and Over-All Grade Average at the End of Freshman Year

Zero-order coefficients of correlation were computed between scores made by members of Groups I and II on each of the eleven tests given at the beginning of the freshman year in college and the subsequent grade average made by members of the same groups at the end of the freshman year in college. The relationships reflected between test scores and the over-all grade averages are shown in Table XX.

TABLE XX

CORRELATIONS OF SCORES ON ELEVEN TESTS WITH CRITERION OF GRADE POINT AVERAGE
AT THE END OF THE FRESHMAN YEAR IN COLLEGE

Tests	Group I	Group II
	N = 31 males r	N = 75 females r
1. Verbal	.08	.53
2. Quantitative	.19	.43
3. Total (SCAT)	.16	.54
4. Vocabulary	.39	.38
5. Paragraph Meaning	.26	.43
6. Total (Nelson-Denny)	.35	.42
7. Mathematics	.58	.48
8. Science	.33	.40
9. Social Studies	.44	.52
10. English (EHSCB)	.37	.59
11. Attitude (MTAI)	.02	.10

Referring to Group I, Table XX, one may discover that the mathematics and social studies tests of the EHSCB appear to have a moderately high predictive value, yielding a coefficient of .58 and .44, respectively. The Nelson-Denny vocabulary, paragraph meaning, and total scores and the science and English tests of the EHSCB reflect a definite but small relationship with coefficients of .39, .26, .35, .33, and .37. The verbal, quantitative, and total SCAT scores and the Minnesota Teacher Attitude score appear to have a slight, almost negligible predictive value yielding r 's of only .08, .19, and .16, respectively.

In the case of Group II, Table XX, the verbal, quantitative, and total SCAT scores, the paragraph meaning and total Nelson-Denny scores and the mathematics, science, social studies, and English tests of EHSCB appear to have a moderately high predictive value, yielding correlation coefficients of .53, .43, .54, .43, .42, .48, .40, .52, and .59, respectively. The vocabulary of the Nelson-Denny test yields a definite but small relationship with a correlation coefficient of .38, and the Minnesota Teacher Attitude Test appears to have a slight, almost negligible predictive value, yielding an r of only .10.

It is interesting to note that the mathematics and social studies tests of the EHSCB appear to have a fairly high predictive value for both groups, whereas the Minnesota Teacher Attitude score has little, if any, predictive value in either group.

The question at this point is which combination of the eleven scores will give us the highest possible correlation for prediction purposes. The Wherry-Doolittle Test Selection Method was utilized to analytically select and add one test at a time until a maximum \bar{R} was obtained.

For the male subjects in Group I, Table XX shows the mathematics test of the EHSCB to correlate .58 with the criterion. The next highest correlation is the social studies test of the same battery yielding an r of .44. When the social studies test was added to the mathematics test, the resulting multiple was actually lowered to .57. Consequently, it appears that the addition of another test does nothing to enhance the predictive value of the first test. This outcome renders it impossible to proceed further in the selection of an optimum test combination and, consequently, makes prediction highly unstable.

For the female subjects of Group II, Table XX reflects the English test of the EHSCB to correlate .59 with the criterion. When the total SCAT score is added, the \bar{R}^1 is raised to .62; when the mathematics test of the EHSCB is added, the \bar{R} is again elevated to .63. The addition of the test with the next highest correlation, the social studies test of the EHSCB, lowers the \bar{R} slightly and is, therefore, not included in the optimum combination of tests. The combination then of the English test of the EHSCB, the total SCAT score, and the mathematics test of the EHSCB, according to this study, gives the best predictive index for the criterion of Group II.

In order to predict the criterion with the greatest accuracy from the scores obtained on the tests, the multiple \bar{R} is used to develop a regression equation. For Group II the regression equation in raw score form is as follows:

$$X_c = .001 (X_7) + .007 (X_3) + .01 (X_{10}) - 1.785$$

The values .001, .007, and .01 are the weights by which the scores of

¹ When using the Wherry-Doolittle Method of Test Selection a shrunken \bar{R} is obtained. This corrects for the operation of chance errors and is a better estimation of the correlation in the true population from which the sample was drawn.

Tests 7, 3, and 10 are multiplied. The X represents the raw score of the particular test. These products and the constant (-1.785) are summed, thereby yielding the predicted over-all grade average, X_c . The standard error of estimate associated with the above equation is .14, the probability being that approximately sixty-eight times in a hundred the actual grade average will fall within the interval of the predicted grade plus or minus the standard error of estimate.

The selected optimum combination of tests accounts for forty-two per cent of the variance of the criterion. Also, the three tests, the EHSCB mathematics, total SCAT, and the EHSCB English contribute seven per cent, eleven per cent, and twenty-two per cent, respectively, to the variance of the criterion. The remaining fifty-eight per cent of the variance of the criterion must be attributed to factors not measured in this problem.

For an over-all look at the outcome of the problem, Table XXI reflects the mean grade point averages and mean scores for each test in the battery, the respective standard deviations for each test, and the t values. On all of the verbal tests the females scored significantly higher than the males. On the Nelson-Denny vocabulary and total scores and the EHSCB English scores the t values were statistically significant at the one per cent level. On the attitude score and the criterion the t values were also statistically significant at the one per cent level. This study, therefore, indicates a true sex difference on scores obtained by males and females. The difference in the numbers included in each group may account for some of the existing difference; however, it seems doubtful that it would account for as much as is reflected.

TABLE XXI

MEANS AND STANDARD DEVIATIONS ON ELEVEN TESTS AND THE CRITERION FOR GROUPS I AND II AND THE *t* VALUES REFLECTING THE DIFFERENCE BETWEEN MEANS

Variable	Group I N = 31 males		Group II N = 75 females		t values
	Mean	S.D.	Mean	S.D.	
Criterion	1.98	.64	2.62	.55	4.74 **
Tests					
1. Verbal	278.1	43.9	294.0	12.6	1.95 *
2. Quantitative	286.9	47.4	297.0	18.0	1.13
3. Total (SCAT)	283.8	37.2	296.0	16.7	1.50
4. Vocabulary	24.3	9.6	34.9	12.2	4.71 **
5. Paragraph Meaning	34.2	11.1	40.2	12.2	2.42 *
6. Total (Nelson-Denny)	58.2	19.7	74.3	23.3	3.57 **
7. Mathematics	28.7	10.5	26.9	7.8	.75
8. Science	39.0	11.5	38.5	9.4	.21
9. Social Studies	38.5	11.8	41.5	11.0	1.20
10. English (EHSCB)	165.2	24.2	190.6	20.3	5.07 **
11. Attitude (MTAI)	96.4	23.7	121.0	26.9	4.62 **

* Significant at the five per cent level of confidence.

** Significant at the one per cent level of confidence.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

This chapter is devoted to a summary of the stated objectives under investigation, conclusions drawn therefrom, and recommendations for further research in the same or similar areas.

Review of Objectives

The problem of this study was to determine the extent to which the first year's training in the College of Education can be predicted by scores obtained on the Cooperative School and College Ability Tests (SCAT), the Nelson-Denny Reading Test, the Essential High School Content Battery (EHSCB), and the Minnesota Teacher Attitude Inventory (MTAI).

The following information was obtained from each of two groups:

Group I (males) and Group II (females):

1. The correlation coefficient between scores on the SCAT and the criterion, i.e., college freshman over-all grade point average.
2. The correlation coefficient between scores on the Nelson-Denny Reading Test and the criterion.
3. The correlation coefficient between scores on each subtest of the Essential High School Content Battery and the criterion.
4. The correlation coefficient between scores on the Minnesota Teacher Attitude Inventory and the criterion.

5. The coefficient of multiple correlation between the optimum combination of the above tests and the criterion.

6. The multiple regression equation for prediction of the criterion from known scores on the above tests.

Summary of Results

1. For Group I the correlation coefficients between scores on the SCAT and the criterion were as follows: verbal, .08; quantitative, .19; and total, .16. Similar correlations for Group II were: verbal, .53; quantitative, .43; and total, .54.

2. For Group I the correlation coefficients between scores on the Nelson-Denny Reading Test and the criterion were as follows: vocabulary, .39; paragraph meaning, .26; and total, .35. Similar correlations for Group II were: vocabulary, .38; paragraph meaning .43; and total, .42.

3. For Group I the correlation coefficients between scores on the Essential High School Content Battery and the criterion were: mathematics, .58; science, .33; social studies, .44; and English, .37. Similar correlations for Group II were: mathematics, .48; science, .40; social studies, .52; and English, .59.

4. For Group I the correlation coefficient between scores on the Minnesota Teacher Attitude Inventory and the criterion were .02 for Group I and .10 for Group II.

5. For Group I only one test, the mathematics test of the EHSCB, yielded a high enough correlation to be of much predictive value. When the test with the next highest relationship was added, the \bar{R} was lowered. Since the accuracy of prediction is considered unstable when using only

one test, no regression equation was developed. For Group II, the optimum combination of tests selected by the Wherry-Doolittle method included the English and mathematics test of the EHSCB and the total SCAT score. This combination yielded a multiple \bar{R} of .63 with the criterion.

6. A regression equation for the purpose of predicting the over-all grade point average criterion from known test scores was developed for Group II. The standard error of estimate of the equation was .14 grade points.

Conclusions

The following represents conclusions and generalizations which appear to be warranted on the basis of data obtained in this study. Needless to say, these statements are restricted to freshmen students in the College of Education and should not be taken as necessarily valid for any other group on the campus of the Oklahoma State University. Further caution is added since this study does not include a cross-validation study and, consequently, should be taken only as a foundation for more extensive study in the area.

1. Scores on the SCAT obtained when a girl enters her freshman year in the College of Education seem to offer the counselor helpful information regarding the success in the academic program through the freshman year. However, scores on the SCAT seem to have so little relationship with grade point average for boys that they offer the counselor little or no help regarding success in the academic program through the freshman year.

2. Scores on the Nelson-Denny Reading Test obtained when a student enters his freshman year in the College of Education seem to offer the counselor some help regarding success in the academic program through the

freshman year. However, since here the highest relationship for the females is .43 and for the males, .39, it is doubtful if these scores would be of much value unless a good many other factors were taken into consideration.

3. Scores on the Essential High School Content Battery seem to offer the counselor helpful information regarding the success of a girl through her academic program of the freshman year. The mathematics test and the social studies test appear to offer more help than the other two tests of the battery in predicting the success of a boy through his academic program of the freshman year.

4. The scores on the Minnesota Teacher Attitude Inventory appear to be of little, if any, help to the counselor. The relationships are so slight as to appear negligible.

5. The combination of the total SCAT score, the English test, and the mathematics test of the Essential High School Content Battery appear to be better predictors of success of women in the academic program as outlined by the College of Education than do any of the tests taken singly.

6. In order to predict the criterion with the greatest accuracy from scores obtained on the optimum combination of tests for the girls, the following regression equation was developed:

$$X_c = .001 (X_7) + .007 (X_3) + .01 (X_{10}) - 1.785$$

The values .001, .007, and .01 are weights by which the scores of Test 7, the mathematics test of the Essential High School Content Battery; Test 3, total SCAT; and Test 10, English test of the Essential High School Content Battery, are multiplied. The X represents the raw score of the particular test. These products and the constant (-1.785) are summed, thereby yielding the predicted over-all grade point average, X_c . The standard error of estimate associated with this equation is plus or minus .14.

Recommendations

Several of the tests under investigation in this study are relatively new and present a good problem for further research. The validity studies available on the SCAT and the Essential High School Content Battery are particularly scarce, and, since they do seem to bear a moderate to moderately high relationship with the criterion, further investigation of this general type would do much to enhance their value as predictive indices of probable success in college.

As stated previously, this study should be considered groundwork only. Because of the size of the study groups, particularly Group I, the results obtained herein may be somewhat unstable. The following suggestions are submitted as possible ways in which this investigation could be improved or extended.

1. Since Group I constitutes such a small sample and may not be representative of the freshmen men in the College of Education, it is recommended that the complete study be repeated for this group based on a sample of not less than seventy-five men.

2. If the results of Group I make possible the selection of an optimum combination of tests for the prediction of grade point average, it is recommended that at a later date a cross-validation study be conducted to determine the stability of this investigation.

3. It is further recommended that a study be initiated to investigate possible reasons for the existing differences between sexes. In addition to the aptitude and achievement tests, some concern might be centered around values and attitudes, personality characteristics, and methods of study.

BIBLIOGRAPHY

- Anderson, Irvin H. and Walter F. Dearborn. "Reading Ability as Related to College Achievement." Journal of Psychology, XI (1941), 387-396.
- Berdie, Ralph F. "Prediction of College Achievement and Satisfaction." Journal of Applied Psychology, XXVIII (June, 1944), 234-245.
- Blum, Milton L. and Benjamin Balinsky. Counseling and Psychology. New York: Prentice-Hall, 1951.
- Boring, Edwin G. History of Experimental Psychology. New York: Appleton-Century-Crofts, Inc., 1950.
- Buros, Oscar K. Fourth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1953.
- _____. Fifth Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1959.
- _____. Nineteen Forty Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1941.
- Cook, Walter W., Carroll H. Leeds, and Robert Callis. The Minnesota Teacher Attitude Inventory Manual. New York: The Psychological Corporation, 1951.
- Cooperative School and College Ability Tests: Manual for Interpreting Scores. Princeton, New Jersey: Educational Testing Service, 1957.
- Cooperative School and College Ability Tests: Technical Manual. Princeton, Princeton, New Jersey: Educational Testing Service, 1957.
- Crawford, A. B. and P. S. Burnham. Forecasting College Achievement. New Haven: Yale University Press, 1946.
- Dolansky, Marie P. "Essential High School Content Battery as a Predictor of College Success." Journal of Educational Psychology, XLIV (October, 1953), 361-365.
- Downie, N. M. Fundamentals of Measurement. New York: Oxford University Press, 1958.
- Evenson, A. B. and D. E. Smith. "Study of Matriculation in Alberta." Alberta Journal of Educational Research, IV (March, 1958), 67-83.

- Fishman, Joshua. "Minnesota Teacher Attitude Inventory in an American Minority Group Setting." Journal of Educational Psychology, XLVIII (January, 1957), 41-51.
- Frederiksen, N. and W. B. Schrader. "American Council on Education Psychological Examination and High School Standing as Predictors of College Success." Journal of Applied Psychology, XXXVI (August, 1952), 261-265.
- Frick, J. W. "Improving the Prediction of Academic Use of the Minnesota Multiphasic Personality Inventory." Journal of Applied Psychology, XXXIX (February, 1955), 49-52.
- Fuller, Elizabeth M. "Use of Teacher Pupil Attitudes, Self Ratings and Measurements of Ability in the Preservice Selection of Nursery School-Kindergarten-Primary Teachers." Journal of Educational Research, XLIV (May, 1951), 675-686.
- Garrett, Harley F. "Review and Interpretation of Investigations of Factors Related to Scholastic Success in College of Arts and Science." Journal of Experimental Education, XVIII (December, 1949), 91-131.
- Garrett, Henry E. Statistics in Psychology and Education. New York: Longmans, Green and Company, Inc., 1958.
- Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill Book Company, 1956.
- Harry, David P., and Walter N. Durost. Essential High School Content Battery Manual of Directions. New York: World Book Company, 1951.
- Held, Omar C. "Nelson-Denny Reading Test as an English Placement Test." School and Society, XLIX (January, 1939), 64-65.
- Manual, Herschel T. "Aptitude Tests for College Admission." Fourteenth Yearbook: National Council on Measurements Used in Education. New Jersey, 1957.
- Nelson, M. J. and E. C. Denny. Nelson-Denny Reading Test. New York: Houghton Mifflin Company, 1929.
- North, Robert D. "Comparison of SCAT and ACE." Educational Records Bulletin, LXVII (February, 1956), 65-72.
- _____. "Further Report on SCAT." Educational Records Bulletin, LXIX (February, 1957), 60-62.
- Rogers, Vincent R. and James H. Smith. "Professional Attitudes of Students in an Intensive Teacher Training Program." Elementary School Journal, LVII (November, 1956), 100-101.

- Stone, Joics B. "Differential Prediction of Academic Success at Brigham Young University." Journal of Applied Psychology, XXXVIII (April, 1954), 109-110.
- Thorndike, Robert L. and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc., 1955.
- Thurston, Louis Leon. Measurement of Values. Chicago: University of Chicago Press, 1959.
- Traxler, Arthur E., Margaret Selover, and Agatha Townsend. Introduction to Testing and Use of Test Results in Public Schools. New York: Harper, 1953.
- Traxler, Arthur E. "One Reading Test Serves the Purpose." Clearing House, XIV (March, 1940), 419-421.
- _____. "Should SCAT Scat ACE?" Educational Records Bulletin, LXVII (February, 1956), 51-63.
- _____. Techniques of Guidance. New York: Harper, 1945.
- Vineyard, Edwin E. and Harold W. Massey. "Interrelationships of Certain Linguistic Skills and Their Relationship with Scholastic Achievement with Intelligence Ruled Constant." Journal of Educational Psychology, XLVIII (May, 1957), 279-286.

VITA

JO STRIBLING DORRIS

Candidate for the Degree of

Master of Science

Thesis: THE RELATIONSHIP OF A BATTERY OF SELECTED TESTS TO SUCCESS IN THE FIRST YEAR PROGRAM IN THE COLLEGE OF EDUCATION AT THE OKLAHOMA STATE UNIVERSITY

Major Field: Psychology

Biographical:

Personal Data: Born at Walters, Oklahoma, March 4, 1932.

Undergraduate Study: Oklahoma College for Women, Chickasha Oklahoma, 1950-1954.

Graduate Study: Oklahoma State University, Stillwater, Oklahoma, 1958-1960.

Experience: Legal Secretary for Phillips Petroleum Company, Legal Department, Bartlesville, Oklahoma, 1954-1955; Claims Secretary for Phillips Petroleum Company, Tax, Insurance and Claims Department, Oklahoma City, Oklahoma, 1955-1958; Night Instructor at Hills' Business College, Oklahoma City, Oklahoma, 1955-1958; Head Counselor at Stout Hall, Oklahoma State University, Stillwater, Oklahoma, 1958-1960.

Member of the Oklahoma State Psychological Association and Psi Chi.

Date of Final Examination: April, 1960