1962

Some California Test Bureau Measures Predictive of Ninth Grade and Eleventh Grade Academic Achievement

Daniel Joseph Hartigan

Loyola University Chicago

Recommended Citation
http://ecommons.luc.edu/luc_diss/640

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1962 Daniel Joseph Hartigan
SOME CALIFORNIA TEST BUREAU MEASURES PREDICTIVE
OF NINTH GRADE AND ELEVENTH GRADE
ACADEMIC ACHIEVEMENT

by

Rev. Daniel Joseph Hartigan, O.S.A.

A Thesis Submitted to the Faculty of the Graduate School
of Loyola University in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Education

June
1962
Daniel Joseph Hartigan was born in Boston, Massachusetts, June 15, 1917.

He was graduated from Boston Trade High School, Boston, Massachusetts, June, 1934, and from Villanova University, Villanova, Pennsylvania, June, 1942, with a degree of Bachelor of Arts in Liberal Arts. He was graduated from Catholic University of America, Washington, D.C., June 1946, with a degree of Master of Science in Mathematics; and from Loyola University, Chicago, Illinois, February, 1959, with the degree of Master of Education.

From 1946 to 1951 the author taught at St. Rita High School in the City of Chicago, Illinois. From 1951 to 1961 he taught and served in the administration of Mendel Catholic High School in the City of Chicago, Illinois. In 1961 he began his present assignment as professor in Mathematics at Tolentine College, Olympia Fields, Illinois. He began his present graduate studies at Loyola University in September, 1956.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I.</strong></td>
<td><strong>II.</strong></td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td><strong>REVIEW OF THE LITERATURE</strong></td>
</tr>
<tr>
<td></td>
<td><strong>A.</strong></td>
</tr>
<tr>
<td><strong>A.</strong> Purpose and justification of the study</td>
<td><strong>A.</strong> Survey of psychological data</td>
</tr>
<tr>
<td><strong>B.</strong> Scope and limits of the study</td>
<td><strong>1.</strong> Intelligence</td>
</tr>
<tr>
<td><strong>C.</strong> Methods to be used in the study</td>
<td><strong>2.</strong> Personality</td>
</tr>
<tr>
<td></td>
<td><strong>3.</strong> Interests</td>
</tr>
<tr>
<td></td>
<td><strong>4.</strong> Multiple-score tests of aptitudes</td>
</tr>
<tr>
<td></td>
<td><strong>5.</strong> Yearly grade averages</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>B.</strong> Predictive studies</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. GRADE EQUIVALENTS</td>
<td>57</td>
</tr>
<tr>
<td>II. NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST SCORE EQUIVALENTS</td>
<td>58</td>
</tr>
<tr>
<td>III. COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND FRESHMAN YEARLY GRADE AVERAGES</td>
<td>77</td>
</tr>
<tr>
<td>IV. COEFFICIENT OF CORRELATION BETWEEN TEACHERS' CHARACTER RATINGS AND FRESHMAN YEARLY GRADE AVERAGES</td>
<td>78</td>
</tr>
<tr>
<td>V. COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST TOTAL SCORES</td>
<td>79</td>
</tr>
<tr>
<td>VI. COEFFICIENT OF CORRELATION BETWEEN TEACHERS' CHARACTER RATING AND NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST TOTAL SCORES</td>
<td>80</td>
</tr>
<tr>
<td>VII. COEFFICIENT OF CORRELATION BETWEEN NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST SCORES AND FRESHMAN YEARLY GRADE AVERAGES</td>
<td>81</td>
</tr>
<tr>
<td>VIII. MATRIX OF INTERCORRELATIONS</td>
<td>83</td>
</tr>
<tr>
<td>IX. ELIMINATION OF PREDICTOR OF LEAST VALUE BY MEANS OF MULTIPLE CORRELATION OF SIX VARIABLES WITH YEARLY GRADE AVERAGE</td>
<td>85</td>
</tr>
<tr>
<td>X. ELIMINATION OF PREDICTOR OF LEAST VALUE BY MEANS OF MULTIPLE CORRELATION OF SIX VARIABLES WITH N.M.S.Q.T. TOTAL SCORES</td>
<td>88</td>
</tr>
<tr>
<td>XI. MATRIX OF INTERCORRELATION</td>
<td>90</td>
</tr>
<tr>
<td>XII. COEFFICIENT OF MULTIPLE CORRELATION BETWEEN FRESHMAN YEARLY GRADE AVERAGE AND A BATTERY OF PREDICTORS FOR THE &quot;CROSS-VALIDATION GROUP&quot;</td>
<td>92</td>
</tr>
<tr>
<td>XIII. COEFFICIENT OF MULTIPLE CORRELATION BETWEEN N.M.S.Q.T. TOTAL SCORES AND A BATTERY OF PREDICTORS FOR THE &quot;CROSS-VALIDATION GROUP&quot;</td>
<td>93</td>
</tr>
</tbody>
</table>
XIV. MATRIX OF INTERCORRELATION - FRESHMAN YEARLY GRADE AVERAGE • "CROSS-VALIDATION GROUP" ......................... 94

XV. T TEST FOR SIGNIFICANCE BETWEEN ORIGINAL GROUP AND FRESHMAN YEARLY GRADE AVERAGE "CROSS-VALIDATION GROUP" ............... 95

XVI. NATIONAL MERIT TOTAL SCORE MATRIX OF INTERCORRELATIONS - "CROSS VALIDATION GROUP" ................................. 96

XVII. T TEST FOR SIGNIFICANCE BETWEEN ORIGINAL GROUP AND N.M.S.Q.T. "CROSS-VALIDATION GROUP" ..................... 97

XVIII. F TEST BETWEEN ORIGINAL AND "CROSS-VALIDATION" SAMPLES FOR FRESHMAN YEARLY GRADE AVERAGE .............. 98

XIX. F TEST BETWEEN ORIGINAL AND "CROSS-VALIDATION" SAMPLES FOR TOTAL SCORE ON NATIONAL MERIT BATTERY .......... 99
CHAPTER I

INTRODUCTION

Traditionally in America, the learning experiences of elementary school children follow routines and regulations which have been similar to most children of like age or grade upon entering the high school. However, there are urgent and distinct individual differences which can no longer be satisfied by like training for all. For that reason, many educators agree that at the end of the eighth grade and about the age of fourteen, it is an appropriate and practical time, both in terms of individual psychological development and in administrative organization, to provide an adjusted educational program designed to meet an ever increasing variety of demands for instruction on the high school level. It is generally accepted that the growth of intellectual abilities, as reflected by a student's standardized test scores, has been stabilized at this age period to the extent that an eighth grade test is likely to prove almost as reliable as a 12th grade test, for purposes of evaluating innate individual potentials.1 Diagnostic and prognostic testing has become quite popular for the purpose of predicting an individual's strength and his potential success.

Particularly today, with the increased emphasis upon guidance and counseling at the secondary school level, it seems appropriate and even necessary to evaluate some of the instruments that have been and are still being used for evaluation of a student's assets or limitations. Especially in private schools, such as Mendel Catholic High School, where elaborate testing involves a fee, it seems to be almost a matter of conscience to have reasonable certitude that the tests are valid tests; and the evaluations and predictions that are given by the psychologist to the parents be defensible predictions.

John L. Homes urges that the school offer broad, functional guidance programs which will enable teachers and administrators to interpret each student's abilities, personality, interests, and achievement status. This should be done, he continues, in ways that will also enable each individual student to understand his abilities, interests and personality characteristics; to develop them as far as possible; to relate them to his life goals; and finally arrive at a state of complete and mature self-guidance, as a contributing citizen of a democratic society.²

A continuous, broad evaluation program includes a wide selection of standardized tests that yield meaningful information about many aspects of an individual's past, present and future status. The following areas are basic to such an evaluation program:³

³Ibid. 2.
1) Mental Maturity: A test of mental maturity is the best general predictor of the educational level at which each student is able to achieve. By the use of a mental maturity test, individuals of high promise can be identified early in their school life. Discrepancy between students' intelligence scores and their level of achievement has significance for the counselor. A student may have a high test score and a low academic achievement. This would indicate the presence of some causative factor other than lack of ability, such as poor habits of study, personality difficulties or other problems which the counselor should investigate if he is to apply a remedy. The reverse condition, that is, a high scholastic standing with a low intelligence score may also need attention. There are pupils who over-tax their ability and injure their health in attempting to secure high grades.

2) Personal and Social Adjustment: The importance of good personal and social adjustment to the mental health of an individual and to vocational success is obvious. The fact that there is difficulty in defining it and lack of agreement concerning the components, should not be an insuperable obstacle to its measurement. The more common assumption is that personality is a more or less definite entity made up largely of a number of generalized traits. The California Test of Personality has been structured along this basic assumption.

In the complex problem of making a vocational choice, an adequate appraisal of a student's personality characteristics is indispensable to the counselor and to the counselee. Teachers and counselors should make every effort to understand the personality characteristics of their students. Through observation and the skilful use of standardized personality tests, the teacher or counselor is better able to identify the student with difficulties
and to counsel him; to identify the individual whose failure or underachievement is due more to personal problems than to low ability or to lack of interest.

3) **Interests:** Measures of occupational interests are useful: a) in identifying unknown or unrecognized interests, b) in confirming stated interests, c) in identifying potential or actual conflicts between stated and tested interests, d) in identifying discrepancies between interests and aptitudes, e) in verifying the absence of well-defined interest patterns, and f) in identifying educational or vocational maladjustment due to inappropriate interests. Then, too, an interest inventory is a useful motivating device to get pupils thinking seriously and objectively about their occupational interests.

4) **Special Abilities and Aptitude Tests:** These are similar to tests of intelligence and achievement. The content of a multifactor aptitude battery may be a longer version of an intelligence test. An achievement test may be used as an aptitude test to predict future behavior in a particular subject or activity. Although individuals who make high scores on intelligence tests often make high scores on all the tests in an aptitude battery, most individuals do not have equally strong aptitudes in all areas. Nor are low scores on an aptitude battery always indicative of low intelligence. The results of ability and aptitude tests help to give a clearer and more specific picture of the students' strengths and weaknesses.

5) **Achievement:** With the possible exception of mental maturity, achievement is the most important area of appraisal in an evaluation program. The cumulative evaluation of a class of students by competent teachers, - all
things being equal - should be the measure against which we could assess the predictive value of all tests. It is a fact that any given test inventory is administered on a particular day, and its results might be influenced by any number of intangibles or contingencies. Whereas, since these intangibles are not present on other days and under other circumstances, the classroom teacher is able to observe and adjust particular evaluations. Hence, the Yearly Grade Average (YGA) of a student should be a rather reliable estimate of academic achievement.

The National Merit Scholarship Qualifying Test (NMSQT) is a battery of five tests used to measure broad aspects of students' general educational development. Some students participate with the hope of qualifying for one of the many Merit Scholarships. Others write the examination to learn more about their educational strengths and weaknesses. All juniors at Hendel Catholic High School write the examination. Because the NMSQT measures the cumulative effect of educational experiences over a long period of time, and because the tests emphasize the ability to use what has been learned, the NMSQT program serves both as a guidance instrument and as a scholarship selection device.

The five areas examined by the NMSQT are: English Usage, Mathematics Usage, Social Studies Reading, Natural Science Reading, and Word Usage. The NMSQT also provides a Composite Score, the average of the five test scores. The Composite Score is considered an index of total educational development.

---

and it has proved in other educational development batteries to be the best predictor of freshman success in college.

The following standardized predictor tests were investigated in this study: 1) The California Short Form Test of Mental Maturity, (CSFTMM): Advanced - Grades nine to adult, S-Form, 1950. This test was devised by Elizabeth T. Sullivan, Willis W. Clark, and Ernest W. Tiegs. 2) The California Test of Personality, (CTP): Secondary - Grades nine to college - Form AA, 1953 Revision. The authors are Louis P. Thorpe, Willis W. Clark, and Ernest W. Tiegs. 3) The Occupational Interest Inventory, (OII): Advanced - Grades nine to adult, 1956 Revision. This test was devised by Edwin A. Lee and Louis P. Thorpe. 4) The California Test Bureau's Multiple Aptitude Tests, (MAT): The authors are David Segel and Evelyn Raskin: 1955.

Hereinafter, in referring to the above-mentioned tests, we shall frequently use the initials indicated in parentheses after each title. Also, when referring to Yearly Grade Average, we shall use YGA.

The term "personality" has many accepted meanings. We will accept the definition as expressed in the manual of administration for the personality test which is used as one of the instruments in this study:

"Personality is not something separate and apart from ability or achievement but includes them; it refers rather to the manner and effectiveness with which the whole individual meets his personal and social problems, and indirectly the manner in which he impresses his fellows."5

By "interests" we do not mean a test of occupational abilities or skills;

but an assessment of occupational preferences. Our purpose is to appraise and analyze the vocationally significant interests of the individuals.

Aptitudes may be considered to be a measure of present characteristics that have been found to be predictive of capacity to learn. English and English, in their Comprehensive Dictionary of Psychological and Psychoanalytical Terms, define aptitude as: "The capacity to acquire proficiency with a given amount of training, formal or informal." According to the authors of the California Multiple Aptitude Tests, their instrument is designed primarily to provide comprehensive differential aptitude data which can be used to help individuals understand their aptitudes and enable them to make wiser decisions in relation to the myriad of vocational, academic, and professional opportunities offered by our modern world.

Purpose and Justification of the Study

The purpose of this study is to investigate the predictive value of a combination of California Test Bureau standardized pencil-and-paper tests in relation to subsequent academic achievement, in order that more effective guidance will be possible to high school freshmen soon after they begin their high school career. This study considers the following questions:

1) How valid are the various "measures" as criteria for predicting students' subsequent academic achievement as a criterion?

2) What are the most effective means of employing these scores or

---

6 E. A. Lee and L. P. Thorpe, Occupational Interest Inventory, California Test Bureau, California, 1956, 2.

7 D. Segel and E. Raskin, Multiple Aptitude Tests, California Test Bureau, California, 1955, 2.
measures to obtain the most reliable prediction of subsequent academic achievement?

3) How large are the errors of the obtained estimates?

4) What recommendations can be made toward more adequate programs of guidance for freshmen high school students to whom these California Test Bureau instruments are administered?

Basically, this study is a statistical investigation of the relative validity of the factors evaluated by the CSFTMM, the CTP, and OII and the NAT with respect to the subsequent academic achievement as revealed by the factor scores on the NMSQT and the YCA scores throughout the students' four years at Mendel Catholic High School. As a cross validating technique with the CTP or as other correlative factors, Character Rating scores by Teachers (ChRT) were likewise investigated as to their relative validity in predicting academic success.

This study is concerned with 210 freshmen boys who elected to be measured by the CSFTMM, the CTP, the OII, and the NAT at Mendel Catholic High School from the years 1955 - 1958. All were not school contemporaries. The scores which the pupils obtained on these tests were studied in reference to subsequent academic achievement-data, earned prior to their graduation from Mendel. These relationships were then considered in the light of previous research and reputable estimates regarding various investigations concerning parts of this study. As will be seen in Chapter II, authorities are divided regarding the prospect of predicting academic success from the results of standardized personality tests and inventoried interests. Yet, according to modern principles of guidance, results of these instruments are quite generally
expected as standard information in each student's cumulative folder for guidance purposes. Assuming that the primary purpose of the high school is to develop and fructify academic abilities, and facilitate academic success, a study of this nature is justified in that it aims at justifying the value of this combination of tests as predictors of subsequent achievement and specifically, how they may be used best for this purpose.

It is commonly thought by teachers - and even by administrators - that individuals who would score above average in a personality test and also obtain above average scores on an interest inventory, and in addition have about average I.Q., would be most likely to succeed academically throughout their high school years. If, as an additional bit of information, it were known that such individuals also obtained above average scores on a multiple aptitude battery, then the expectation would be that these pupils would be above average in academic achievement. As a matter of fact, this has been the assumption of school officials at Mendel Catholic High School in promoting the "experimental" administration of the California Test Batteries to students who elect to take them.

Various facets of the problem comprehended by this dissertation have been studied separately; but no one has ever investigated the total and cumulative predictive value of all four batteries when administered to freshmen boys in a Catholic high school.

Scope and Limits of the Study

Although this study investigates the predictive value of certain paper-and-pencil test scores in relation to subsequent academic achievement, it has definite limitations. No attempt was made to classify the individuals in the
sample according to levels of I.Q. (high, average, or low). Reasons for this
stem from previous studies mentioned in Chapter II relating to these
classifications.

The data used in this study are properly called secondary source data.
The various tests were administered by a competent school psychologist. The
grades and character ratings were determined by the school's teaching personnel.

Over the last several years, the annual enrollment at Mendel Catholic High
School has been approximately 1600 boys. Students are drawn from the local
elementary schools, both parochial and public. As far as the socio-economic
status of the pupils is concerned, it can be said that the majority come from
the middle and upper-middle classes. Very few come from the wealthy class and
a negligible number can be classified as poor or very poor. On the whole, the
largest number of them come from well regulated homes in which both parents are
interested in their children, and eager to give them the best their means allow
in the line of food, clothing, shelter and education. Approximately twenty-five
percent of each freshman class elects to take the California Test Bureau
batteries.

The four tests, namely the CSF'TMM, the CTP, and OII, and the MAT, were
administered each year over a span of three days. The first test administered
was the CSF'TMM; the second test was the CTP; the third, the OII; and the last
was the MAT.

In order to further define the nature of our sample, it should be noted
that the 210 pupils comprising the sample were not school contemporaries. While
all of them wrote the tests during their freshman year, the actual testing was
spread out over four years. Since the individuals in the sample were tested in
their first year of high school, we believe that they were very receptive to the testing program and were still unsophisticated enough to be sincere in personality and interests responses.

Furthermore, Mendel has had a turn-over among the lay faculty of approximately from eight to thirteen teachers each year. The majority of freshman instructors are lay teachers. This fact, coupled with the previous one would seem to give us a random sampling of both students and teachers.

In view of the fact that the California Test Bureau instruments were administered during the freshman year, it was thought that for the purpose of more valid comparison, the ChQT for their freshman year should be used throughout the study. Thus, we have a personality profile as obtained through the paper-and-pencil test, (CTP), and a personality or character profile as given by teachers who actually observed the students in the classroom environment throughout most of their freshman year.

A further limitation upon this study was seen to be the time-difference as to when the CTP was administered and when the ChQT was made. In most instances, there would be a differential of two to four months in actual time evaluation.

Also, this study is being conducted with the assumption that the I.Q. remained constant throughout the four years of high school. As a matter of fact, modern thought holds that the I.Q. could vary in this four year span; and no doubt an average of several I.Q.'s would have yielded a different and better I.Q. score for the four year period of study. This is indicated as another limitation.

The only formal follow-up on the results of the California Test Bureau batteries was done by the school psychologist in a single interview with the boy
and his parents. This would appear to be significant in the interpretation of the various correlations between these factor scores and subsequent academic achievement.

Finally, the chronological age of the pupils was not considered. While it is true that an I.Q. reports a pupil's mental ability in relation to other pupils of the same chronological age, it was felt that in any school grade there is usually a similar range in the chronological ages of the pupils. From a random check of freshman ages, it was considered reasonable to estimate an average chronological age of fourteen years. (Even if we could have obtained the chronological ages, we could not use another variable; since the IBM-650 can handle only fifty variables, and we already had that many.)

It must be realized that the generalizations of this study can be applied only to a limited extent to students enrolled under like circumstances in similar situations.

Methods to be Used in the Study

Most of the studies concerning the CSFTMM, the CTP, the OII, and the MAT have been essentially statistical studies which investigated only limited aspects of this problem. The present study will attempt to investigate the broader problem of prediction of academic achievement from the results of all of these tests.

A complete evaluation of any prediction of academic success should include consideration of the many known or assumed components which comprise the entire process: intelligence, personality, interests, and various aptitudes. Therefore, it was the intention of the author to investigate the predictive value of the entire problem as it applied to the 210 Mendel graduates, rather than a
few isolated factors of the problem. The procedure of predicting academic success from the above mentioned tests was approached through the following steps:

1) A review of previous research on the problem of predicting academic success by means of any one or combination of the following measures: Intelligence Tests, Personality Tests, Interests Inventories, and Aptitude Tests.

2) All of these tests had been administered at least four years ago and the results were on file at Mendel High School. The first order of business was to gather the data and uniformly represent it on appropriate file cards. The yearly grade averages were obtained from the individual student's permanent record cards in the main office. The Character Ratings by Teachers were obtained from the Guidance files. The achievement scores on the National Merit Scholarship Qualifying Tests were obtained from the permanent record of each of the students. Finally, all of this data was entered on standard I.B.M. cards.

3) Employing an I.B.M.-650, all possible Pearson-product-moment coefficients of correlation were determined. There then followed a study of relationships between the predictor variables (intelligence quotient, personality scores, interests scores, aptitude scores, ChRT scores) and the criterion variables (NMSQT scores and YGA scores). Appropriate statistical techniques were used with a view to optimum prediction and/or classification according to future performance in academic endeavor. The DuBois multiple regression equation technique was employed.

4) After determining the best multiple regression equations, it would appear that we could regard them as useful prediction formulae which would
enable administrators to foresee degrees of academic achievement.

Guidance personnel and teachers could then concentrate more completely and thoroughly upon the developing of these particular factors in all of the students, especially those whose records show them to be deficient in any of the factors of the multiple regression equations.

It was thought best to develop two multiple regression equations. Since the California Test Bureau instruments were administered during the freshman year, one multiple regression equation would establish the best combination of factor-scores for academic achievement in that particular year. In the junior year, the National Merit Scholarship Qualifying Test is administered; hence, a second multiple regression equation would identify the best combination of factor-scores for optimum achievement in the NMSQT. Other multiple regression equations were considered possible for predicting the optimum achievement at the end of the sophomore year, at the end of the junior year, and at the end of the senior year.
CHAPTER II

REVIEW OF THE LITERATURE

Certain basic information about students is needed if we are to understand them and be able to assist them. Undoubtedly, the most important information is about the intelligence and aptitudes of these students. Without knowing something about their potential, we cannot determine what to expect from them. We certainly need to know their achievement in terms of objective measurements. How does the student stand in comparison with others of his own age and grade? It is also important to know the teacher's estimate of what a student has accomplished in terms of grades or scholarship.

Less tangible, but nevertheless very important information about the student concerns his interests. It must be recognized that interests change and mature, are subject to flights of fancy, and are sometimes attached to factors which obscure their real nature. It is also important to know how a person feels about himself and toward others. Even though there are differences in regard to theories of personality, the teacher and the counselor must assist in the adjustment of the individual in this important area.

All of the above factors are involved more or less in this study; and it is important to clarify the meanings of these factors. The first part of this

---

The current emphasis on guidance in the secondary schools usually highlights the importance of a battery of tests which can give a comprehensive array of guidance data. In the past, the emphasis in measurement has been in the direction of survey and administrative uses of the test results. These are desirable purposes. Administrators and supervisors use test data to relate the accomplishments of students in their schools to those of the average or typical school—the norm—and also to obtain information about the characteristics of their school populations. However, the primary purpose of testing is essentially an instructional one. Testing is a means of getting as much information as possible about pupils. Such information, when properly used, can serve as the basis for directing further learning.

Survey of Psychological Data

Intelligence

The science of mental measurement began to take definite shape about 1900. Three key ideas or trends were mainly responsible for the development of intelligence testing. First was the concept of mental maturity, implying a process of mental development taking place from within the growing child. Secondly, the systematic study of individual differences, which began in the nineteenth century and developed rapidly through the use of statistical

techniques of measurement. The third was the need for classifying pupils in terms of their abilities.

Not all psychologists agree as to the exact nature of intelligence. Some have maintained that there is only one kind of mental ability (general intelligence), that there is simply a gradient of brightness-dullness, that the bright students can learn anything better than the less bright. Others insist that there are many different kinds of abilities, and that a high degree of one kind tells us nothing about the degrees of other kinds of abilities.\(^3\) The problem of classifying abilities has many practical aspects, as well as aspects of theoretical interest. First, the construction and use of achievement and aptitude tests depend on knowledge about the kinds of abilities. If there is only one important mental ability, we need only one kind of test to measure it. If many, we need many kinds of tests to assess properly the abilities of students and to plan their educational experiences.

Terman declares "an individual is intelligent in proportion as he is able to carry on abstract thinking."\(^4\) Spearman understood intelligence to be comprised of two factors, one being general and the other factor being specific.\(^5\) Thurstone presented the "multiple factor" theory which stressed intelligence as consisting of several abilities, seven of which have been


\(^4\)L.M. Terman, "Intelligence and Its Measurement", *Journal of Educational Psychology*, March, 1921, 128.

identified as (1) number ability, (2) word fluence, (3) space, (4) memory, (5) perceptual speed, (6) reasoning, and (7) verbal comprehension. Guilford, in a rather recent article, indicates the multiplicity of factors comprising intelligence. Using factor analysis, he discerned at least 150 distinct elements which are measurable. Hence, he concludes, no one test could ever hope to thoroughly evaluate intelligence.

Intelligence has frequently been defined "as the ability to adjust to the environment or to learn from experience." As Garrett has pointed out, this definition is too broad to be very helpful in practical work. A more psychological and more generally applicable definition is suggested by Garrett. "Intelligence", he states, "includes at least the abilities demanded in the solution of problems which require the comprehension and use of symbols."

Super, commenting on this definition, affirms its operational nature in that it is based on an analysis of the tasks involved in solving the problems presented by an intelligence test. It is broader than some test-based definitions because it applies not only to the tasks presented by the test, but also to the tasks presented by the school or college courses, success in which, it is designed to predict. It has the additional advantage of taking into account the important work of the past ten or fifteen years which demonstrates that intelligence is no one aptitude, but a constellation of aptitudes.


8 Donald E. Super, Appraising Vocational Fitness, New York, 1949, 86.
"Ability" commonly refers to something a person can do here and now. "Aptitude", on the other hand, usually has a future reference, and in the general sense has been defined as a "condition or set of characteristics indicative of ability to learn something." 9

Super summarized general conclusions emerging from studies by Stoddard10, McNemar11, Thorndike12, and Wellman13, regarding methods and results of nature-nurture concepts of intelligence, by saying:

"...Whereas both nature and nurture play a part in the development of intelligence, mental ability as indicated by the intelligence quotient is relatively constant from the time a child enters elementary school until late adulthood. It is true that the obtained I.Q. will vary some after the age of six, but this is generally more a function of the tests ... than of the individual. Some changes which are too great to be explained by these causes are the result of emotional conditions which invalidate the score of one test, or of organic changes resulting from disease or injury. That there are other changes, not explained by any of these factors and attributable to changes in the environment which modify the functioning intelligence, has not been demonstrated to the satisfaction of all competent judges with persons of elementary school age or older."14

The role played by intelligence in educational achievement has been studied


14Super, Appraising Vocational Fitness, 86-87.
by Pintner and Strang. Despite some results revealed by Pintner, Super maintains that the correlation between intelligence tests and grades is not especially high. Numerous summaries of the subject show that in high school they tend to range from .30 to .80, and in college from .20 to .70, the modal range being between .40 and .50 in high school and .30 and .50 in college.

The relationship of intelligence test scores to educational achievement has been demonstrated in one other type of study, in which a genetic approach has related intelligence to amount of education obtained. Studies reviewing the period 1917-1930 make it clear that, on the whole, those who are most able, obtain the most education.

PERSONALITY

Scholastic ability test scores are reasonably good indicators of probable success or failure in certain kinds of scholastic activities. But such scores are not always accurate predictors of academic success, because success depends not only upon what is measured by the test but also upon personality factors.

Guidance workers have two major reasons for studying the personality of students. First, they must make some attempt to determine the appropriateness

---


17 Super, *Appraising Vocational Fitness*, 90.

18 W.M. Proctor, "Intelligence and Length of Schooling in Relation to Occupational Levels", *School and Society*, 1935, 783-786.

of a student's personality make-up for the occupational field which he may be planning to enter. Second, they must try to identify the maladjusted student. His maladjustments may be due not to lack of aptitude but rather to other factors which are preventing him from using his aptitude effectively. Among such factors are the personality traits exhibited in excessive daydreaming, excessive shyness, bullying or other overaggressive behavior, and attention-getting devices including those used to seek sympathy.

One point of view looks upon the personality as a whole global unit, complex in nature, and therefore virtually unanalyzable. This is the gestalt protest against the behaviorists' atomistic approach. Another concept, generally credited to Thorpe\(^{20}\), is known as the "social stimulus value" of the individual. Under this definition of personality, the individual would be observed in groups, and interviews or written reports would be obtained from persons who know him. It is obvious that this definition is limited by its empiricism. A third definition is that given by Allport:\(^{21}\) "Personality is the dynamic organization within the individual of those psychophysical systems that determine his unique adjustment to his environment." Those who hold to this definition would concentrate on interviews or projective techniques. A fourth definition considers personality as a pattern of traits or ways of reacting to external stimuli.\(^{22}\) This point of view is unitary and therefore

\(^{20}\) Louis P. Thorpe, Psychological Foundations of Personality, New York, 1938, 64.


\(^{22}\) Clyde Kluckhohn and Henry Murray, Personality in Nature, Society and Culture, New York, 1949, 44.
If we accept the more common assumption that personality is a more or less definite set of attributes made up largely of a number of generalized traits, the fact that there is difficulty in defining it and lack of agreement concerning its components should not be an insuperable obstacle to measurement. A similar situation exists with respect to intelligence, yet tests in that field have been in general use for years and have shown their worth in a guidance program. Trowler defines personality as "the sum total of an individual's behavior in social situations." Behavior includes not only overt acts but inward feeling-tones produced by the situation as interpreted by the individual through introspection.

In testing personality and predicting from personality tests, we still do not know how much of a given personality trait or what combination of personality traits makes for successful school achievement or successful job adjustment. One reason for this is that people vary in their definitions of a given trait and in their opinions of what behavior indicates that trait. Another reason for this inability to relate personality traits to success in school or outside is the difficulty in agreeing upon standards of "good" or "bad" performance for the scores made on personality tests.

Personality traits warrant careful study not only because they are important in day to day behavior, but also because they are related to interests. Research studies which revealed interest types showed also that people who had particular types of interests tended to have personality

---

patterns related to these interests. 24

Such evidence indicates that occupational adjustment is based not only upon mental ability and special achievement but also upon occupational interest and, to a certain extent, upon personality traits. These interests and personality traits can be estimated or measured with a certain amount of accuracy. 25

According to Traxler, most personality tests and interest measures are reliable enough for group studies, and they may be used for research if they can be shown to be reasonably valid. The majority of these instruments, however, are not reliable enough to be very satisfactory for use in individual diagnosis. Individual scores on personality tests can contribute materially nevertheless, to the guidance program of the school. 26

INTERESTS

Interests have probably received more attention from vocational psychologists during the past generation than any other single type of human characteristic, including intelligence, aptitudes and personality traits.

There have been four major interpretations of the term interest connected with as many different methods of obtaining data. Super classifies them as expressions, manifestations, tests and inventories of interests. Inventoried interest is assessed by means of lists of activities and occupations which bear

24 Froehlich, Studying Students, 306.

25 Ibid., 307.

26 Traxler, Techniques of Guidance, 108.
a superficial resemblance to some questionnaires for the study of expressed interests, for each item in the list is responded to with an expression of preference. The essential and all-important difference is that in the case of the inventory, each possible response is given an experimentally determined weight, and the weights corresponding to the answers given by the person completing the inventory are added in order to yield a score which represents, not a single subjective estimate as in the case of expressed interests, but a pattern of interests which research has shown to be rather stable.  

Out of the variety of means of ascertaining interest grows one of the most perplexing characteristics of interest research. According to this research, the various means of identifying interests do not yield comparable data. In fact, at times, they provide conflicting information.  

Lee Cronbach  opens his chapter on interests with the following statement: "An interest may be defined as a tendency to seek out an activity or object, or a tendency to choose it rather than some alternative." Hahn and MacLean state: "Interests are an aspect of personality development shaped by both hereditary and environmental factors." Strong claims that interest is an aspect of behavior, a response to a liking, and "since interest involves

---

27 Super, Appraising Vocational Fitness, 376.
28 Froehlich, Studying Students, 279.
30 Milton E. Hahn and Malcolm S. MacLean, General Clinical Counseling, New York, 264.
reactions to specific things, they must all be learned. Accordingly, they may be modified later on by re-education." According to Froehlich, they may be thought of as one of the forces that motivate activity. In other words, they represent a tendency to select one activity or thing in preference to something else, to choose one instead of another. Said even more simply, interests are likes.

According to Super, an objective theory would recognize the fact of multiple causation, the principle of interaction, and the joint contribution of nature and nurture. He says, "Interests are the product of interaction between inherited aptitudes and endocrine factors, on the one hand, and opportunity and social evaluation on the other." ... "For most persons, adolescent exploration is an awakening to something that is already there."

Interests are very important. It is important to be interested in the educational program one is following. Lack of interest contributes to poor performance or to dropping out of school before graduation. It is important to be interested in the occupation one selects, for one may pursue the same occupation for fifty years. Happiness in any occupation is very closely related to one's interest in the activities and associations of that occupation.

Childhood interests do change radically. Empirical evidence obtained over a period of years has indicated that although interests of some high school

---


33 Super, *Appraising Vocational Fitness*, 406.
freshmen are assuming fixed patterns, there are still many students who change during the sophomore year. During the junior year, or when the students are about 17, interests begin to crystallize. The relationship between interest patterns and other personality factors is attested by a growing body of research; persons of different interests have been shown to differ in varying degrees on such basic personality traits as extroversion, sociability, and masculinity-femininity.34

The relationship between interest and ability is surprisingly less clear. On a priori grounds, we should assume a relationship to exist for two reasons: 1) High ability in some activity should lead to success and satisfaction, thence to increased interest; while low ability should lead to failure, dissatisfaction, and loss of interest. 2) High interest in an activity should tend to persist and become stronger when ability is high, since superior ability will ordinarily lead to more satisfying consequences. We would thus expect a high degree of relationship between literary interests and ability in reading and writing. On this kind of reasoning, Strong has maintained that "interests reflect inborn abilities." But correlations between interests and measures of ability have been almost uniformly low; some relationship exists, but not much. In any case, we cannot assume that scores on interest inventories are indirectly measuring abilities, or that abilities directly influence such scores.35

The purpose of interest testing is to predict satisfaction; the test of

34Stoops, Principles and Practices in Guidance, 57.

35Commins, Principles of Educational Psychology, 385.
validity for an interest inventory is how well a high score predicts a liking for an occupation or a certain field of study. Conversely, low scores should be associated with dislike or dissatisfaction.

Multiple-Score Tests of Aptitudes

The importance of the results gained from factor analysis is just beginning to make itself felt in educational and vocational guidance. Item and factor analyses have indicated that tests do not have the unitary wholeness formerly taken for granted. It is coming to be realized that "intelligence" is a constellation of abilities. In the introduction to the Examiner's Manual for the Tests of Primary Mental Abilities, the Thurstones36 stated specifically, "By use of the statistical tool of factor analysis, psychologists have found that intelligence is made up of a number of more or less independent describable, testable, and significant abilities."

As indicated previously, the content of a multi-factor aptitude battery may be a longer version of an intelligence test. If a counselor is interested only in determining the general level of ability in an individual, one of the tests of general mental ability will suffice. If it is recognized that these tests of general mental ability fail to reveal needed facts about the individual, tests providing language and non-language scores, linguistic and quantitative, or verbal and performance scores would be desired. None of these, however, give all the information required by the counselor who is

engaged daily in educational and vocational guidance. Simple observation of
the heterogeneous population in our schools will show other types of ability
not measured in our tests of general ability or scholastic aptitude. Since
only about forty percent of the student group, on the average, will go on to
college, assistance is needed in assessing the aptitudes of the remaining sixty
percent.37

The same author stresses that a great deal more needs to be done to
validate multiple-score tests and to develop norms that will be helpful in
educational and vocational guidance. These tests certainly give a great deal
more information about the individual than was obtained from the single-score
tests.

In a discussion of the Differential Aptitude Tests, Bennett, Seashore, and
Wesman further amplified the need for getting away from the single all-embracing
comprehensive score:38

"...the use of a batter of tests illuminates cases which were quite
obscure when only a single-score IQ was available. For example, two
students who are of the same chronological age and earn the same total
score on a test composed of verbal and numerical items, will receive
the same IQ. Yet one of these students may have answered all the
numerical items correctly and done but indifferently well on the verbal
items; the other may have answered few of the numerical items correctly,
but handled the verbal items well. Can anyone doubt that these boys,
described as identical by the IQ, will perform quite differently in a
course in algebra—or that they may need to consider different careers?

"...It is becoming more and more apparent that a sensible attempt
to obtain a rating on more than one of the facets of the mind will
provide more meaningful information for guidance purposes than will a
single score."


38George K. Bennett, Harold G. Seashore, and Alexander O. Wesman,
A Casebook for the Differential Aptitude Tests, The Psychological Corporation,
New York, 1951, 9-10.
According to Bingham, a student's readiness to develop an interest in his potential ability, his ability to become thoroughly engrossed in his work, and his ability to perform at a satisfactory level of competence may be considered as measures of aptitude. The same author explains: "In referring to a person's aptitude for mathematics or art ..., we are looking to the future. His aptitude is, however, a present condition, a pattern of traits, deemed to be indicative of his potentialities."  

The results of multiple aptitude testing serve as one possible source of information giving a general indication of an individual's degree of talent, but not as the sole criterion of his promise. Due consideration must be given also to other such data about the individual as intelligence test scores, achievement test scores, interest test scores, and evidence concerning personality traits. The measuring of aptitudes is a matter of predicting future achievement from present performance.

Yearly Grade Averages - Indicators of Scholarship

Scholarship, meaning grades earned by students, is separated from achievement on standardized tests because of important implications. While abilities and achievement can be measured by objective means and are very important factors to be known by the counselor, the degree to which the individual has been able to utilize both his ability and his knowledge in specific instances

---

40 Ibid., 16.
is sometimes more important than the first two factors. An individual who is unable to perform according to his expected ability and achievement certainly should have the special attention of the counselor.

In most instances, a student's grade in algebra may include, in addition to the individual's knowledge of algebra, his general acceptance by the instructor and the class, his manner of turning in his homework, or lack of homework, and evaluation of his verbosity during class, and possibly other factors. Teachers have been heard to say: "Fifty percent of your grade in this class will be determined by your attitude." Yet, with all these factors represented in a grade, according to authorities on the subject, it is still one of the best indicators of future academic success.

Everyone concerned with education is anxious to know how much students learn in school. For years, teachers have made judgments about students' levels of achievement, using marks to record these judgments. But making such judgments has often plagued teachers—and not without reason. No exact or completely satisfactory yardstick with which to judge a student's achievement is available. In fact, all of the yardsticks now in use are open to some criticism.

The unreliability of teachers' marks was demonstrated in a pioneer study by Starch and Elliot. They had a number of English teachers grade the same final examination paper, written by a student. Working independently, each

---

41 Stoops, Principles and Practices in Guidance, 51.
42 Froehlich, Studying Students, 238.
teacher studied the paper and assigned a mark to it. Their marks ranged from 50 to almost 100 percent. The results of this and many other similar studies have raised many questions about the so-called "essay" type exam and teachers' evaluations of students' achievement.

The N.M.S.Q.T. - Measures of Standardised Achievement

While the declared purpose of these tests is to measure broad aspects of students' general educational development, they may properly be regarded as achievement tests.

Typically, tests of this type cover large areas, such as social studies, mathematics, English, word usage and natural sciences. Since these tests are broad in scope, they do not place too much emphasis on determining the range of a student's knowledge or the level of his skills. They attempt, rather, to measure his ability to apply his knowledge and skills to the solution of a new problem, or to the interpretation or evaluation of unfamiliar material.

Evidence has long been available to indicate that such achievement tests are satisfactory indicators of academic success. In one study, correlations between Iowa test scores and school marks ranged from .75 to .90.44

Diagnostic achievement tests are similar in many respects to subject-matter tests; the principal differences between the two types of tests are the kinds of scores that can be obtained. Diagnostic achievement tests are arranged so that they yield not only a total achievement score, but also individual scores for

---

each type of knowledge or skill measured. The National Merit Scholarship Qualifying Test provides a score for each of these areas: English usage, mathematics usage, social studies reading, natural science reading and word usage along with a composite score.

PREDICTIVE STUDIES

As modern society has become more conscious of the importance of fully utilizing its human resources as well as natural resources, psychological testing has come into increasing use. One of the techniques which has been developed is that of administering a battery of psychological tests to a large number of individuals in an attempt to determine the patterns of abilities necessary for success in various fields of endeavor.

While there have been many investigations of the merits of specific instruments as predictors of academic success at the high school level, the author was unable to find any which investigated the predictiveness of a complete battery of instruments measuring all the facets of psychological development which are considered important by guidance experts. Furthermore, no studies were found which confined such investigations exclusively to the specific instruments of a single test publisher. The areas regarded by guidance experts as significantly important in the evaluation of high school students are: intelligence, personality, interests, special aptitudes and achievement.

Attempts to predict academic success are many and varied. May  is one of

---

the forerunners. He investigated the predictability of freshman honor points by means of a combination raw score on the Miller Mental Test, hours spent in study, admission units, the Dartmouth Completion of Definitions Test, and the high-school record. Employing partial correlation technique, he found intelligence to be the most important factor; high school averages came second; units offered, third; and application, fourth.

On the secondary-school level, Rector\(^\text{46}\) testing by correlation the relative validity of three predictive criteria—group intelligence quotients (Army Alpha), scholarship, and application ratings—concluded that all three were invalid for predicting success in high school as measured by the teachers' final marks.

Ross\(^\text{47}\), in a study of the diagnostic value of individual grammar-school records, found a higher correlation between teachers' marks and high school success than between the latter and intelligence test scores. He concluded that a combination of intelligence scores, achievement test scores, and teachers' ratings would make the best prophecy.

Ross and Hooks\(^\text{48}\), reviewing correlation results of other studies, concluded that the grammar-school record, including age, grade progress, attendance, and teachers' marks, was the best means of predicting future academic success.


\(^{47}\)Clay C. Ross, \textit{The Relation Between Grade-School Record and Achievement}, Teachers College, Columbia University, New York, 1925, 44.

Gates and La Salle's study\(^{49}\) was regarded in the 1930's as an excellent example on the elementary level. Using the Stanford Binet Intelligence Test, the National Group Intelligence Examination, and certain educational tests in reading, arithmetic, and spelling, in grades III to VI, these investigators concluded that the best predictor in any of these subjects would probably be an educational test in that subject.

Coe\(^{50}\) using a sample of 230 pupils, the majority of whom were resident students, administered Terman's Group Test of Mental Ability along with ten standardized subject-achievement tests. He concluded that an average of all subject-grades from grammar school predicts first-year high-school average (success) better than all of the predictor measures used. He goes so far as to say that this combined average of grammar school grades will predict freshman success better than either seventh grade English or seventh grade mathematics marks predict first-year high-school English and mathematics marks, respectively.

Jensen\(^{51}\) investigated the relationship between school achievement and scholastic aptitude. High School students in four California high schools were administered the following tests: The Progressive Achievement Test, the S.R.A.


\(^{50}\) Roger L. Coe, \textit{Predicting First-Year High-School Success in a County School System}, Published Doctoral Dissertation, George Peabody College for Teachers, Nashville, 1934, 46.

Primary Mental Abilities Test as measures of intelligence; the California Test of Personality and the Kuder Preference Record as measures of personality and occupational preference. The pupils were classified into three groups — those whose average mark scores exceeded their intelligence test scores by a specified amount were classified as "over-achievers"; those whose intelligence test scores exceeded their average mark scores by the same amount as the average marks of the "over-achievers" exceeded their intelligence scores were classified as "under-achievers". "Par-achievers" were those pupils whose average mark score and intelligence score matched each other within a narrow range.

Jensen found the correlation between scholastic aptitude and school marks to be high for all three groups. The coefficients were as follows: over-achievers, .91; par-achievers, .09; under-achievers, .78. These correlations, he concluded, indicate that scholastic aptitude is not a casual factor in under- and over-achievement. In addition, he found that the correlations between reading and school marks were as follows: over-achievers, .05; par-achievers, .80; under-achievers, .65. These correlations, it was concluded, indicate that reading skill or lack of it was not a critical factor in under-achievement and over-achievement.

Comparison of the averages of the three groups on the California Test of Personality showed nothing that would indicate any relationship between under- and over-achievement and scores on this test. Finally, he concludes that the data of this investigation shows that high ability pupils under-achieve and low ability pupils over-achieve.
Hillebrand\textsuperscript{52}, studying seventy-four "gifted" students, all of whom were appraised by the California Short Form Test of Mental Maturity, the California Test of Personality and the Occupational Interest Inventory, concluded that the possession of intelligence does not imply personal or social adjustment; nor does it relate significantly with interest factors. As evidence, the correlations between intelligence and personality factors are cited:

**Personal Adjustment**
- Self-Reliance
  \( r = .190 \)
- Sense of Personal Worth
  \( r = .099 \)
- Sense of Personal Freedom
  \( r = -.016 \)
- Feeling of Belonging
  \( r = -.050 \)
- Withdrawing Tendencies
  \( r = -.019 \)
- Nervous Symptoms
  \( r = -.303 \)

**Total Personal Adjustment**
\( r = .140 \)

**Social Adjustment**
- Social Standards
  \( r = .112 \)
- Social Skills
  \( r = .162 \)
- Anti-Social Tendencies
  \( r = .118 \)
- Family Relations
  \( r = .090 \)
- School Relations
  \( r = -.0004 \)
- Community Relations
  \( r = .079 \)

**Total Social Adjustment**
\( r = .117 \)

**Total Adjustment**
\( r = .150 \)

The correlations obtained between intelligence and interests are summarised as follows: The range of the correlations for Fields of Interest was from -.227 in the natural field of interest to .28 in the scientific field. Types of Interest correlations ranged from .018 in verbal interests to .281 in

computational interests. 53

Turcich 54 employed the California Short Form Test of Mental Maturity, the California Test of Personality and the Occupational Interests Inventory to investigate the relationship of intelligence to personality traits and interest patterns. Using a sample of forty-three ninth grade boys in a private school, having low average and below average intelligence quotients, he concluded that the correlation between intelligence and personality components indicate that there is a low or possibly inverse relationship between the two. The coefficients of correlation between intelligence and the Personal Adjustment components range from -.288 to .281. The coefficients of correlation between intelligence and the Social Adjustment components ranged from -.013 to .148. None of the correlations were significant even at the five percent level.

From the correlations between intelligence and the different Fields of Interest, it was seen that they ranged from -.012 in the Natural Field of Interest to .222 in the Sciences. Among Types of Interest, the coefficients ranged from -.019 in Manipulative Interest to .028 in Computational Interest.

To determine whether or not success in dental school is related to personality and interest factors was the basis of a study by Thompson. 55

53 Ibid., 41.


The California Test of Personality and the MacQuarrie Test for Mechanical Ability were administered to 158 freshmen. The California Test of Personality and MacQuarrie Test for Mechanical Ability and the Kuder Preference Record were administered to 66 seniors. Criteria for the test evaluations were made by separating completely, qualitative from quantitative grades. This was done by calling cumulative points earned on product, or work done, quantitative; and cumulative points earned in process, or technique and theory, qualitative.

Correlations obtained for the 158 freshmen showed that there is a positive relationship between quantitative criterion scores and Social Adjustment and Total Adjustment scores on the California Test of Personality. All correlations between quantitative criterion scores and Social Adjustment, Self Adjustment, and Total Adjustment are statistically significant. The correlation between Social Adjustment and qualitative criterion scores is statistically significant.

These findings tend to support the contention that personality factors make a definite contribution to success in dental school.

LaBelle surveyed the relationship of personality adjustment to academic achievements. In her study, she took a group of 56 pupils ranging in age from 12 to 17 years, and whose intelligence quotients ranged from 65 to 126. She found that regardless of a given intelligence quotient, there is no escape from the fact that there is some form of personality maladjustment. Those whose intelligence quotients were high appeared more aware of their limitations and

perhaps for that reason revealed a great lack in personal security. The findings in regard to the relationship between personality and academic achievement was negligible for the group.

The effectiveness of remedial reading programs on reading and personality adjustment has been the source of an increasingly large number of research studies. Many of the studies showed a concomitant improvement in reading and personality adjustment. Schonell\(^57\) reports good results for retarded readers who were given a comprehensive remedial program featuring meaningful reading experiences. Those who had remedial training showed significant progress over those who were left in regular classes without any special help. Remedial work in some cases even tended to erase minor neuroses and unhappiness.

Ardizzone\(^58\) studied 700 pupils in his investigation of the differences in reading, intelligence and personality among high, average and slow groups of sixth and seventh grades. As appraising instruments, he used the California Mental Maturity Test, the California Test of Personality and the California Achievement Tests. He defined an advanced reader as one who is at least one and one-half years in reading achievement beyond actual grade placement; and a retarded reader as one whose achievement is at least one and one-half years below actual grade placement. The table following presents the chief results of Ardizzone's study.


\(^{58}\) Frank Ardizzone, Differences in Reading, Intelligence, and Personality Among High, Average, and Slow Groups of the Sixth and Seventh Grades, Unpublished Master's Thesis, DePaul University, 1959, 3.
## Results of Ardizzone's Study of I.Q. and Personality

<table>
<thead>
<tr>
<th>Group</th>
<th>Intelligence Quotient</th>
<th>Reading Ability</th>
<th>Correlation between I.Q. and Personality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>115 and over</td>
<td>Advanced readers</td>
<td>.24</td>
</tr>
<tr>
<td>2</td>
<td>90 - 115</td>
<td>Advanced readers</td>
<td>.09</td>
</tr>
<tr>
<td>3</td>
<td>90 - 115</td>
<td>Retarded readers</td>
<td>.24</td>
</tr>
<tr>
<td>4</td>
<td>70 - 90</td>
<td>Retarded readers</td>
<td>.27</td>
</tr>
</tbody>
</table>

He found that behavior problems in the form of personal and social mal-adjustment are more numerous among retarded readers and those of low intelligence than those of average and above average ability and intelligence. Evidence in the above table leads to the conclusion that intelligence has very little influence on an individual's personality.

Hallock⁵⁹ endeavored to determine whether or not there is a relationship between reading achievement and certain attitudes children hold about themselves and their environment, which could be measured by a paper-pencil test of personality. While his investigation concerned fourth, sixth and eighth grade children, it is enlightening for the present study. His measuring instruments

⁵⁹George A. Hallock, Attitudinal Factors Affecting Achievement in Reading, Unpublished Doctoral Dissertation, Wayne State University, 1958, 89.
were the California Short-Form Test of Mental Maturity, the California Reading Test and the California Test of Personality. He found significant relationships between success in reading and certain attitudes which children have toward themselves and their environment-family relations, self-reliance, anti-social tendencies, feeling of belonging, withdrawing tendencies, school relations, nervous symptoms, and feeling of personal worth. Community relations, social skills, and sense of personal freedom appear to be less significant. He concludes that certain attitudes as measured by the California Test of Personality bear a relationship to reading achievement; that when these attitudes are favorable, reading achievement is satisfactory; that when these attitudes are unfavorable, children fail to learn to read. He states that "the California Test of Personality will not enable the teacher to make a clinical evaluation for each child in her class, but it will give her additional clues to add to the information she already possesses about a particular child. This may enable her to better help him once he has acquired a reading disability, or, better still, help prevent the disability from occurring."

Forlano and Wrightstone used the "California Test of Personality" to obtain self-descriptive ratings from students and character ratings by teachers having direct relationships with the students (along with other sociometric measures) as correlatives. His abridgment of the California Test of Personality was original; and hence, not available for practical use. They drew the following tentative conclusions: "If rated teachers' observations are considered as
giving a fairly accurate and valid estimate of pupil adjustment and maladjust-
ment, as observed at schools, it may be tentatively concluded that the 
California Test of Personality or either of its two parts, correlates highly
enough with teachers' ratings to be used as part of a battery to identify pupils
with personality problems."

It is interesting to note that Carroll states that there is difficulty in
obtaining valid results from a personality test because of the attitudes of
the examinee. Individuals are hesitant about revealing their emotional
problems and their adjustment difficulties. They tend, comments Carroll, to
feel that they must protect themselves. A desire to maintain self-respect and
to win the respect of others is usually stronger than the desire to be
completely honest. Reliability is difficult to establish. Personal behavior
is both consistent and inconsistent. Some traits show more consistency than
others. Personality tests have a value in certain situations when used with
discrimination. Tests are valuable as supplementary tools if the person using
them constantly keeps in mind that the results are subject to many reservations
and qualifications.

IN SUMMARY: The literature in the field dealing with the relationship of
intelligence to personality factors and interests patterns is wide and varied.
For the most part, writers deal with a correlation of intelligence to one phase
of personality or achievement such as reading.

61 H. A. Carroll, Mental Hygiene. The Dynamics of Adjustment, Third Edition,
New Jersey, 1950, 356-357.
Charles 62 investigation of the relationship between Level of Interest and academic achievement among 960 graduating senior students from 34 small and medium-sized high schools, led to the conclusion that a highly significant relationship exists between level of interest on the Occupational Interest Inventory and school achievement.

Smith 63, after investigating several preference scales in their relation to academic success, came to the conclusion that certain preference scales materially improve the prediction of grades for students selecting various major courses at the A. & M. College of Texas. He comments that this sort of increase is understandable because the preference scales usually overlap very little with measures of ability.

Chilese 64 studied the vocational characteristics of mentally gifted students. One hundred and twenty-seven subjects selected for the study were within the I.Q. range of 127 to 146. These subjects took twenty-two different tests measuring their vocational characteristics. He found that on the Bernreuter Personality Inventory and the Adams-LePeley Personality Audit these mentally gifted students had scores which approximated a normal distribution.

62 Ramon L. Charles, A Study of the Relationship Between Level of Interest and Academic Achievement, Unpublished Study, Division of Guidance and Personal Services, Kansas State Department of Public Instruction, February, 1957, 44.


and were average in all characteristics except for Firmness and Tolerance. For these two personality characteristics, the scores were exceptionally high.

As regards vocational characteristics, Chiles found that these mentally gifted subjects scored high in the social service and scientific vocational fields. On most of the tests, the results showed low interests in the mechanical, clerical and computational fields. Also, the results were low in the practical affairs of the business world and in accounting.

"Mentally gifted students are highly verbal, have vocational interests in the scientific and social service areas, and tend to make high scores in all tests given at the office of Vocational Guidance and Placement, except in those areas in which interests along clerical and computational lines are measured."

McGhee and Lewis felt a need for a larger sampling over a greater geographical area. The authors accumulated a population sampling of 4,797 boys and 4,264 girls drawn from a total population of about 50,000 covering 36 states. The appraising instruments were the Kuhlmann-Anderson Intelligence Tests and the Maller and Boyton Personality Inventories with the addition of teachers' rating of personality traits. Using the upper ten percent of the sample as the superior group and the lower ten percent as the retarded group intellectually, they concluded that the retarded group revealed significantly more maladjustment than the superior group.

In still another group, Connors made some observations of the personality


traits and vocational interests of some continuation school students. He administered the California Test of Personality to one hundred pupils having a mean I.Q. of 85. The results of the tests indicate that the pupils concerned were not adjusting themselves to the conditions which confront them. They are not a normal, happy group.

Strong 67 rather recently observed that "... interest tests do not correlate to any practical degree with measures of success."

Racky 68, in an aptitude - interest study concerning the ninth grade woodshop performance of students, found statistically significant correlations between the semester's grade and the S.R.A. Primary Mental Abilities Test; the Kuhlmann-Anderson Intelligence Test; ... the Garretson Interest Questionnaire for High School Students; and the Kuder Preference Record mechanical interest, literary interest, and musical interest. Interest as measured by the Kuder Preference Record for the outdoor, computational, scientific, persuasive, artistic, social service, or clerical areas was not significantly related to the grades.

Racky's study has additional importance in that it represents a modern technique of administering a battery of psychological tests to a reasonably large number of individual's in an attempt to determine the patterns of


abilities necessary for success in a specific field of endeavor. His multiple regression equation predicting success in ninth grade woodshop included five variables (psychological evaluations) and under the test of cross-validation, proved to be of diagnostic value in indicating pupils who might experience difficulty in ninth grade woodshop.

Layton and Swanson investigated the predictiveness of the Differential Aptitude Tests battery by studying students in twenty-seven schools randomly chosen from among the one hundred and eight Minnesota high schools which had given the DAT in grade nine, more than two years earlier. Scores on Verbal Reasoning and Numerical Ability were the best predictors of academic rank more than two years later. The coefficients of correlation are presented below:

<table>
<thead>
<tr>
<th>Test</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N=628)</td>
<td>(N=532)</td>
</tr>
<tr>
<td>Verbal Reasoning</td>
<td>.56</td>
<td>.58</td>
</tr>
<tr>
<td>Numerical Ability</td>
<td>.57</td>
<td>.51</td>
</tr>
<tr>
<td>Abstract Reasoning</td>
<td>.45</td>
<td>.43</td>
</tr>
<tr>
<td>Space Relations</td>
<td>.36</td>
<td>.32</td>
</tr>
<tr>
<td>Mechanical Reasoning</td>
<td>.24</td>
<td>.38</td>
</tr>
<tr>
<td>Clerical Speed and Accuracy</td>
<td>.32</td>
<td>.15</td>
</tr>
<tr>
<td>Language Usage</td>
<td>Scores not available</td>
<td></td>
</tr>
</tbody>
</table>

Multiple Correlation
VR and NA vs. Rank .63 .61

Ferguson\textsuperscript{70} uses a quotation from a book by Strong\textsuperscript{71} as a springboard for his analysis of "ability, interest and aptitude". Ferguson begins by acknowledging that there are formal definitions for ability, interest and aptitude; but immediately adds: "We still have much to learn about the ways they interact and affect, or affect behavior. Does a person first, have ability and then develop interest? Or does he, first have interest and then develop ability? Or does he, first have aptitude, and then develop interest and/or ability? Can ability be defined in terms of interest and aptitude? Or can we define interest in terms of ability and aptitude? Or aptitude in terms of ability and interests?"

Using a geometric presentation of correlations for his data, he proceeds to explain previous research done over the last five years, and concludes that logic leads to the belief that "interest is to be defined in terms of ability and aptitude."

Carter\textsuperscript{72} investigated variables in pupils, other than mental ability, which enter into the assignment of marks by teachers of beginning algebra. The factors especially investigated were socio-economic status, interest, and personality. Working with a sample of 235 pupils (135 boys and 100 girls), he administered the following battery: Otis Quick Scoring Mental Ability Test, the


\textsuperscript{71} E. K. Strong, Vocational Interests 18 Years After College, Minneapolis: University of Minnesota Press, 1955, 155. "Perplexing to many a psychologist has been the determination of the exact inter-relationships among ability, interest and aptitude."

Colvin-Schramm Algebra Test, the California Test of Personality, the Sims Score Card for Socio-Economic Status, and the Garretson-Symonds Interest Questionnaire. The correlation coefficients he obtained for the total group are reported as follows:

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.39</td>
<td>.51</td>
<td>.18</td>
<td>.14</td>
<td>-.31</td>
<td>.16</td>
</tr>
<tr>
<td>2</td>
<td>.52</td>
<td>.13</td>
<td>.16</td>
<td>.22</td>
<td>-.13</td>
<td>.13</td>
</tr>
<tr>
<td>3</td>
<td>-.32</td>
<td>.15</td>
<td>.17</td>
<td>.45</td>
<td>.07</td>
<td>.07</td>
</tr>
</tbody>
</table>

.129 is significant at the five percent level.
.169 is significant at the one percent level.

His general findings indicate that teachers' marks represent achievement, but they give evidence of the effects of intelligence, the socio-economic status, the personality, the interests, and the age of the student upon the teacher. Of these variables, the first three were more influential than the last two.

D'Amico, Bryant and Prohl collaborated in an investigation of the relationship between Multiple Aptitude Test scores and achievement in Junior College subjects. Approximately 450 Flint Junior College students were used as basis for determining the predictive validity of the nine tests of the Multiple

---

Aptitude Test battery. They concluded, in general, that Multiple Aptitude Test scores have from moderate to high validity in relation to achievement in Junior College subjects. Test 3 (Language Usage) of the MAT contributed significantly to prediction of academic success in Biology, Psychology, Foreign Languages, English, General Biology and Drawing. Test 6 (Arithmetic Computation) of the MAT was next highest in its significant contribution to the prediction of success in Mathematics, Social Sciences, Accounting and Business Mathematics. These researchers sum up their findings by saying: "It may be concluded that Multiple Aptitude Test scores have sufficiently high validities to warrant use for prediction of achievement in Junior College Subjects."

An extensive follow-up study recently completed by SRA, involved 517 graduates from Oregon high schools who were enrolled in a large university in Oregon. These students had taken the National Merit Scholarship Qualifying Test in 1958 while juniors in high school. On entering college in the fall of 1959, they took a battery of tests including the Minnesota Scholastic Aptitude Test, the Cooperative Mechanics of Expression Test, and the Cooperative Elementary Algebra Test. In addition, an over-all quality rating of high school grades and first-semester freshman grade-point averages were obtained for each student in the sample. These variables, including the six NMSQT scores, were intercorrelated. The table on the following page shows the intercorrelations.

---

<table>
<thead>
<tr>
<th>Test and Measures</th>
<th>Intercorrelations</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. NMSQT English Usage</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2. NMSQT Math Usage</td>
<td></td>
<td>33</td>
<td>36</td>
<td>45</td>
<td>28</td>
<td>67</td>
<td>45</td>
<td>42</td>
<td>35</td>
<td>74</td>
<td>35</td>
</tr>
<tr>
<td>3. NMSQT Social Studies</td>
<td></td>
<td>54</td>
<td>36</td>
<td>63</td>
<td>58</td>
<td>81</td>
<td>60</td>
<td>53</td>
<td>51</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>4. NMSQT Natural Sciences</td>
<td></td>
<td>44</td>
<td>45</td>
<td>63</td>
<td>61</td>
<td>82</td>
<td>58</td>
<td>56</td>
<td>41</td>
<td>42</td>
<td>34</td>
</tr>
<tr>
<td>5. NMSQT Word Usage</td>
<td></td>
<td>60</td>
<td>28</td>
<td>58</td>
<td>61</td>
<td>77</td>
<td>74</td>
<td>64</td>
<td>56</td>
<td>25</td>
<td>39</td>
</tr>
<tr>
<td>6. NMSQT Composite</td>
<td></td>
<td>73</td>
<td>67</td>
<td>81</td>
<td>82</td>
<td>77</td>
<td>78</td>
<td>70</td>
<td>66</td>
<td>55</td>
<td>51</td>
</tr>
<tr>
<td>7. Minn. Scholastic Apt. Test</td>
<td></td>
<td>67</td>
<td>45</td>
<td>60</td>
<td>58</td>
<td>74</td>
<td>78</td>
<td>85</td>
<td>70</td>
<td>48</td>
<td>51</td>
</tr>
<tr>
<td>8. MSAT Reading</td>
<td></td>
<td>55</td>
<td>42</td>
<td>53</td>
<td>56</td>
<td>64</td>
<td>70</td>
<td>85</td>
<td>58</td>
<td>48</td>
<td>45</td>
</tr>
<tr>
<td>9. Coop. Mechan. of Exp.</td>
<td></td>
<td>77</td>
<td>35</td>
<td>51</td>
<td>41</td>
<td>56</td>
<td>66</td>
<td>70</td>
<td>58</td>
<td>34</td>
<td>58</td>
</tr>
<tr>
<td>11. High School Rating</td>
<td></td>
<td>55</td>
<td>35</td>
<td>36</td>
<td>34</td>
<td>39</td>
<td>51</td>
<td>51</td>
<td>45</td>
<td>58</td>
<td>32</td>
</tr>
<tr>
<td>12. College GPA</td>
<td></td>
<td>49</td>
<td>37</td>
<td>41</td>
<td>37</td>
<td>41</td>
<td>53</td>
<td>56</td>
<td>50</td>
<td>50</td>
<td>36</td>
</tr>
</tbody>
</table>

**NOTE:** All decimal points have been omitted from the correlations.

* These scores are reported in stanines.

Many interesting relationships are revealed by these data. The NMSQT composite score and English Usage score prove to be fairly good predictors of first-semester freshman college grade point averages almost two years later. The NMSQT composite score predicts college grade point averages about as well as the Minnesota Scholastic Aptitude Test and the Cooperative Mechanics of Expression Test, even though the latter tests were administered some seventeen months later than the NMSQT. The NMSQT also predicts college grade-point averages as well as does the quality rating of high school grades, which was determined some fourteen months after the NMSQT testing.

The study goes on to say that the prediction of college grade-point average can be improved by using both the NMSQT composite score and high school rating. The multiple correlation of these two predictors with grade-point average is
increased to .62. This is higher than when the Minnesota Scholastic Aptitude Test and the MMCT composite score are used as predictors.

The New York State Education Department recently made a research study on the comparability of scores of the examinations utilized by the three programs which reach the largest numbers of senior high school students in the state: its own Regents Scholarship (RSE), the College Entrance Examinations Board's Scholastic Ability Test (SAT) and the National Merit Corporation's Scholarship Qualifying Test (NSQT). The purpose of the study was to explore the possibility of determining some method of equivalent scores, so that the college admissions officer, for example, might reasonably accept an applicant's scores on one examination in lieu of those on another examination.

Using a sample of 1,132 students who took all three tests, they arrived at a table of equivalent scores computed by the equi-percentile method. They stress the fact that because one score is equivalent to another this does not mean that the one predicts the other; but only that the scores are equally good scores on the respective tests in which they were achieved. (A bowling score of 290 is as great an achievement as a batting average of .400 in baseball, but one does not predict the other.) However, in tests with high intercorrelations, a student's score on one test will ordinarily be close to his equivalent score on the other.

To find the extent to which the three tests measure the same traits, the correlations in the following table were corrected for attenuation. It was

---

The table below presents the uncorrected Pearson intercorrelations of these tests:

<table>
<thead>
<tr>
<th>TEST</th>
<th>RSE</th>
<th>SAT-V</th>
<th>SAT-M</th>
<th>SAT-V+M</th>
<th>NMSQT</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSE</td>
<td>0.87</td>
<td>0.74</td>
<td>0.90</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>SAT-V</td>
<td>0.87</td>
<td>0.60</td>
<td>0.89</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>SAT-M</td>
<td>0.74</td>
<td>0.60</td>
<td>0.89</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>SAT-V+M</td>
<td>0.90</td>
<td>0.89</td>
<td>0.89</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>NMSQT</td>
<td>0.83</td>
<td>0.79</td>
<td>0.67</td>
<td>0.81</td>
<td></td>
</tr>
</tbody>
</table>

found that there is considerable overlap in the factors being measured by the three tests.

The findings of this study indicate that colleges and other users of any of these tests may well contribute to a lessening of the problem by accepting test score "legal tender" where feasible. The study indicates that there is statistical justification for using interchangeable test score currency. If users can see their way clear to adopting flexible but not equivalent criteria, and if those who deal with the secondary school student can guide him into a proper approach regarding the preparation for the taking of external tests, the problem of multiplicity of testing would be greatly lessened.

Tests of "general educational development" usually show substantial positive correlations with tests of general mental ability, and many students whose scores on the NMSQT are low will be identified simply as students who would not be expected to show major strengths in their educational development. Even with these students, however, the profile should be reviewed for any positive evidence of relative strengths. Such evidence should be considered
for possible implications for the remainder of the student's high school program and his post-high school plans.

Although the NMSQT tests are called measures of educational development, they are similar in many respects to the usual tests of scholastic aptitude. An analysis of the questions in various scholarship and college admission tests reveals that many of the questions are similar or even identical in form and purpose. This observation is supported by research literature. The NMSQT is actually an adaptation of the Iowa Tests of Educational Development (ITED), as is likewise true of the United States Armed Forces Institute General Educational Development Tests (USAFT-GED). It is interesting to report that the ITED battery, the USAFT-GED tests and the NMSQT usually show a correlation of .70 to .90 with other explicit tests of scholastic aptitude, such as the College Board SAT, the Ohio State University Psychological Test, and the ACE Psychological Examination. There are, however, some important differences from tests of scholastic aptitude.

CHAPTER III

METHOD OF INVESTIGATING THE DATA

While the literature abounds with studies aimed at predicting various types of success using a variety of instruments in select combinations, the author could find none that specifically investigated the predictive value of the four California Test Bureau instruments studied herein, in relation to academic success during high school. Patterson,¹ after reviewing a considerable number of studies, believed that it would be possible to select a battery of tests which would combine to yield fair prediction of success in any public or private school. He suggests that the nature of such a battery might consist of a verbal intelligence test, a test of mechanical information, a test of spatial ability, and possibly an interest test. Racky² increased the number of predictor measures and used the battery to derive a multiple regression equation for the prediction of success in ninth grade woodshop. However, the present author was primarily interested in prediction of academic success; and since comparable tests published by a single Test Bureau had already been administered to a sizeable group of young men, and these results were available,


² Donald Racky, Aptitude and Interest Measures Predictive of Ninth Grade Woodshop Performance, Unpublished Doctoral Dissertation, Loyola University, 1959, 76.
he felt that an investigation as to their total and integrated prediction value in relation to proximate and remote academic success would be of value to administrators and guidance personnel. The so-called immediate academic success was regarded to be the student's freshman yearly grade average; and the future academic success was limited specifically to the sophomore, junior, senior Yearly Grade Averages and to the six scores he would achieve as a junior on the National Merit Scholarship Qualifying Test.

Description of the Variables

With this objective in mind, five independent variables were considered in this study to predict academic successes. All but one of the measures used in this study were administered in October of the respective freshman years to all students in the sample. That one, number five below, was obtained approximately in January of the freshman year. These measures were as follows:

1. Intelligence
2. Personality
3. Occupational Interests
4. Various Aptitudes
5. Character Rating by Teachers

The criterion variables or, more simply, the criteria, were the variables to be predicted. As already mentioned, these were the freshman, sophomore, junior and senior Yearly Grade Averages, and the six scores obtained during the junior year in the National Merit Scholarship Qualifying Test.

Collection of the Data

The students whose test measures and other evaluation scores constitute
the data for this study had already graduated from Mendel Catholic High School when this investigation began. Hence, properly speaking, all data is of a secondary-source nature. A competent psychologist had personally administered and scored the following tests:

- The California Test of Personality
- The California Occupational Interest Inventory
- The California Multiple Aptitude Tests
- The California Short Form Test of Mental Maturity
- The Character Rating by Teachers was obtained by averaging the specific ratings by current teachers of each student. The traits measured were:
  - Initiative
  - Industry
  - Loyalty
  - Cooperation
  - Sociability

Each teacher graded the student as high, average, or low in each of these traits. A numerical score was obtained by assigning a value of one for low; a value of two for average; and a value of three for high. These data on file in the Guidance Office.

All data used as criteria was on file among the permanent records of the school. Yearly grade averages were obtained by an arithmetic average of the semester grades for each year.

The grading system employed at Mendel Catholic High School makes use of a letter grade and its numerical range. Although no yearly grade averages were assigned a plus or minus value, the regression value found for an individual
may fall within such a plus or minus range. A true interpretation of a predicted score must be defined within certain limits for each letter grade. These limits as well as the letter grade and a description of the grade are shown in Table I.

**TABLE I**

**GRADE EQUIVALENTS**

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Description</th>
<th>Limits of Numerical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Excellent</td>
<td>93 - 100</td>
</tr>
<tr>
<td>B</td>
<td>Good</td>
<td>85 - 92</td>
</tr>
<tr>
<td>C</td>
<td>Fair</td>
<td>77 - 84</td>
</tr>
<tr>
<td>D</td>
<td>Passing</td>
<td>70 - 76</td>
</tr>
<tr>
<td>E</td>
<td>Conditional</td>
<td>70 (only at first semester)</td>
</tr>
<tr>
<td>F</td>
<td>Failure</td>
<td>0 - 69</td>
</tr>
</tbody>
</table>

The six scores for each student on the National Merit exam were gathered from the permanent records. It is well to emphasize that different segments of the sample wrote National Merit Scholarship Qualifying Tests in four successive years. This reveals the likelihood of shifting norms. In Table II, an estimation of the limits of the National Merit Scholarship Qualifying Tests total scores is presented with their equivalent location on the normal curve.
### Table II

**National Merit Scholarship Qualifying Test Score Equivalents**

<table>
<thead>
<tr>
<th>Total Standard Score</th>
<th>Quarter</th>
<th>Quartile</th>
<th>Limits of Total Standard Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 - 36</td>
<td>Highest</td>
<td></td>
<td>19 - 36</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>Third</td>
<td></td>
</tr>
<tr>
<td>16 - 18</td>
<td>High Middle</td>
<td></td>
<td>15 - 19</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>Second</td>
<td></td>
</tr>
<tr>
<td>13 - 14</td>
<td>Low Middle</td>
<td></td>
<td>12 - 15</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>First</td>
<td></td>
</tr>
<tr>
<td>0 - 11</td>
<td>Lowest</td>
<td></td>
<td>0 - 12</td>
</tr>
</tbody>
</table>

Standard Error = 1.

Measures of the predictor and criterion variables revealed fifty-four factor scores, distributed as follows:

**Intelligence (Five Factor Scores):**


**Personality (Fifteen Factor Scores):**

- Self-Reliance, Personal Worth, Personal Freedom, Feeling of Belonging, Withdrawal Tendencies, Nervous Symptoms, Self-Adjustment (Sub-total score), Social Standing, Social Skills, Anti-Social Tendencies, Family Relations, School Relations, Community Relations, Social Adjustment (Sub-total score),
and Total Adjustment (Total score).

Occupational Interests (Ten Factor Scores):

Personal-Social, Natural, Mechanical, Business, Arithmetic, Scientific, Verbal, Manipulative, Computational and Level of Interest.

Various Aptitudes (Nine Factor Scores):


Character Rating by Teachers (Five Factor Scores):

Initiative, Industry, Loyalty, Co-operation and Sociability.

Yearly Grade Averages (Four Factor Scores):

Freshman, Sophomore, Junior and Senior.

National Merit Scholarship Qualifying Test: (Six Factor Scores)

English Usage, Mathematics Usage, Social Studies Readings, Natural Sciences Readings, Word Usage and Total Score (an arithmetic average of the five scores).

An appropriate 6 x 9 card was designed to exhibit all available raw scores for each student qualifying for the sample. After eliminating incomplete or otherwise faulty data, a final sample of two hundred and ten students was obtained.

The opportunity presented itself of having an IBM-650 computer calculate all possible correlations; and hence the data had to be re-copied on standard I.B.M. punch cards. Since the machine could handle only fifty factor-scores, a
decision had to be made as to which scores would be eliminated from the study. It was decided to use I.Q. alone, as the measure of intelligence, thus eliminating the factor scores of Verbal Concepts, Numerical Reasoning, Logical Reasoning and Spatial Relationships. The reason for this choice was the fact that it was felt that these measures were part of the Multiple Aptitude Tests.

The machine calculated coefficients of correlation revealed existing relationships between all predictor factors and the criteria factors for this specific sample of two hundred and ten freshmen at Mendel. As a visual aid in analyzing patterns of significant correlation, all one thousand two hundred and twenty-five coefficients were displayed in matrix form.

From visual evidence, the decision was made to employ only Freshman Yearly Grade Average and Total Score on the National Merit Battery as criteria, against which to test the validity of prediction from particular combinations of predictor factors. Again, from visual evidence, it was decided that for both criteria, greatest hope rested with I.Q. and certain factors of the Multiple Aptitude battery. As integrated predictors of Freshman Yearly Grade Average, the following factors were selected: Arithmetic Computation, I.Q., Arithmetic Reasoning, Spatial Relations (3-Dimensional), Paragraph Meaning and Language Usage. For the prediction of Total Score on the National Merit battery the following factors appeared most valid: Word Meaning, I.Q., Language Usage, Paragraph Meaning, Arithmetic Computation and Arithmetic Reasoning. The fact that all final predictor-factors were from two California Test Bureau instruments, was seen as an advantage - both in simplicity and economy, should the prediction equations prove significant.

It often happens that a limited number of variables can be selected that
that will have a multiple $R$ almost as high as that of a total group of predictors. DuBois' approach to this process appeared practical; and hence, his method$^2$ was employed. DuBois has presented a routine in matrix format for computing multiple $R$, together with notation indicating the statistical meaning of each cell entry. The steps implicitly indicated have counterparts in explicit formulas in common or scalar notation, which the matrix format replaces. Multiple $R$'s are obtained from the matrix routine by the reduction of criterion variances.

This method of multiple correlation is adjustable to incorporate the feature of identifying at any stage the variable which will make the greatest contribution to the multiple. This procedure tends toward a selection of a limited number of variables having highest multiple correlation with the criterion. In this selection process, variables which duplicate the functions of variables already chosen tend to drop out. Selection is accomplished by a process essentially identical with that used in the Wherry-Doolittle method, but within the framework of the DuBois method.

By beginning with the factor exhibiting the highest correlation with the criterion variable, and adding successively the next highest factors, while examining the resultant coefficients of correlation, one can discern the relative advantage of eliminating certain predictor variables.

For the purpose of this study, no predictor variables were dropped completely. In instances where one or more could reasonably be dropped, a

---

complete multiple regression equation in the original six variables was still
determined and exhibited; then alternate reduced equations were calculated to
embody the advantage of having fewer predictor variables.

Employing the predictor variables in the combinations mentioned above,
multiple regression equations were developed for the prediction of Freshman
Yearly Grade Average and Total Score on the National Merit battery.

In order to test the predictive value of these equations, a random selection of student-data cards was made from the original sample period. Using
the specific raw scores found thereon as values for the particular variables of
the equations, predicted scores were determined.

Since a true interpretation of a predicted score must be defined within
certain limits, all predicted scores were interpreted in the light of the
pertinent limits shown in Tables I and II.

Cross-Validation

In order to test the validity of the obtained multiple regression equations
with their respectively differentiated beta weights for each of the predictor
variables, a new sample of student data was obtained from the files of Mendel Catholic High School. The sixty-two individuals of this sample graduated from Mendel one year later than the most recently graduated boys in the original sample. This sample differed in its composition from the original, in that all of them wrote exactly the same National Merit Scholarship Qualifying Test; whereas elements of the original sample were spread over four different National Merit Scholarship Qualifying Tests. There are other differences among the two

samples which will be explained in a later chapter.

The multiple regression equations obtained from the original sample were applied to students having the same card numbers in the validation sample. A comparison of these predicted scores was then made with the permissible ranges as shown in Tables I and II.

Tests for Significance Between Original Group and Validation Group

In order to ascertain significance levels regarding the comparability of the two samples, various statistical tests were applied, involving the measures of central tendency. Principal among these were: The $t$ - Test, the $F$ - Ratio, and the total variances of the related multiple regression equations.
CHAPTER IV

ANALYSIS OF THE DATA

In order to meaningfully interpret the one thousand two-hundred and fifty coefficients of correlation displayed on the matrix, it was necessary first to determine minimum figures for the commonly used significance levels. The size of the sample being two hundred and ten, it was determined that for a correlation to be significant at the five percent level of confidence, it had to be .1360. A correlation of .1780 yielded significance at the one percent level of confidence.

RELATIONS BETWEEN PREDICTOR VARIABLES AND CRITERIA

N.M.S.Q.T. - English Usage Score

Seventeen predictor factors correlated significantly with this score above the one percent level; five additional factors were significant above the five percent level, while eighteen failed to meet the test of significance even at the five percent level. The independent predictor variable having the greatest number of factors significant above the one percent level was Aptitude. Eight of its nine factors were thus significant. Character Rating by Teachers was next having all five factors significant above the one percent level. Intelligence and three factors of the Occupational Interest variable also correlated above the one percent level of significance.
N.M.S.Q.T. - Mathematics Usage Score

Fourteen predictor factors emerged as significant above the one percent level of confidence, distributed as follows: All nine factors of the Multiple Aptitude battery, I.Q., one of the Occupational Interest factors, (Scientific Interest), and three of the five factors comprising Character Rating by Teachers. Three additional factors were significant at the five percent level, while twenty-three failed to meet the test of significance even at the five percent level.

N.M.S.Q.T. - Social Studies Reading Score

Herein twelve predictor factors met the test of significance at the one percent level; four additional scores were significant at the five percent level, while twenty-four failed to correlate significantly even at the five percent level. The independent predictor variable having the greatest number of factors significant above the one percent level was again that of Aptitude. Six of its factor scores were thus significant. Character Rating by Teachers had four factor scores similarly significant; and the others were I.Q. and the Scientific factor of the Occupational Interests battery.

N.M.S.Q.T. - Natural Science Readings

Again, all nine of the multiple Aptitude factor scores correlated significantly at the one percent level with this criterion. A total of fifteen thus correlated and were distributed as follows: I.Q., the Scientific factor of the Occupational Interests battery and Level of Interests, along with three factors of the Character Rating by Teachers. Four additional factors met the test of significance at the five percent level.
N.M.S.Q.T.- Word Usage Score

There were sixteen factors correlating significantly at the one percent level with this criterion. Six were factors of the Multiple Aptitude battery; and all five of the Character rating scores were in this category; and four factors of the Occupational Interests battery along with I.Q. likewise qualified. Four additional scores were significant at the five percent level, while eighteen failed to meet the test of significance even at the five percent level.

N.M.S.Q.T. - Total or Composite Score

As a consequence of visual analysis of the correlation matrix and other reasons, this composite score was ultimately selected as one of the two final criterion variables. Nineteen factor scores correlated significantly with it at the one percent level. Nine of these were the factor scores of the Multiple Aptitude battery. The five factors of Character Rating likewise qualified. The remaining five were I.Q. and four factors of the Occupational Interests battery. Three additional factors met the five percent level, while the remaining eighteen did not.

Freshman Yearly Grade Average

The only predictor scores which correlated significantly above the one percent level with this criterion were I.Q., six factors of the Multiple Aptitude battery, and three factors of Character Rating. With the exception of two factors of the Occupational Interests battery, no other score met the test of significance even at the five percent level. For reasons which will now be given, this variable was selected as the second of the two final criteria used in this study. The remaining three possible criteria, Sophomore, Junior and Senior Grade Averages had to be rejected, since, with the exceptions of two I.Q.
scores, and four Multiple Aptitude scores which met the test of significance at the five percent level, no other scores showed any significant relationships. No one of these possible criteria revealed more than two predictor factors significant at the five percent level.

In summary, I.Q. correlated significantly far above the one percent level with all six scores of the National Merit Test. Its highest correlation, .6593, was with the Total Score of the N.M.S.Q.T.; its lowest was .4530, with the Social Science Reading factor. Only one factor of the Personality battery was able to meet even the five percent level of confidence. The Scientific factor of the Occupational Interest battery was the only score therein to correlate significantly at the one percent level with all six N.M.S.Q.T. scores. The highest of these correlations was .3305, with the Natural Science Reading factor. The range of the many significant correlations between the Multiple Aptitude factors and the six scores of the N.M.S.Q.T. was .1406 - .7821. Only three factors of the Multiple Aptitude battery failed to meet even the five percent level of significance with all scores of the N.M.S.Q.T. These were the Routine Clerical Facility with the Word Usage score, the Applied Science and Mechanics with the Social Science Readings score, and the Spatial Relations (three dimensional) with the Social Readings score. Among the N.M.S.Q.T. criterion variables, the one having the highest average correlation with the nine factors of the Multiple Aptitude battery was the Total or Composite score. This was an added reason for selecting it as one of the final criteria for this study. The greatest reason, however, was the fact that, being a composite score, it comprehended the implications of all factor scores of the N.M.S.Q.T.
While all of the Character Rating factors correlated significantly at least above the five percent level, it was decided not to use any of them as final predictor variables. The reason for this was the fact that their correlations did not appear high enough. Added to this fact was a personal awareness of the manner in which some teachers evaluated the various traits comprising the rating scale. Another, perhaps weaker, reason was the awareness that to include them might involve over-extending what now appeared to be an already unified and adequate investigation. Aptitude factors presented a unified and more significant approach to the prediction of the two final criterion variables selected on their own merits.
SUMMARY OF INTERCORRELATIONS WITHIN EACH PREDICTOR BATTERY

In general, all of the predictor batteries manifested significantly high intercorrelations (above the one percent level).

California Test of Personality

This test purports to measure two general facets of personality, vis. Self Adjustment and Social Adjustment. Six presumed factors constituting each of these are investigated by means of distinct test items. Hence the California Test of Personality provides twelve factor scores under two headings. The total of each of these two categories constitutes sub-total scores. The total of both of these categories makes up the individual's Total Personality Adjustment score. Counting the sub-total scores and the combined total score, there are sixteen scores included.

The highest intercorrelation among all these measures was .8800 and the lowest was .1323. The median correlation was .4317. The total range was calculated to be .7478. The total range of the six Self-Adjustment factors was .1891, since the highest intercorrelation was .7346 and the lowest, .5456. The intercorrelations of the Six Social Adjustment factors revealed a high of .5235 and a total range of .2135. Self Adjustment correlated with Social Adjustment at .7165; and each of these correlated with the Total Personality Adjustment score at .8654 and .8800 respectively.

Occupational Interest Inventory

A total of ten scores emerge from this battery. The Level of Interest score is somewhat of a total score arrived at by a formula involving the three scores measuring Types of Interest. The remaining six scores are under the
heading of Fields of Interest.

The total range among the intercorrelations of Fields of Interest factors was .5156; the highest correlation being -.5894, between Personal-Social Interests and Mechanical Interests. The correlation between Mechanical Interest and Scientific Interest appeared as the least, viz., -.0739.

Types of Interest scores intercorrelated from a high of .4227 to a low of .2716. Verbal and Computational correlated highest, while Verbal and Manipulative showed up as lowest.

The intercorrelations between Fields of Interest and Types of Interest ranged from a high of .9002 (Business Interest with Computational Interest) to a low of .1040 (Business Interest with Manipulative Interest).

Level of Interest correlated highest with Scientific Interest (.4139); and lowest with Business Interest (.0374).

The median correlation for all the scores was .2676. Approximately half of the intercorrelations were negative.

Multiple Aptitude Tests

As has already been indicated, certain factors of this battery, along with I.Q. score, were finally selected as the predictor variables for this study. It comprises nine distinct tests, each yielding a factor score. Of the thirty-six intercorrelations emerging, twenty-seven were above the one percent level of significance, six were above the five percent level and three were insignificantly related.

The highest relationship revealed was between Word Meaning and Paragraph Meaning (.6909) and the lowest was between Routine Clerical Facility and Applied Science and Mechanics (.0217). The median correlation was .2652.
Character Rating by Teachers

All intercorrelations between the traits measured by this Scale were far above the one percent level of significance. As the author has indicated earlier, he is inclined from personal experience to regard the validity of these evaluations with suspicion. Experience has similarly indicated that they tend to be reliable.

The highest intercorrelation was between Initiative and Industriousness (.7720), while the lowest turned out to be .4592, between Initiative and Sociability. The median correlation was .6408.

SUMMARY OF INTERCORRELATIONS BETWEEN CRITERION VARIABLES

National Merit Scholarship Qualifying Test

This battery is comprised of five distinct tests each yielding a score, and a composite score which is the average of the five scores. Without exception, the most significant intercorrelations occurred between each of the test scores in conjunction with the composite score.

The intercorrelations between all six scores of this battery ranged from .8246 to .4579. The median correlation was .6099. The highest correlation was between Natural Science Reading and the composite score; the lowest involved Word Usage and Mathematics Usage.

Yearly Grade Averages

Herein were discovered the highest intercorrelations of the entire matrix. The range went from nearly perfect correlation (.9988) to the very high correlation of .8871. The median correlation was .9528.
RESUME OF CORRELATIONS BETWEEN THE PREDICTOR BATTERIES

Intelligence Quotient and Test of Personality

No factor of the Personality battery met the test of significance even at the five percent level of confidence. The highest correlation was .1072, while the lowest was .0056. The median correlation was calculated as .0258.

Intelligence Quotient and Occupational Interest Inventory

Scientific Interest was the only factor significant at the one percent level; and no other factor met even the five percent level. Scientific Interest correlated at .3123; and the lowest correlation appeared to be .0126. The median correlation was .0850.

Intelligence Quotient and Multiple Aptitude Tests

All nine factors of the Multiple Aptitude battery correlated significantly above the one percent level. The highest correlation was with the aptitude, Word Meaning, (.6305), while the lowest was .2413 (Routine Clerical Facility). The median correlation was calculated to be .5264.

Intelligence Quotient and Character Rating by Teachers

Initiative and Industriousness correlated significantly at the one percent level; and one other, Loyalty, met the test of significance at the five percent level. The total range of all five correlations was .0993, the highest correlation being .1958. The median correlation was .1368.

California Test of Personality and Occupational Interest Inventory

Only ten correlations between factors of the Test of Personality and the Interest Inventory were significant at the one percent level. Twenty-one met the test for significance at the five percent level; and the remaining one hundred and eighteen correlations were insignificant. The total range for all
correlations was .2548, with the highest correlation being .2574. The median correlation was .0844.

**California Test of Personality and Multiple Aptitude Tests**

Social Standing correlated significantly at the five percent level with Applied Science and Mechanics of the Aptitude battery. Otherwise, all correlations were insignificant. The highest correlation was .1597 and the total range was .1536. The median correlation was .0493. The total number of correlations was one hundred and thirty six.

**California Test of Personality and Character Rating by Teachers**

One of the correlations was significant at the one percent level; one met the test of significance at the five percent level, and the remaining seventy-three were not significant. The highest correlation was .1731; and the lowest showed up as .0010. The median correlation was calculated to be .0437.

**Occupational Interests Inventory and Multiple Aptitude Tests**

Sixteen correlations between these batteries were significant at the one percent level; twelve met the test of significance at the five percent level, while the remaining sixty-two correlations appeared as insignificant. The total range of correlations was .3710, with the upper correlation being .3714. The median was calculated to be .0776.

**Occupational Interest Inventory and Character Rating by Teachers**

Only one factor met even the five percent level. This was Manipulative Interest correlated with the trait of Cooperation. Its value, .1466, along with the lowest correlation, (Business Interest with the trait of Loyalty) .0032, yielded a total range of .1435. The median correlation was .0448. There were fifty correlations involved.
Multiple Aptitude Tests and Character Rating by Teachers

Fifteen correlations met the test of significance at the one percent level; six reached the five percent level, and the remaining twenty-four failed to meet the test of significance even at the five percent level. The highest correlation was .2713; and the range was calculated to be .2457, and the median correlation was .1285.

RESUME OF CORRELATIONS BETWEEN PREDICTOR BATTERIES AND THE CRITERIA BATTERIES

Intelligence Quotient and National Merit Battery

All six correlations are considerably above the one percent level. The highest correlation was .6503 and the lowest was .4530, yielding a range of .2064. The median correlation was calculated to be .5736.

Intelligence Quotient and Yearly Grade Averages

Intelligence significantly correlated at the one percent level with the Freshman Yearly Grade Average. It did so at the five percent level for the Sophomore and Senior Yearly Grade Averages. Its correlation with the Junior Yearly Grade Average was the lowest, viz., .1263. The highest correlation was .3018. The median correlation was .1440.

California Test of Personality and National Merit Battery

Of the ninety correlations involved, only one met even the five percent level. This was School Relations with the factor of Word Usage. The range is indicated by the limits of the correlations, .1453 - .0006. The median correlation was .0533.

California Test of Personality and Yearly Grade Averages

No significant correlations appeared between personality factors and any
of the yearly grade averages. An examination of the sixty correlations revealed
a range from .1163 - .0019, and a median correlation of .0384.

Occupational Interest Inventory and National Merit Battery

Fifteen correlations significant at the one percent level and fourteen
which met the five percent level constitute the significant relationships
between these two batteries. The highest among the sixty correlations was
.3634, and the least was .0102. The median correlation was .1208.

Occupational Interest Inventory and Yearly Grade Averages

Only two correlations were able to meet the five percent level. The median
among the forty correlations was .0486, the extreme correlations were .1769 and
.0004.

Multiple Aptitude Tests and National Merit Battery

All except seven correlations appeared significant at the one percent level,
while four additional correlations met the five percent level. Only three
failed to meet even the five percent level of confidence. The range was .7052,
determined by the high of .7821 and the low of .0769. The median was .3774.

Multiple Aptitude Tests and Yearly Grade Averages

Six correlations were significant at the one percent level, while four
additional ones met the test of significance at the five percent level. Twenty-
six correlations failed to meet even the five percent level. The highest
correlation (.3150), was between Arithmetic Computation and Freshman Yearly
Grade Average. The lowest (.0258), appeared between Word Meaning and Junior
Yearly Grade Average. The median correlation was .1036.

Character Rating by Teachers and National Merit Battery

All of the correlations were significant; twenty-five at the one percent
level and the remaining five at the five percent level. The highest correlation was between Industriousness and English Usage, (.3784), and the lowest appeared as Sociability with Natural Science Reading, (.1473). The median correlation was .2518.

Character Rating by Teachers and Yearly Grade Averages

Three correlations, vis., Initiative, (.1834), Industriousness, (.1961), and Loyalty, (.1939) respectively, with Freshman Yearly Grade Average, were significant at the one percent level. Two others met the five percent level. The lowest correlation was .0955 between Sociability and the Junior Yearly Grade Average. The median correlation was .1262. There were twenty correlations in all.

RESUME OF CORRELATIONS BETWEEN CRITERION VARIABLES

National Merit Battery and Yearly Grade Averages

Seven correlations were significant at the one percent level, while five additional correlations met the five percent level. The largest correlation was between English Usage and the Freshman Yearly Grade Average, (.3249). The smallest was between Social Studies Reading and Junior Yearly Grade Average, (.0148). The median correlation was calculated to be .1222. There were twenty-four correlations.
### TABLE III

**COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND FRESHMAN YEARLY GRADE AVERAGES**

<table>
<thead>
<tr>
<th>Test</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>California Short Form Mental Maturity - Total Score</td>
<td>.3018</td>
</tr>
<tr>
<td>California Test of Personality</td>
<td></td>
</tr>
<tr>
<td>Self Reliance</td>
<td>-.0237</td>
</tr>
<tr>
<td>Personal Worth</td>
<td>.0545</td>
</tr>
<tr>
<td>Personal Freedom</td>
<td>.0773</td>
</tr>
<tr>
<td>Feeling of Belonging</td>
<td>.0766</td>
</tr>
<tr>
<td>Withdrawal Tendencies</td>
<td>.0059</td>
</tr>
<tr>
<td>Nervous Symptoms</td>
<td>-.0447</td>
</tr>
<tr>
<td><strong>Self-Adjustment (Sub-Total Score)</strong></td>
<td></td>
</tr>
<tr>
<td>Social Standing</td>
<td>-.0719</td>
</tr>
<tr>
<td>Social Skills</td>
<td>-.1163</td>
</tr>
<tr>
<td>Anti-Social Tendencies</td>
<td>-.0487</td>
</tr>
<tr>
<td>Family Relations</td>
<td>.0679</td>
</tr>
<tr>
<td>School Relations</td>
<td>-.0343</td>
</tr>
<tr>
<td>Community Relations</td>
<td>.0903</td>
</tr>
<tr>
<td><strong>Social Adjustment (Sub-Total Score)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL ADJUSTMENT (Total Score)</strong></td>
<td>-.0063</td>
</tr>
<tr>
<td><strong>California Occupational Interest Inventory</strong></td>
<td></td>
</tr>
<tr>
<td>Fields of Interest</td>
<td></td>
</tr>
<tr>
<td>Personal-Social</td>
<td>.0405</td>
</tr>
<tr>
<td>Natural</td>
<td>-.0989</td>
</tr>
<tr>
<td>Mechanical</td>
<td>-.0229</td>
</tr>
<tr>
<td>Business</td>
<td>-.0408</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>.0113</td>
</tr>
<tr>
<td>Scientific</td>
<td>.1769</td>
</tr>
<tr>
<td>Types of Interest</td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>-.0397</td>
</tr>
<tr>
<td>Manipulative</td>
<td>.1655</td>
</tr>
<tr>
<td>Computational</td>
<td>.0650</td>
</tr>
<tr>
<td>Level of Interest</td>
<td>.0129</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level
0.178 is significant at the one percent level
### Table III (continued)

**COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND YEARLY GRADE AVERAGES**

<table>
<thead>
<tr>
<th>Test</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>California Multiple Aptitude Tests</strong></td>
<td></td>
</tr>
<tr>
<td>Word Meaning</td>
<td>.2070</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>.2392</td>
</tr>
<tr>
<td>Language Usage</td>
<td>.2197</td>
</tr>
<tr>
<td>Routine Clerical Facility</td>
<td>.1066</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>.2596</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>.3150</td>
</tr>
<tr>
<td>Applied Science and Mechanics</td>
<td>.0712</td>
</tr>
<tr>
<td>Spatial Relations (2 dimensional)</td>
<td>.1116</td>
</tr>
<tr>
<td>Spatial Relations (3 dimensional)</td>
<td>.2500</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level

0.178 is significant at the one percent level

### Table IV

**COEFFICIENT OF CORRELATION BETWEEN TEACHERS' CHARACTER RATINGS AND FRESHMAN YEARLY GRADE AVERAGES**

<table>
<thead>
<tr>
<th>Trait</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiative</td>
<td>.1834</td>
</tr>
<tr>
<td>Industry</td>
<td>.1961</td>
</tr>
<tr>
<td>Loyalty</td>
<td>.1939</td>
</tr>
<tr>
<td>Co-operation</td>
<td>.1524</td>
</tr>
<tr>
<td>Sociability</td>
<td>.1620</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level

0.178 is significant at the one percent level
### TABLE V

**COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST TOTAL SCORES**

<table>
<thead>
<tr>
<th>Test</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>California Short Form Mental Maturity - Total Score</td>
<td>.6593</td>
</tr>
<tr>
<td>California Test of Personality</td>
<td></td>
</tr>
<tr>
<td>Self Reliance</td>
<td>-.0390</td>
</tr>
<tr>
<td>Personal Worth</td>
<td>.0365</td>
</tr>
<tr>
<td>Personal Freedom</td>
<td>-.0212</td>
</tr>
<tr>
<td>Feeling of Belonging</td>
<td>.0452</td>
</tr>
<tr>
<td>Withdrawal Tendencies</td>
<td>.0775</td>
</tr>
<tr>
<td>Nervous Symptoms</td>
<td>.1095</td>
</tr>
<tr>
<td><strong>Self Adjustment</strong> (Sub-Total Score)</td>
<td></td>
</tr>
<tr>
<td>Social Standing</td>
<td>.0603</td>
</tr>
<tr>
<td>Social Skills</td>
<td>.0008</td>
</tr>
<tr>
<td>Anti-Social Tendencies</td>
<td>.0553</td>
</tr>
<tr>
<td>Family Relations</td>
<td>.0162</td>
</tr>
<tr>
<td>School Relations</td>
<td>.1010</td>
</tr>
<tr>
<td>Community Relations</td>
<td>.0559</td>
</tr>
<tr>
<td><strong>Social Adjustment</strong> (Sub-Total Score)</td>
<td>.0477</td>
</tr>
<tr>
<td><strong>TOTAL ADJUSTMENT</strong> (Total Score)</td>
<td>.0731</td>
</tr>
<tr>
<td>California Vocational Interest Inventory</td>
<td></td>
</tr>
<tr>
<td><strong>Fields of Interest</strong></td>
<td></td>
</tr>
<tr>
<td>Personal-Social</td>
<td>.1463</td>
</tr>
<tr>
<td>Natural</td>
<td>-.1878</td>
</tr>
<tr>
<td>Mechanical</td>
<td>-.1567</td>
</tr>
<tr>
<td>Business</td>
<td>-.0653</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>.0428</td>
</tr>
<tr>
<td>Scientific</td>
<td>.3634</td>
</tr>
<tr>
<td><strong>Types of Interest</strong></td>
<td></td>
</tr>
<tr>
<td>Verbal</td>
<td>-.0102</td>
</tr>
<tr>
<td>Manipulative</td>
<td>.1925</td>
</tr>
<tr>
<td>Computational</td>
<td>.1571</td>
</tr>
<tr>
<td>Level of Interest</td>
<td>.1805</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level

0.178 is significant at the one percent level
TABLE V (continued)

COEFFICIENT OF CORRELATION BETWEEN PSYCHOLOGICAL TEST SCORES AND NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST TOTAL SCORES

<table>
<thead>
<tr>
<th>Test</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>California Multiple Aptitude Tests</td>
<td></td>
</tr>
<tr>
<td>Word Meaning</td>
<td>.6765</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>.6015</td>
</tr>
<tr>
<td>Language Usage</td>
<td>.6336</td>
</tr>
<tr>
<td>Routine Clerical Facility</td>
<td>.2131</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>.5431</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>.5982</td>
</tr>
<tr>
<td>Applied Science and Mechanics</td>
<td>.2161</td>
</tr>
<tr>
<td>Spatial Relations (2 dimensional)</td>
<td>.2704</td>
</tr>
<tr>
<td>Spatial Relations (3 dimensional)</td>
<td>.2950</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level
0.178 is significant at the one percent level

TABLE VI

COEFFICIENT OF CORRELATION BETWEEN TEACHERS' CHARACTER RATING AND NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST TOTAL SCORES

<table>
<thead>
<tr>
<th>Trait</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiative</td>
<td>.3547</td>
</tr>
<tr>
<td>Industry</td>
<td>.3582</td>
</tr>
<tr>
<td>Loyalty</td>
<td>.2971</td>
</tr>
<tr>
<td>Co-operation</td>
<td>.2541</td>
</tr>
<tr>
<td>Sociability</td>
<td>.2419</td>
</tr>
</tbody>
</table>

0.136 is significant at the five percent level
0.178 is significant at the one percent level
### TABLE VII

**COEFFICIENT OF CORRELATION BETWEEN NATIONAL MERIT SCHOLARSHIP QUALIFYING TEST SCORES AND FRESHMAN YEARLY GRADE AVERAGES**

<table>
<thead>
<tr>
<th>Test</th>
<th>Coefficient of Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Usage</td>
<td>.3249</td>
</tr>
<tr>
<td>Mathematics Usage</td>
<td>.2976</td>
</tr>
<tr>
<td>Social Studies Readings</td>
<td>.1533</td>
</tr>
<tr>
<td>Natural Science Readings</td>
<td>.2346</td>
</tr>
<tr>
<td>Word Usage</td>
<td>.2893</td>
</tr>
<tr>
<td><strong>TOTAL SCORE</strong></td>
<td><strong>.2749</strong></td>
</tr>
</tbody>
</table>

.136 is significant at the five percent level.

.138 is significant at the one percent level.
DuBois has presented a routine in matrix format for computing multiple $R$, together with notation indicating the statistical meaning of each cell entry$^1$. The steps implicitly indicated have counterparts in explicit formulas in common or scalar notation, which the matrix format replaces. Multiple $R$'s are obtained from the matrix routine by the reduction of criterion variance.

This method of multiple correlation can be readily adjusted to incorporate the feature of identifying at any stage the variable which will make the greatest contribution to the multiple, and of using that variable as the next one to be eliminated$^2$. This procedure will tend toward a selection of a limited number of variables having highest multiple correlation with the criterion. Thus, in some cases, a limited number of variables so selected will have a multiple almost as high as that of the total group of predictors, resulting in economy of application to practical situations (of use).

In this selection process, variables which duplicate the functions of variables already chosen, tend to drop out. Selection is accomplished by a process essentially identical with that used in the Wherry-Doolittle Method, but within the framework of the DuBois method. By beginning with the factor exhibiting the highest correlation with the criterion variable, and adding successively the next highest factors, while examining the resultant coefficients of multiple correlation, one can discern the relative advantage of eliminating certain predictor variables.

$^2$Ibid., 36.
DEVELOPMENT OF PREDICTION EQUATION FOR FRESHMAN YEARLY GRADE AVERAGE

Intercorrelations of Predictor Variables

The following Table exhibits the intercorrelations between the predictor variables. The mean and standard deviation of each of the variables were included in the matrix.

TABLE VIII

MATRIX OF INTERCORRELATIONS

<table>
<thead>
<tr>
<th></th>
<th>Arithmetic Computation</th>
<th>Intelligence Quotient</th>
<th>Arithmetic Reasoning</th>
<th>Spatial Relations 2-dim.</th>
<th>Paragraph Meaning</th>
<th>Language Usage</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman YGA</td>
<td>0.3150</td>
<td>0.3018</td>
<td>0.2596</td>
<td>0.2500</td>
<td>0.2392</td>
<td>0.2197</td>
<td>81.995</td>
<td>10.198</td>
</tr>
<tr>
<td>Arithmetic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computation</td>
<td>0.5623</td>
<td>0.6427</td>
<td>0.4065</td>
<td>0.4509</td>
<td>0.4732</td>
<td>25.666</td>
<td>6.923</td>
<td></td>
</tr>
<tr>
<td>I.Q.</td>
<td></td>
<td>0.5589</td>
<td>0.2897</td>
<td>0.5264</td>
<td>0.5449</td>
<td>116.528</td>
<td>14.623</td>
<td></td>
</tr>
<tr>
<td>Arithmetic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reasoning</td>
<td>0.3255</td>
<td>0.4540</td>
<td>0.4214</td>
<td>15.504</td>
<td>7.071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial Relations 3-dim.</td>
<td>0.2613</td>
<td>0.2031</td>
<td>13.035</td>
<td>4.532</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>0.4991</td>
<td>35.200</td>
<td>6.164</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Usage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75.742</td>
<td>15.000</td>
<td></td>
</tr>
</tbody>
</table>

An examination of Table VIII reveals that all variables correlated positively with the others and the range was from 0.2031 to 0.6427. The smallest
correlation, .2031 was between spatial relations 3-dim. and language usage. The largest correlation was between Arithmetic Computation and Arithmetic Reasoning (.6427).

Accordingly, in the investigation of certain factors of the MAT in their relationship to freshman YGA, it was found that a correlation of .3731 was obtained by a 5-predictor battery with the criterion of Freshman YGA. The variables of the predictor battery, in their order of contribution to multiple R, were: Arithmetic Computation, I.Q., Spatial Relation - (3-Dimensional), Paragraph Meaning, and Arithmetic Reasoning.

But when the sixth predictor variable, Language Usage, was added to the routine, the multiple R increased by only .0003. The results of the selection process are shown in Table IX. Actually, the result of adding the last two predictor variables increased the multiple R only .0005. Consequently, a multiple regression equation utilizing four variables appears reasonably accurate for prediction purposes.
TABLE IX

ELIMINATION OF PREDICTOR OF LEAST VALUE BY MEANS OF MULTIPLE CORRELATION OF SIX VARIABLES WITH YEARLY GRADE AVERAGE

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>Coefficient of Multiple Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic Computation, Intelligence Quotient,</td>
<td>.3734</td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional,</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning, Arithmetic Reasoning, Language Usage</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Computation, Intelligence Quotient,</td>
<td>.3731</td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional,</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning, Arithmetic Reasoning</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Computation, Intelligence Quotient,</td>
<td>.3729</td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional,</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Computation, Intelligence Quotient,</td>
<td>.3696</td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Computation, Intelligence Quotient,</td>
<td>.3492</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>.3150</td>
</tr>
</tbody>
</table>

The DuBois method was employed to calculate the differential weights for each of the 6 predictors for use in a multiple regression equation. This equation is shown below:

\[
\text{Predicted grade} = 0.2070 \times (\text{Arithmetic Computation Score}) + \\
0.0958 \times (\text{Intelligence Quotient Score}) + \\
0.2878 \times (\text{Spatial Relations Score, 3-Dimensional}) + \\
0.0863 \times (\text{Paragraph Meaning Score}) + \\
0.0277 \times (\text{Arithmetic Reasoning Score}) + \\
0.0122 \times (\text{Language Usage Score}) + 57.1540
\]
The equation on the previous page can be expressed in simplified form as:

\[ \bar{x} = 0.2070x_6 + 0.0958x_5 + 0.2878x_4 + 0.0863x_3 + 0.0277x_2 + 0.0122x_1 + 57.154 \]

where:

\[ \bar{x} = \text{Predicted Grade} \]

\[ x_6 = \text{Arithmetic Computation Score} \]

\[ x_5 = \text{Intelligence Quotient Score} \]

\[ x_4 = \text{Spatial Relations - 3-Dimensional Score} \]

\[ x_3 = \text{Paragraph Meaning Score} \]

\[ x_2 = \text{Arithmetic Reasoning Score} \]

\[ x_1 = \text{Language Usage Score} \]

Using this equation, predicted grades were calculated for the first three students in the group. These students were the cards numbered 1, 2 and 3. These students received a predicted grade of 90%, 91% and 87% respectively after rounding off decimals. The actual grades received by these students were 84%, 88% and 88% respectively. A comparison of the predicted grades with the "Limits of Numerical Values" as shown in Table I, Chapter III, discloses that the predicted grades for students numbered 2 and 3 fall within the range allowed for the given letter grade. Student 1 failed by one percent to fall within the range allowed for the given letter grade.

By eliminating variable \( x_1 \) (Language Usage Score) from the battery of predictors, the following multiple regression equation was obtained:

\[ \bar{x}' = 0.2070x_6 + 0.0958x_5 + 0.2878x_4 + 0.0863x_3 + 0.0277x_2 + 58.0780 \]

Applying this equation to the same cards described above, predicted scores were obtained as follows: Card 1 received a predicted grade of 90% again after rounding off decimals; Card 2 remained 91% after rounding off decimals;
Card 3 decreased in predicted score to 86%. Only Card 1 missed the range allowed for the given letter grade; and again, it was by one percentage point.

Finally, by eliminating variable \( x_2 \) (Arithmetic Reasoning Score) a multiple regression equation in four unknowns was obtained. The four variable multiple regression equation is presented below, wherein all \( x_1 \) have the same interpretation as above.

\[
\bar{X}'' = 0.2070x_6 + 0.0958x_5 + 0.2878x_4 + 0.0863x_3 + 58.5030
\]

Applying this equation to the same cards, predicted scores of 89%, 90% and 86% respectively were obtained. There was no change in relative positions.

**DEVELOPMENT OF PREDICTION EQUATION FOR TOTAL N.M.S.Q.T. SCORE**

Applying the same techniques to another set of six predictors from the California Multiple Aptitude Tests in relation to the criterion of total score on the National Merit Scholarship Qualifying Test, the following facts became evident. First of all, the elimination test for the variable of least value disclosed that all six predictor variables contributed significantly; and hence were used in the development of a multiple regression equation.

The variables of the predictor battery in their order of contribution to multiple \( R \), were Word Meaning, Intelligence Quotient, Language Usage, Arithmetic Computation, Paragraph Meaning, and Arithmetic Reasoning. The results of the selection routine are shown in Table X.
TABLE X

ELIMINATION OF PREDICTOR OF LEAST VALUE BY MEANS OF MULTIPLE CORRELATION OF SIX VARIABLES WITH N.M.S.Q.T. TOTAL SCORE

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Coefficient of Multiple Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Meaning, Intelligence Quotient, Language Usage, Arithmetic Computation, Paragraph Meaning, Arithmetic Reasoning</td>
<td>.8005</td>
</tr>
<tr>
<td>Word Meaning, Intelligence Quotient, Language Usage, Arithmetic Computation</td>
<td>.7987</td>
</tr>
<tr>
<td>Word Meaning, Intelligence Quotient, Language Usage, Arithmetic Computation</td>
<td>.7930</td>
</tr>
<tr>
<td>Word Meaning, Intelligence Quotient, Language Usage</td>
<td>.7738</td>
</tr>
<tr>
<td>Word Meaning, Intelligence Quotient</td>
<td>.7400</td>
</tr>
<tr>
<td>Word Meaning</td>
<td>.6765</td>
</tr>
</tbody>
</table>

Employing the DuBois method to calculate the differential weight for each of the 6 predictor variables, the following multiple regression equation was obtained:

Predicted score = .1092 x (Word Meaning Score) +
                 .0562 x (Intelligence Quotient) +
                 .0623 x (Language Usage Score) +
                 .1021 x (Arithmetic Computation Score) +
                 .0824 x (Paragraph Meaning Score) +
.0426 x (Arithmetic Reasoning Score) + 

-3.356

The above equation can be expressed in simplified form as:

\[ \bar{X} = .1092x_8 + .0562x_5 + .0623x_4 + .1021x_3 + .0824x_2 + .0426x_1 - 3.356 \]

where:
- \( \bar{X} \) = Predicted Grade
- \( x_8 \) = Word Meaning Score
- \( x_5 \) = Intelligence Quotient Score
- \( x_4 \) = Language Usage Score
- \( x_3 \) = Arithmetic Computation Score
- \( x_2 \) = Paragraph Meaning Score
- \( x_1 \) = Arithmetic Reasoning Score

This equation was then used to predict National Merit Scholarships qualifying Test total scores for the same sample of students. These were cards numbered 1, 2 and 3. These students received a predicted score of 21, 23 and 22 respectively. The actual scores received by these three students were 22, 29 and 22 respectively. A comparison of the predicted scores with the "Limits of Total Standard Scores" as shown in Table II, Chapter III, reveals that the predicted scores fell within the range allowed for the particular quarter.

**Correlation of Predictor Variables**

Table XI shows the intercorrelation between the predictor variables used with the National Merit Scholarship Qualifying Test Total Score. The mean and standard deviation of each of the variables have been included in the matrix.
<table>
<thead>
<tr>
<th></th>
<th>Word Meaning</th>
<th>Intelligence Quotient</th>
<th>Language Usage</th>
<th>Paragraph Meaning</th>
<th>Arithmetic Computation</th>
<th>Arithmetic Reasoning</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N.M.S.Q.T. Total Score</td>
<td>.6765</td>
<td>.6593</td>
<td>.6338</td>
<td>.6015</td>
<td>.5982</td>
<td>.5431</td>
<td>18.038</td>
<td>4.123</td>
</tr>
<tr>
<td>Word Meaning</td>
<td></td>
<td>.6305</td>
<td>.5661</td>
<td>.6709</td>
<td>.4803</td>
<td>.4467</td>
<td>36.128</td>
<td>8.485</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>.5449</td>
<td></td>
<td>.5264</td>
<td>.5623</td>
<td>.5589</td>
<td>116.523</td>
<td>14.628</td>
<td></td>
</tr>
<tr>
<td>Language Usage</td>
<td></td>
<td>.4991</td>
<td></td>
<td>.4732</td>
<td>.4214</td>
<td>75.742</td>
<td>15.000</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td></td>
<td></td>
<td></td>
<td>.4509</td>
<td>.4540</td>
<td>35.200</td>
<td>6.164</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>.6427</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.666</td>
<td>6.928</td>
<td></td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.504</td>
<td>7.071</td>
</tr>
</tbody>
</table>

As evidenced by the table, all variables correlated positively with the other variables; and this correlation was far above the one percent level of confidence. The highest correlation was between the total score of the National Merit Scholarship Qualifying Test and the Word Meaning test of the Multiple Aptitude Tests (.6765). Exhibiting least correlation were the language usage and arithmetic reasoning factors of the Multiple Aptitude Tests (.4214).
Cross-Validation

To test the validity of the obtained multiple regression equations with their differentiated weights for each of their predictors, the equations were applied to a new sample of sixty-two boys enrolled in September, 1958, at Mendel Catholic High School, one year after the latest group included in the original sample.

The coefficient of multiple correlation between the final predictors and the Freshman Yearly Grade Average is presented in Table XII. The sequence of the predictors vary in the several intermediate batteries. The coefficient of multiple correlation for the entire battery was .7061, considerably higher than the multiple $R$ for the original sample.

Applying the multiple regression equation for the prediction of Freshman Yearly Grade Average to the cards in the validation sample, numbered exactly the same as the ones previously tested, the grades of 87%, 84% and 83% were predicted. The actual grades received by these three students were 83%, 84% and 85% respectively. Employing the multiple regression equation for the prediction of Total Score on the National Merit Scholarship Qualifying Test to the very same students' predictor scores, yielded scores of 24, 20 and 19, respectively. The actual scores received were 25, 17 and 19 respectively. The first and third predictions of Freshman Yearly Grade Average barely missed falling in the required range. The second was a completely accurate prediction. The second card used for prediction of the National Merit Total Score barely missed falling in the required range; while the other two met this requirement. Card number three was an accurate prediction.
TABLE XII

COEFFICIENT OF MULTIPLE CORRELATION BETWEEN FRESHMAN YEARLY GRADE AVERAGE AND A BATTERY OF PREDICTORS FOR THE "CROSS-VALIDATION GROUP"

<table>
<thead>
<tr>
<th>Test Battery</th>
<th>Coefficient of Multiple Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery A</td>
<td>.7061</td>
</tr>
<tr>
<td>Battery B</td>
<td>.7046</td>
</tr>
<tr>
<td>Battery C</td>
<td>.7042</td>
</tr>
<tr>
<td>Battery D</td>
<td>.7036</td>
</tr>
<tr>
<td>Battery E</td>
<td>.6365</td>
</tr>
<tr>
<td>Battery F</td>
<td>.5789</td>
</tr>
</tbody>
</table>

**Code**

Battery A: Arithmetic Computation, Arithmetic Reasoning, Paragraph Meaning, Language Usage, Intelligence Quotient, Spatial Relations - 3-Dimensional.


Battery E: Arithmetic Computation, Arithmetic Reasoning.

Battery F: Arithmetic Computation.
TABLE XIII

COEFFICIENT OF MULTIPLE CORRELATION BETWEEN N.M.S.Q.T. TOTAL SCORES AND A BATTERY OF PREDICTORS FOR THE "CROSS-VALIDATION GROUP"

<table>
<thead>
<tr>
<th>Test Battery</th>
<th>Coefficient of Multiple Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery A</td>
<td>.7781</td>
</tr>
<tr>
<td>Battery B</td>
<td>.7775</td>
</tr>
<tr>
<td>Battery C</td>
<td>.7707</td>
</tr>
<tr>
<td>Battery D</td>
<td>.7570</td>
</tr>
<tr>
<td>Battery E</td>
<td>.6832</td>
</tr>
<tr>
<td>Battery F</td>
<td>.6350</td>
</tr>
</tbody>
</table>

Code


Battery E: Paragraph Meaning, Word Meaning.

Battery F: Paragraph Meaning.
The intercorrelations between the variables in the cross-validation group and the criterion, Freshman Yearly Grade Average, along with the means and standard deviations are shown in Table XIV. Except for a slight exception in the criterion (YGA), the mean of each variable was higher in the original sample. The relative importance of these differences will be brought out in the next table.

**TABLE XIV**

**MATRIX OF INTERCORRELATION - FRESHMAN YEARLY GRADE AVERAGE "CROSS-VALIDATION GROUP"**

<table>
<thead>
<tr>
<th></th>
<th>Arithmetic Computation</th>
<th>Intelligence Quotient</th>
<th>Arithmetic Reasoning</th>
<th>Spatial Relations 3-Dimensional</th>
<th>Paragraph Meaning</th>
<th>Language Usage</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman YGA</td>
<td>.5789</td>
<td>.4218</td>
<td>.5699</td>
<td>.3067</td>
<td>.5547</td>
<td>.4235</td>
<td>82.741</td>
<td>6.480</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>.4119</td>
<td>.6288</td>
<td>.2887</td>
<td>.3409</td>
<td>.3193</td>
<td>25.306</td>
<td>6.403</td>
<td></td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional</td>
<td>.5262</td>
<td>.0791</td>
<td>.4601</td>
<td>.4942</td>
<td>109.338</td>
<td>12.529</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>.3923</td>
<td>.5030</td>
<td>.4449</td>
<td>12.145</td>
<td>6.244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Usage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>62.596</td>
<td>14.352</td>
</tr>
</tbody>
</table>
TABLE XV

T TEST FOR SIGNIFICANCE BETWEEN ORIGINAL GROUP
AND FRESHMAN YEARLY GRADE AVERAGE
"CROSS-VALIDATION GROUP"

<table>
<thead>
<tr>
<th>Variable</th>
<th>t ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic Computation</td>
<td>0.094</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>0.426</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>1.416</td>
</tr>
<tr>
<td>Spatial Relations 3-Dimensional</td>
<td>0.909</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>0.971</td>
</tr>
<tr>
<td>Language Usage</td>
<td>1.217</td>
</tr>
<tr>
<td>Freshman Yearly Grade Average</td>
<td>0.062</td>
</tr>
<tr>
<td>Complete battery of six predictors</td>
<td>3.299</td>
</tr>
</tbody>
</table>

Code

1.972 if significant at the five percent level

2.600 is significant at the one percent level

The standard error of the mean for the original sample is 0.221 and 0.326 for the validation sample.
The intercorrelations between the variables in the cross-validation group and the criterion, Total Score on the N.M.S.Q.T., along with the means and standard deviations are shown in Table XVI. The mean of the criterion is 1.574 higher than in the original sample; while the mean of each of the predictor variables is lower. The relative importance of these differences will be brought out in the next table.

TABLE XVI

NATIONAL MERIT TOTAL SCORE
MATRIX OF INTERCORRELATIONS - "CROSS-VALIDATION GROUP"

<table>
<thead>
<tr>
<th></th>
<th>Word Meaning</th>
<th>Intelligence Quotient</th>
<th>Language Usage</th>
<th>Paragraph Meaning</th>
<th>Arithmetic Computation</th>
<th>Arithmetic Reasoning</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N.M.S.Q.T.</td>
<td>.6315</td>
<td>.4953</td>
<td>.5566</td>
<td>.6350</td>
<td>.5459</td>
<td>.6271</td>
<td>19.612</td>
<td>3.872</td>
</tr>
<tr>
<td>Word Meaning</td>
<td>.5056</td>
<td>.5993</td>
<td>.6936</td>
<td>.3862</td>
<td>.5645</td>
<td></td>
<td>27.548</td>
<td>10.535</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>.4942</td>
<td>.4601</td>
<td>.4119</td>
<td>.5262</td>
<td></td>
<td></td>
<td>109.338</td>
<td>12.529</td>
</tr>
<tr>
<td>Language Usage</td>
<td>.6096</td>
<td>.3193</td>
<td>.4449</td>
<td></td>
<td></td>
<td></td>
<td>62.596</td>
<td>14.352</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>.3409</td>
<td>.5030</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.306</td>
<td>7.000</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td></td>
<td>.6288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.306</td>
<td>6.403</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.145</td>
<td>6.244</td>
</tr>
</tbody>
</table>
### TABLE XVII

**t** TEST FOR SIGNIFICANCE BETWEEN ORIGINAL GROUP  
AND N.M.S.Q.T. "CROSS-VALIDATION GROUP"

<table>
<thead>
<tr>
<th>Variable</th>
<th>t ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Meaning</td>
<td>1.660</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>0.428</td>
</tr>
<tr>
<td>Language Usage</td>
<td>1.217</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>0.972</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>0.094</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>1.416</td>
</tr>
<tr>
<td>N.M.S.Q.T. Total Score</td>
<td>0.575</td>
</tr>
<tr>
<td>Complete battery of six predictors</td>
<td>0.403</td>
</tr>
</tbody>
</table>

**Code**

1.972 is significant at the five percent level.

2.600 is significant at the one percent level.

The standard error of the mean for the original sample is 0.140; and for the validation sample, 0.252.
<table>
<thead>
<tr>
<th>Variable</th>
<th>F ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshman Yearly Grade Average</td>
<td>2.4767</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>1.1659</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>1.3631</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>1.2824</td>
</tr>
<tr>
<td>Spatial Relations (3-Dimensional)</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>#1.2896</td>
</tr>
<tr>
<td>Language Usage</td>
<td>1.0923</td>
</tr>
</tbody>
</table>

1.43 is significant at the five percent level
1.76 is significant at the one percent level

*Indicates the following criteria for significance levels.
1.42 is significant at the five percent level
1.62 is significant at the one percent level

The standard error of the standard deviation for the original sample is 0.498; and for the validation sample, it is 0.412.
### Table A.1

**F Test Between Original and "Cross-Validation" Samples for Total Score on National Merit Battery**

<table>
<thead>
<tr>
<th>Variable</th>
<th>F Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>N.M.S.Q.T. Total Score</td>
<td>1.134</td>
</tr>
<tr>
<td>Word Meaning</td>
<td>*1.542</td>
</tr>
<tr>
<td>Intelligence Quotient</td>
<td>1.363</td>
</tr>
<tr>
<td>Language Usage</td>
<td>1.092</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>*1.290</td>
</tr>
<tr>
<td>Arithmetic Computation</td>
<td>1.171</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>1.262</td>
</tr>
</tbody>
</table>

1.48 is significant at the five percent level

1.76 is significant at the one percent level

*Indicates the following criteria for significance levels

1.42 is significant at the five percent level

1.62 is significant at the one percent level

The standard error of the standard deviation for the original sample is 0.201; and for the validation sample, it is 0.246.
CHAPTER V

EVALUATION AND HYPOTHESES CONCERNING STATISTICAL RESULTS

The purpose of this study was to investigate the predictive value of the California Short Form Test of Mental Maturity, the California Test of Personality, the Occupational Interest Inventory and the Multiple Aptitude Tests in relation to subsequent academic achievement as revealed by the factor scores and composite score in the National Merit Scholarship Qualifying Test and the Yearly Grade Averages throughout the four years at Mendel Catholic High School. As a possible cross-validating battery with the California Test of Personality, Character Ratings by Teachers were also investigated as to their relative validity in predicting academic success. A perusal of that part of Chapter IV, entitled "Resume of Correlations Between Predictor Batteries and the Criteria Batteries" reveals the reasons why the California Short Form Test of Mental Maturity and the Multiple Aptitude Tests were selected as the best source of predictor variables; and the advantages in selecting the NMSQT Total Score and the Freshman Yearly Grade Average as the sole criterion variables.

Tables III and IV exhibit the lack of significant correlations between factors of the California Test of Personality and both criteria. This situation appears reasonable in the light of Hillebrand's and others' conclusions\(^1\). Hillebrand found that the possession of intelligence does not imply

---
\(^1\) Hillebrand, The Correlation of Personality Factors and Interest Patterns with the Intelligence Quotients of Some Gifted High School Pupils, 38.
personal or social adjustment as measured by this Test; nor does it relate significantly with the interest factors of the Occupational Interest Inventory. But, intelligence is significantly related to achievement. Hence it seems reasonable not to expect many, if any at all, significant correlations between personality factors or interest patterns and the achievement-criteria of this study; since the same instruments were used under similar circumstances. The same reasoning offered above appears sufficient to explain the paucity of significant correlations between the Occupational Interest Inventory and the selected criteria of this study. Strong's recent observation that interest tests do not correlate to any practical degree with measures of educational success, lends further credibility to the reasoning presented. Scientific Interest and Manipulative Interest patterns correlated above the one percent level of confidence with the criterion of Total Score on the NMSQT; and above the five percent level with the criterion of Freshman Yearly Grade Average. It may be concluded, therefore, that both of these interest patterns would very likely contribute to an improved estimate of an individual's chances for academic success. The same can be said of Level of Interest Score in relation to Total Score on the NMSQT; since this interest score correlated therein beyond the one percent chance expectancy.

Of the seventy-five correlations determined between factors of the

---

2Super, Appraising Vocational Fitness, 90.

California Test of Personality with Character Trait Ratings by Teachers, only two met even the five percent level of confidence. A comparison of the traits measured by the latter with the qualities presumably measured by the California Test of Personality led to the hypothesis that cross-validation was not feasible; nor was it justified. However, a glance at Table IV informs the reader that all of the Character Ratings by Teachers significantly correlated above the five percent level of probability with the criterion of Freshman Yearly Grade Average. It may thus be inferred that Character Ratings of individuals in this sample were rather accurate predictions of the Freshman Yearly Grade Average. Similarly, Table VI reveals that all of the Character Ratings by Teachers showed significant correlations above the one percent level of confidence with Total Score on the NMSQT; and as a consequence, should be regarded as significant predictors of the composite score in the National Merit Scholarship Qualifying Test. Nevertheless, for reasons indicated in Chapter IV, no attempt was made to employ any or all of the Character Ratings by Teachers as predictor variables in a multiple regression equation.

All nine tests of the Multiple Aptitude Tests battery were related significantly above the one percent chance expectancy with the criterion of Total Score in the National Merit Scholarship Qualifying Test. Of these nine tests, six were similarly correlated to the other criterion, Freshman Yearly Grade Average. It was decided, therefore, to concentrate exclusively on the investigation of the predictive value of I.Q. and these six tests in relation to both criterial measures. The relevant tests were: Word Meaning, Paragraph Meaning, Language Usage, Arithmetic Reasoning, Arithmetic Computation and Three-dimensional Spatial Relations.
An examination of Table III reveals that IQ as measured by the California Short Form Test of Mental Maturity, and six aptitudes measured by the MAT exhibit significant correlation above the one percent level of confidence with the criterion of Freshman Yearly Grade Average. If the assumption is made that the tests are valid, that is, that they measure the hypothetical qualities which they were designed to measure, then it may be inferred that IQ and these six aptitudes are highly significant factors in scholastic achievement as indicated by the Freshman Yearly Grade Average. An examination of Table V further reveals that all of the selected predictor variables showed significant correlation above the one percent level of probability with the total score in the NMSQT. If the assumption is made that the tests are valid, then it may be inferred that IQ and the above mentioned six aptitudes of the MAT play a highly significant part in the Total Score in the National Merit Scholarship Qualifying Test.

A publication of the California Test Bureau advises: "The correlation between intelligence test scores and academic achievement will, in all probability, be slightly lower, as one might expect, than the correlation between academic aptitude tests and academic achievement."4 The correlations obtained in the present study do not support this expectation.

The lowest correlations obtained between aptitudes and both criteria occurred when the aptitudes of Routine Clerical Facility and Applied Science and Mechanics were correlated with the criteria. This was no surprise, however,

---

4Questions and Answers on the California Test of Mental Maturity, by the Staff of the California Test Bureau, California, 1961, 4.
since students having these abilities and ambitions would hardly choose Mendal Catholic High School for their development.

The results shown in Tables III, IV, V, and VI may be summarized by saying that IQ, as measured by the California Short Form Test of Mental Maturity, and aptitudes, as measured by the Multiple Aptitude Tests have a far greater influence on academic achievement, as measured by Freshman Yearly Grade Average and by Total Score in the National Merit Scholarship Qualifying Test, than any of the other factors measured. In addition, they indicate that certain interests, as measured by the Occupational Interest Inventory, and all Character Ratings, as made by current teachers, show significant relationship to academic achievement, as measured by Freshman Yearly Grade Average and by Total Score in the National Merit Scholarship Qualifying Test.

The correlations displayed in Table VII manifest the significant relationship that appears to exist between the criterion measures. Freshman Yearly Grade Average correlated with Total Score in the National Merit Test beyond the one percent chance expectancy; and could likely contribute something to improved estimation of an individual's chances of success in the National Merit Scholarship Qualifying Test. There is nothing unusual about this, however, since both criteria are known measures of scholastic achievement at different levels of schooling.

With the exception of IQ, which maintained a second position in each numerical sequence of predictor variables, with respect to both criterial measures, the remaining predictor variables were practically in an inverse sequence of significance when correlated with each of the alternate criterion variables. The least significant of the predictor variables in each sequence
was eliminated, since it is recommended in practical situations that a multiple regression equation not exceed six predictor variables. This selection determined the final predictor variables exhibited in Tables VIII and XI respectively.

In setting up the predictive regression equations, one looks for independent variables which have low coefficients of correlation between each other and high coefficients of correlation with the criteria.

With this in mind, it became evident that the relatively high inter-correlations between the predictor variables, in comparison to the correlations between them and the criterion of Freshman Yearly Grade Average, (Table VIII), offered little hope of obtaining a highly significant multiple regression equation. However, since the author planned on using the DuBois method of multivariate correlation analysis in order to select a limited number of variables having highest multiple correlation with the criterion, knowing it would be of greater value than mere multiple correlation statements, he decided to obtain and utilize the prediction equation.

It should be pointed out that a comparison of the sample means and standard deviations of the selected independent variables with the means and standard deviations of the normative population reveals from insignificant to very significant departures. The mean IQ for the sample was 116.5, while the standard deviation was 10.2. The table on page 106 reveals the variations of

<table>
<thead>
<tr>
<th></th>
<th>Mean of Sample</th>
<th>Mean of Norm</th>
<th>S.D. of Sample</th>
<th>S.D. of Norm</th>
<th>S.E. Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic Computation</td>
<td>25.7</td>
<td>19.1</td>
<td>6.9</td>
<td>9.6</td>
<td>2.5</td>
</tr>
<tr>
<td>Arithmetic Reasoning</td>
<td>15.5</td>
<td>8.3</td>
<td>7.1</td>
<td>6.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Spatial Relations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three Dimen.</td>
<td>13.1</td>
<td>11.0</td>
<td>4.6</td>
<td>5.0</td>
<td>2.3</td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>35.2</td>
<td>25.9</td>
<td>6.2</td>
<td>7.7</td>
<td>3.4</td>
</tr>
<tr>
<td>Language Usage</td>
<td>75.7</td>
<td>56.4</td>
<td>15.0</td>
<td>14.3</td>
<td>5.3</td>
</tr>
<tr>
<td>Word Meaning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.6</td>
<td>10.8</td>
<td></td>
<td></td>
<td>3.2</td>
</tr>
</tbody>
</table>

the remaining five predictor variables from the established norms.

A comparison of the above statistics inclines to the conclusion that the sampling of students represented in this study was most probably not random. Actually, only approximately 25 percent of each freshman class wrote the various tests. The fact that a fee was expected from each applicant would appear to constitute an economic bias in the sample. Further, as a rule, only those students wrote the tests whose parents deemed the program of great value to their sons. Usually, this would infer greater educational background on the part of such parents.

Table IX exhibits the multiple R's obtained by employing the DuBois method.

---

"Guilford, *Fundamental Statistics in Psychology and Education*, 170."
In this routine, variables which duplicate the functions of variables already chosen tend to drop out. It will be noticed, therefore, that in adding the fifth variable (Arithmetic Reasoning) to the routine, multiple R increased only .0002. And when the sixth variable (Language Usage) was finally included, a further increase of but .0003 was gained. This is a good example of the way in which each additional variable yields diminishing returns in the way of improved predictions. It is certain therefore, that the trouble of measuring and using in the regression equation the two additional variables is not worthwhile. As a further note, it is questionable whether or not the addition of even the fourth variable (Paragraph Meaning) increased the forecasting efficiency of the prediction equation sufficiently to justify its inclusion. Hence, a prediction equation utilizing the three independent variables, Arithmetic Computation, IQ, and Three Dimensional Spatial Relations appears to be the optimum regression equation emerging from this phase of the investigation. Guilford suggests as probable that the efficiency of predictions based upon the average unsystematic interview is less than five percent. The index of forecasting efficiency for the three independent variable regression equation referred to above, is 7.1 percent.

The final group of independent variables, selected and ordered on the basis of significant contribution to multiple R with the criterion of Total Score in the National Merit Scholarship Qualifying Test, is displayed in Table X. Table XI exhibits the correlation of each predictor variable with the criterion,

---

8Ibid., 378.
as well as the intercorrelations among the predictor variables. The highly significant correlations of all predictor variables with the criterion, coupled with the many lower intercorrelations among the predictor variables, presented a favorable condition for the development of a worthwhile multiple regression equation.

The mean and standard deviation of the sample in the National Merit Scholarship Qualifying Test were 18.0 and 4.0 respectively. These statistics compared favorably with the norms (20.3 and 4.1 respectively) available on 335 male participants who subsequently entered a large Catholic University in Missouri. On the basis of this conformity, there is significant probability that the sample can be reasonably characterized as a college ability group. Further evidence of this is seen in the mean IQ (116) of the sample.

Table X reveals the results of the selection process embodied in the DuBois method of multivariate correlation. This procedure tends toward a selection of a limited number of variables having highest multiple correlation with the criterion. In some cases, this reveals a limited number of variables which have a multiple R almost as high as that of the total group of predictors, resulting in economy of application to practical situations. Rounding off the multiple R's in Table X to two places, and calculation of the index of forecasting efficiency for each, reveals a range of indices from 26.8 (Word Meaning alone) to 40.0 (utilizing all six predictor variables). Due to the process of rounding off to two decimal places, no increase in forecasting efficiency
occurred when the fifth variable was added, (Paragraph Meaning); but an increase from 38.8 to 40.0 occurred when the sixth independent variable was included in the routine. This inferred that sixty-four percent of the variance had been explained. By eliminating variables four, five, and six, the index of forecasting efficiency for the resulting three-variable regression equation was reduced to 36.3; which meant that fifty-nine and three-tenths percent of the variances had been accounted for. While there would be significant improvement in the way of prediction through use of the six variable predictor equation, Guilford would be inclined to question whether the trouble of measuring and using the regression equation the additional three variables is worthwhile 9.

Cross-validation is particularly important after we have used statistical data to choose the best tests to make up a battery for use with subsequent applicants or students. A cross-validation experiment on a new group will tell us how good our choice really is. Cross-validation data is valid only when the original tests are administered without change, without further revision or refinement, to an entirely new and independent criterion group.

Adhering to these requirements, the obtained multiple regression equations were applied to a new sample of sixty-two boys enrolled in September, 1958, at Mendel Catholic High School, one year later than the latest group included in the original sample. Table XII displays the new multiple $R$'s obtained between the final predictors and the criterion of Freshman Yearly Grade Average. The sequence of the predictor variables varied in the several intermediate batteries

---

Table XVIII and Table XIX reveal some significant F ratios; hence, it became necessary to investigate further, by using t tests. Table XV exhibits the t ratios obtained by comparing the results of the original sample with the results of the same battery of six predictors as given to sixty-two boys in the validation sample. The answer to the question, "Did the boys studied in the original sample differ significantly from the boys investigated in the validation sample?" is in the negative, since none of the predictor variables nor the criterion variable met even the five percent level of significance. However, the complete battery of six predictors, applied to the original sample, revealed a multiple R of .3734, as compared to .7061 when applied to the validation sample. When the t test was applied to these two correlations, the difference was found to be significant above the one percent level. The regression formula, as developed for the original test group should not be expected to apply to the validation sample.

When the t test comparison was made between the results obtained with the original sample and the identical battery of six predictors as administered to the validation group of sixty-two boys at Mendel, the criterion now being Total Score in the National Merit Scholarship Test, no significant difference was found. The complete battery, which revealed a predictive coefficient of $R = .8005$ for the original group, as compared to $R = .7731$ of the validation group, exhibited no significant difference in the t test comparison of the two groups. Variations which did appear could have arisen from sampling factors. The regression formula with the weights assigned to the various variables withstood the test of cross-validation; and may, therefore, be used again with confidence in similar situations.
CHAPTER VI

SUMMARY OF INVESTIGATION

This study investigated the predictive value of psychological measurements obtained from 210 ninth grade pupils at Mendel Catholic High School in relation to their subsequent academic successes, as measured by their four successive yearly grade averages and their scores on the National Merit Scholarship Qualifying Test during their eleventh grade. The psychological measures of mental maturity, personality, occupational interests and multiple aptitudes were obtained by pencil and paper tests published by the California Test Bureau. As a cross-validating technique with the pencil-and-paper test of personality or as other correlative factors, Character Rating scores by teachers were likewise investigated as to their relative validity in predicting academic success.

All possible Pearson product-moment coefficients of correlation were ascertained by an IBM-650 machine. A thorough study of the "more than 1200 correlations" led to the selection of the significant predictor variables and related criterion variables. One of the criterion variables, Freshman Yearly Grade Average, exhibited highest relationship to mental maturity and the following factors of the Multiple Aptitude battery: Arithmetic Computation, Arithmetic Reasoning, Spatial Relations (3-Dimensional), Paragraph Meaning and Language Usage. The other criterion variable, Composite Score on the National Merit Scholarship Qualifying Test, was most significantly related to mental maturity and the following factors measured by the Multiple Aptitude battery:
Word Meaning, Language Usage, Paragraph Meaning, Arithmetic Computation and Arithmetic Reasoning. The fact that all final predictor-factors were from two California Test Bureau instruments, was seen as an advantage - both in simplicity and economy, should the prediction equations prove significant.

Only one of the factors of the pencil-and-paper test of personality correlated significantly with any of the original criteria. This was School Relations; which correlated at the five percent level of significance with the English Usage score of the N.M.S.Q.T. The Scientific factor of the interest battery was the only score therein to correlate significantly at the one percent level with all six N.M.S.Q.T. scores. The highest of these correlations was .3305, with the Natural Science Reading factor. This Scientific factor correlated with Freshman Yearly Grade Average just below the one percent level of significance; but failed to correlate significantly even at the five percent level with the remaining yearly grade averages. Two Character Ratings by teachers (Initiative and Leadership) correlated significantly with one of the factors of the paper and pencil test of personality (Social Standing); otherwise, the instruments were hardly related. However, all of the Character Ratings by teachers correlated at least above the five percent level of significance with all of the original criterion variables.

Employing the DuBois method of multivariate correlation analysis, a multiple regression equation was developed for optimum prediction of each of the selected criterion variables. It was determined that a prediction equation utilizing the three independent variables, Arithmetic Computation, mental maturity, and Spatial Relations (3-Dimensional) was practically as good as the complete six variable equation, for the prediction of Freshman Yearly Grade
As for the prediction of Total Score on the N.M.S.Q.T., the six variable predictor equation would yield optimum forecasting; however, a three variable predictor equation utilizing the independent variables; Word Meaning, mental maturity, and Language Usage, would be more useful and yet reliable.

To test the validity of the obtained regression formulae with their assigned weights, a second sample consisting of 62 boys from the same school was tested one year later.

The multiple regression equation predictive of Freshman Yearly Grade Average, applied to the original sample, revealed a multiple $R$ of .3734, as compared to .7061 when applied to the validation sample. When the $t$ test was applied to these two correlations, the difference was found to be significant above the one percent level. The regression formula, as developed for the original test group should not be expected to apply to the validation sample.

The multiple regression equation predictive of Total Score on the National Merit battery, applied to the original sample, revealed a predictive coefficient of $R = .8005$, as compared to $R = .7781$ when applied to the validation sample. When the $t$ test was applied to these two correlations, no significant difference appeared from the comparison of the two groups. This regression formula, with the weights assigned to the various variables withstood the test of cross-validation; and may, therefore, be used again with confidence in similar situations. While the predictor value of the battery would not be sufficiently high to justify selection and elimination of pupils on the basis of scores obtained, the information provided by the scores would be of diagnostic value to the teacher, as it would indicate which pupils might experience difficulties so that special attention might be given to these pupils. It would also tend to
indicate those individuals who could most likely score highest in the National Merit Scholarship Qualifying Test.
CHAPTER VII

RECOMMENDATIONS AND SUGGESTIONS FOR ADDITIONAL RESEARCH

In planning this dissertation, the hope was to predict academic success from measures of performance available in student records, by setting up a matrix of intercorrelations between the variables, and from the respective coefficients contained therein, to select batteries of tests or subtests which should predict academic success as measured by the selected criteria. There is a choice in many testing situations as to whether one general test should be used or whether a number of different tests should be administered in a battery. The potential advantage is always with multiple measurements; more information cannot lower prediction. Moreover, there are some circumstances in which it is especially important to measure the individual's differential ability. One such circumstance is vocational guidance. The individual needs advice; and, regardless of how high or low his over-all ability, the problem is to find the occupations that are best suited to his pattern of abilities. Even the person who makes a very low score on a general intelligence test might have certain aptitudes, such as in arithmetic computation or language usage, that would suit him for some jobs.

There is little doubt that well-designed test batteries will eventually come into use in many testing programs. They will pay for themselves by improving the guidance, selection, and classification of individuals. The know-how is already available to construct better test batteries for many uses; and
it will come to fruition when administrative officials in schools, the Armed Services, industry, and other institutions supply the funds and resources.\(^1\)

The newer concept of a program of individualized guidance in the secondary school as contrasted with the older practice of mass education, is concentrating the attention of administrators, teachers, and counselors at all levels of the school on the need for a great variety of information about each individual pupil. In their efforts to obtain records of ability, personality, interests and achievement that have a high degree of objectivity and reliability an increasing number of schools are introducing programs of measurement. The extent of these programs will vary; but usually, guidance experts agree that measures of intelligence, aptitudes, personality, interests and achievement constitute a minimum program.\(^2\). Guidance authorities also advise that tests and inventories will constitute the major part of the material budget for the guidance services.\(^3\)

Donald Super emphasizes in the preface to his book, Appraising Vocational Fitness, that one of the principal weaknesses in the measurement movement has been the excessive individualism of the research which has been carried on. He continues: "Individualism has been good in that it has encouraged branching out in new directions and trying out new possibilities, but it has been bad in that it has resulted in the scattering of efforts and in the frequent dropping of a

\(^1\) Nunnally, Tests and Measurements, New York, 1959, 140.


\(^3\) Ibid., 2:4.
good idea after it has been barely tried. For every research project comparable
to Strong's persistent study and refinement of his Vocational Interest Blank
throughout the past twenty years, there are several like Zve's Scientific
Aptitude Test and Bernreuter's Personality Inventory, whose initial promise
have never been adequately explored. This is partly because the test authors,
often for excellent reasons, did not follow up their initial work, partly
because the research carried on by other people with these instruments has
generally been unco-ordinated and incidental. Super believes, and the present
author strongly recommends this too, that for test development work to be fully
effective, two innovations are needed. One of these is the periodic and
systematic review of work with specific tests or types of tests. This should
be more integrated, critical and creative than the periodic reviews presently
popular. This could best be done through a co-ordinated committee,
representative of the major groups vitally concerned with the problem. The
second major need in the development of vocational testing is an extension of
the function of this committee from systematic review and suggestion to system-
atic planning and execution of research. Such a committee should take the
initiative in encouraging research along needed lines, partly by publications
and talks at professional meetings, and partly by a program of grants-in-aid
for suitable research.

A current cooperative study undertaken by the American Association of
School Administrators, the Council of Chief State School Officers and the
National Association of Secondary School Principals, warned that large-scale

4 Super, Appraising Vocational Fitness, xvi.
external testing programs have gotten out of hand. The booklet advised high schools "to refuse to participate in nationally sponsored tests unless those tests can be demonstrated to have value commensurate with the effort, money, time and emotional strain involved". The report criticized the importance attached to test results, noting that they are not infallible and cannot measure many important qualities. Impliedly, the report is aimed principally at the National Merit Scholarship Qualifying Test, the College Entrance Boards Exam, the Regents Scholarship Exam, and such like.

The recent study by the University of the State of New York, reviewed in Chapter II, appears to indicate a prior awareness on the part of some external testing agencies, of an aggravation which was building up among secondary school officials.

While standardized subject-area tests are becoming more popular in our secondary schools, and to that extent, administrators and guidance personnel are better informed about students' strengths and weaknesses; generally speaking, the sole criterion of academic achievement in the minds of educators is still students' school marks. Aware of this situation, Guilford regards it as a drawback to studies of validity of tests. "The use of school marks," he says, "as criteria of success in training, is in itself a questionable procedure. School marks being derived as they generally are by the indiscriminate pooling of marks from different subjects and from different instructors and treating them as if

---


they were of the same kind of coin, renders them of questionable reliability and
validity and contaminated with irrelevant factors." This situation alone
militates against high validity coefficients for predictive indices.

It is the author's conclusion and recommendation that limited external test-
ing, such as school-wide participation in the National Merit Scholarship
Program, be encouraged and expanded; so that (among other benefits) there be at
least one highly reliable and valid criterion of academic achievement for a
school or a school system.

Brought into focus indirectly by the recent two-year study of external
testing programs in our secondary schools, and highlighted by the conclusions
arrived at in that study, is the great need for a co-ordinated investigation of
our internal testing programs. Since various test publishers supply their own
specific instruments for the variety of measurements required in a good
guidance program, it appears reasonable and urgent that specific batteries
produced by each publisher be investigated as to validities, reliability, and
practicality.

This dissertation has attempted to investigate the predictive validity of a
recommended battery of psychological measuring instruments published by a single
Test Bureau. The deficiencies and strengths have been revealed; and a "best"
use of this battery has been suggested in the presentation of two multiple
regression equations. Similar research under like situations would produce


*Testing, Testing, Testing*. 
comparable data, which in turn would provide a pool of information from which better judgments and selection could be made.

Systematic research into the various validities and reliability of batteries, eclectically formed by selecting particular tests from different publishers could be conducted simultaneously or subsequently. Results from such investigations would provide a wide range of choice for administrative and guidance personnel; and certainly would tend to point the way toward a reduction of any shortcomings revealed by the present investigation.
BIBLIOGRAPHY

BOOKS


Ross, Clay C., *The Relation between Grade-School Record and Achievement*, Teachers College, Columbia University, New York, 1925.


ARTICLES

Alteneder, L. E., "The Value of Intelligence, Personality, and Vocational Interest Tests in a Guidance Program", Journal of Educational Psychology, XXXI, 1940.


Berdie, R. F., Factors Related to Vocational Interests Psychological Bulletin, XLI, 1944.


Carter, Harold D., "Improving the Prediction of School Achievement", Educational Administration and Supervision, XLV, September, 1959, 225-260.


Ellis, Albert, "The Validity of Personality Questionnaires", Psychological Bulletin, XLIII, September, 1946.


McNemar, V. Q., "Critical Examination of the University of Iowa, Studies of Environmental Influences Upon the I.Q.", Psychological Bulletin, XLIII, 1946, 282-374.


Proctor, W. M., "Intelligence and Length of Schooling in Relation to Occupational Levels", School and Society, 1935, 783-786.


Terman, L. M., "Intelligence and Its Measurement", Journal of Educational Psychology, March, 1921, 128-147.


Tiers, Ernest W., Clark W. Willis, and Louis P. Thorpe, "The California Test of Personality", Journal of Educational Research, XXXV, October, 19


UNPUBLISHED MATERIALS

Ardizzone, Frank, Differences in Reading, Intelligence, and Personality Among High, Average, and Slow Groups of the Sixth and the Seventh Grades, Unpublished Master's Thesis, De Paul University, Chicago, 1959.

Charles, Ramon L., A Study of the Relationship Between Level of Interest and Academic-Achievement, Unpublished Study, Division of Guidance and Personnel Services, Kansas State Department of Public Instruction, February, 1957.


MISCELLANEOUS MATERIALS


Clark, W. W., E. W. Tiers, and L. P. Thorpe, California Test of Personality, California Test Bureau, California, 1953.


Questions and Answers on the California Test of Mental Maturity, California Test Bureau, California, 1961, 4.


APPROVAL SHEET

The dissertation submitted by Reverend Daniel Joseph Hartigan, O.S.A. has been read and approved by five members of the Department of Education.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the Degree of Doctor of Education.

6/9/62

Date

Arthur P. O'Mara
Signature of Adviser