AN ANALYSIS OF THE EDUCATIONAL TESTING

PROGRAM OF THE SCHOOL OF POLICE

ADMINISTRATION AND PUBLIC SAFETY

Thesis for the Degree of M. S.

MICHIGAN STATE UNIVERSITY

Donald Jameson Pope

1966

AN ANALYSIS OF THE EDUCATIONAL TESTING PROGRAM

OF THE SCHOOL OF POLICE ADMINISTRATION

AND PUBLIC SAFETY

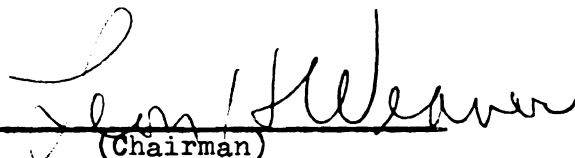By

Donald Jameson Pope

AN ABSTRACT OF A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
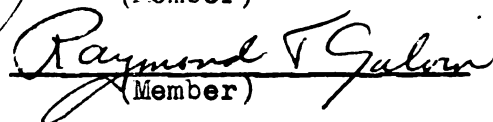for the degree of

MASTER OF SCIENCE

Department of Police Administration and Public Safety

May 1966

APPROVED _____
(Chairman)

_____
(Member)

_____
(Member)

ABSTRACT

AN ANALYSIS OF THE EDUCATIONAL TESTING PROGRAM
OF THE SCHOOL OF POLICE ADMINISTRATION
AND PUBLIC SAFETY

by Donald Jameson Pope

For a total of eleven years the School of Police Administration and Public Safety at Michigan State University administered a battery of tests to all freshmen and transfer students entering the school. This test battery was composed of tests designed to measure academic ability, emotional maturity and stability, and vocational interest. The purpose of this study has been to evaluate the effectiveness of this testing program. This evaluation was approached from two angles. The first of these was the over-all effectiveness of the program and the use made of its results. The second approach concerned the actual data gathered, and how this data might be used to predict the academic success of students enrolling in the School of Police Administration and Public Safety. As the reader will see in succeeding chapters, the study presents some conclusions, but on other points the study has been inconclusive or contradictory.

An evaluation of the whole testing program was not possible because of the somewhat deficient planning and organization behind the program. Those conclusions that can be drawn indicate fairly

clearly that the program was only a qualified success. There was little use made of the data as it was gathered, and it was not organized in a manner to facilitate any study of its over-all meaning or contribution.

Since the total program could not be studied, conclusions as to the predictive ability of its results must be based on a relatively small portion of the data gathered. These conclusions are somewhat tentative and do little more than show that one of the tests studied, the College Qualification Test, bears a small relationship to academic success. This relationship is consistent, but not strong enough to justify its use as a device for screening applicants for admission.

However, the shortcomings of this first testing program do not mean that further efforts should not be considered. If anything, it points out our lack of knowledge in areas where we should know much more than we presently do know. There is a very real need for more exhaustive studies in areas relative to success in higher education. The School of Police Administration and Public Safety can and should assume leadership in studying those areas related to its own unique contribution to the community.

AN ANALYSIS OF THE EDUCATIONAL TESTING PROGRAM

OF THE SCHOOL OF POLICE ADMINISTRATION

AND PUBLIC SAFETY

By

Donald Jameson Pope

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Police Administration and Public Safety

May 1966

ACKNOWLEDGMENTS

If I were to give full and detailed credit to all who contributed to the development of this thesis, the acknowledgments would be longer than the text. This campus seems to have more than its quota of helpful people who courteously answer what very probably seem foolish questions to them.

The staff at the computer center was most helpful, especially, Mrs. Norma Landis. I am deeply indebted to Doctor Karslake of the Psychology Department for his counseling on matters methodological. This study could not have been completed without his invaluable assistance. My special appreciation to Doctor Leon Weaver who has assisted me throughout this study, listened to my tales of woe, and encouraged me to further efforts; and finally, deepest gratitude to my mother, who is even now typing the last pages of text as I write these acknowledgements.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

For a period of eleven years, ending with the Fall term of
1963, the School of Police Administration and Public Safety at Michi-
gan State University administered a mental and psychological testing
program to all entering students. The purpose of this study is to
evaluate the effectiveness and significance of that testing program.
The most important part of this study centers around the search for
components of the test battery that will predict academic success.
There will also be occasion to evaluate the testing program in terms
of how well it was organized and the uses to which its results were
put.

No formal hypothesis is tested in this study. Rather, this
study represents an attempt to organize an already existing body of
data and to search for possible relationships between the data and
academic success in college.

CONTRIBUTIONS OF THE STUDY

The importance of this study lies in its contributions to two
major areas. The first of these is in providing the School of
Police Administration and Public Safety with an evaluation of their
now discontinued testing program. It is hoped that, on the basis
of this evaluation, the school will be encouraged to develop better
organized and more effective testing programs in the future. The

second contribution, hopefully, will be to our general knowledge and understanding of the factors relevant to academic success.

At the present time, the search for factors relating to academic success and attempting to predict academic success receive high priorities from psychologists and educators. Their concern for this area is not without justification, as we know all too little about why some students succeed and others fail in school. We know even less about why some seemingly good students fail while many poorer students succeed. To be sure, we can identify many of the relevant factors operating here, but we are not yet able to determine how and to what extent these factors operate.

LIMITATIONS OF THE STUDY

Any conclusions drawn on the basis of this study will necessarily be limited by the design of the testing program and by the students participating in the program. The design of the testing program imposes several limitations on the study. The first of these is that the complete program could not be studied. One reason for this is that the composition of the test battery was changed in the middle of the eleven-year span of the program, thus making difficult or impossible an evaluation of test results covering the entire eleven year period. The other reason is that many of the students' test records were imcomplete, and therefore un-usable.

Another possible limitation lies in the fact that the students taking the tests are not representative of the University population as a whole. In fact, the Office of Evaluation Services reports on comparative standings on the M. S. U. Orientation Tests

show that students entering the School of Police Administration and Public Safety have average scores much below the University average. The only academic areas scoring consistently lower than Police Administration were Agriculture and Education.

As a result of these limitations, any efforts to extend the conclusions of this study to other testing programs or to other student groups should proceed with caution. There will be some generalisations made in the concluding sections of this study, but they will be somewhat tentative and in general supported by the findings of other researchers.

METHODOLOGY

The general plan of the study was first to divide the records from the testing program into three equal groups, with each group containing an equal number of students from each term. Then, through correlating test scores with academic performance, as represented by all-University and Police Administration Grade Point Averages, the first two groups of students were used to identify test variables which seemed to correlate well with performance. These remaining test variables were then used to predict academic performance for the third group of students. A comparison of their predicted performance with their actual performance would indicate, hopefully, the extent to which the test items related to and were indicative of, a student's level of academic success.

The reader will note in succeeding chapters that the neat study plan outlined above did not always prove workable. However, with modifications in design, the study was still able to reach some meaningful conclusions.

ORGANIZATION OF THE REMAINDER OF THIS THESIS

The remainder of this thesis is organized into five major parts, each part represented by a separate chapter. The next chapter (Chapter II) deals with the nature of the Police Administration testing program. Chapter III is based on a review of the literature concerning attempts to predict academic success on the college level. Chapter IV provides a detailed explanation of the study design. It also covers most of those factors that necessitated changes in the design of the study.

Chapters V and VI deal with the actual results of the study, the former representing the results and the latter analyzing them. In Chapter VI conclusions are stated insofar as generalizations from the data can be justified. The final chapter, in addition to evaluating the testing program, also contains some suggestions concerning content and organization of future studies that the School of Police Administration and Public Safety might attempt.

CHAPTER TWO

THE TESTING PROGRAM

Beginning with the fall term of 1952 and continuing through the fall term of 1963, all new freshmen and transfer students entering the School of Police Administration and Public Safety were given a battery of psychological tests and inventories. The tests were generally given in the introductory course in Police Administration and then scored and interpreted by personnel in the Michigan State University Counseling Center. The final results for each student, reduced to summary form, were then transmitted back to personnel in the School of Police Administration concerned with academic advising. These test scores were intended to aid in the advising process by providing indications as to where each student's strong and weak points lay. (The Appendix shows the report form employed).

In the course of these eleven years, about 2,200 students were enrolled in the introductory Police Administration course. For reasons that have not been documented, almost half of these students had incomplete or totally missing test scores. There are about 1,200 partial or complete sets of test scores available, but as tests taken over some periods were unusable, no exact count of the total was made.

During this eleven year period there were a number of changes made in both the composition of the test battery and in the form of

reporting. It is for this reason that the sample used in this study is so much smaller than the number of students taking the tests. Chapter four explains in detail why the nature of the study necessitated selecting such a relatively small group.

The tests employed in this program were concerned with assessing emotional stability, maturity, vocational preference, level of academic potential or development, and certain aspects of the individual's personality. By covering such a wide scope, the tests employed in this program represent a fair cross section of the testing instruments available. Chapter four lists all the tests and inventories employed throughout the duration of the program. A brief review of the nature of these tests actually covered in this study will provide the reader with some idea of its scope and composition.

THE M. S. U. ORIENTATION TESTS

The M. S. U. Orientation tests comprise a battery of tests presented by the University to all incoming Freshmen and transfer students. In both form and function they are similar to the entrance tests employed by most colleges and universities. They are all attempts to measure the general academic development of the entering college student. During the test period the Univesity employed two different batteries of entrance tests, but this study was limited to just one of them.

The "College Qualification Tests"[1] consist of a three-part

---

[1] College Qualification Tests: Combined Booklet Edition, the Psychologican Corporation, New York, 1959.

multiple choice examination covering verbal ability, mathematical ability and general information. In measuring verbal ability, the student is asked to pick a word that means the same (or opposite) of a given word. Selection is made from a list of four alternatives and the items vary in difficulty. In the mathematics section the student is given a problem and asked to select the solution from a list of alternatives. The general information section simply asks questions drawn from diverse areas and asks the student to choose the correct answers. The sum of these three tests, the CQT total score, is also reported.

In addition to the CQT the entrance battery also includes the "MSU Reading Test"[2] and the "English Placement Test"[3]. The reading test is designed to measure a student's ability to understand and learn from what he reads. It consists of eight reading selections representative of several academic areas here at Michigan State University. After each reading passage there is a series of questions on the meaning and contents of the passage. In addition to measuring the student's ability to remember what he reads, the questions also provide some measure of the factors involved in critical thought.

The English Placement Test consists of thirty-five multiple choice questions representing various aspects of English usage: spelling, capitalization, grammar, punctuation, sentence structure

---

[2]MSU Reading Test, Office of Evaluation Services, Michigan State University, East Lansing, Michigan, Form 1962.

[3]English Placement Test, Benjamin B. Hickok, Department of American Thought and Language, Michigan State University, East Lansing, 1962.

and Organization. This test is used primarily to identify students who need remedial assistance in English or who qualify for honors sections.

OTIS TEST OF MENTAL ABILITY[4]

The Otis Test of Mental Ability is designed to measure reasoning ability, not accumulated learning or knowledge. While it is not possible to measure mental ability directly, one can measure the effect mental ability has on a student's ability to acquire certain knowledge and mental skill. One way of doing this is to ask questions that draw more on a person's innate ability and reasoning power, and less on prior learning. It is admittedly impossible to separate completely inborn skills from those acquired through learning processes, but a careful selection of test items can reduce this problem to a point where meaningful statements can be made about a person's mental ability.

This mental ability is scored and interpreted in terms of an intelligence quotient, or IQ. The IQ score is a function of the ratio of mental age divided by chronological age, multiplied by 100 for ease of numerical representation. Hence a person with a mental age of eighteen and a chronological age of eighteen would have an IQ of 100 which is regarded as "normal". Similarly, a person with a mental age higher than his chronological age would have an above average IQ (over 100).

This particular IQ test consists of eighty multiple choice

---

[4]Otis, Arthur S., _Otis Quick-Scoring Mental Ability Tests: New Edition: Gamma Test: Form EM_, (The World Book Company; New York), 1954.

items dealing with vocabulary, arithmetic reasoning, spatial relations, number series, logical relationships and other areas.

THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY (MMPI)[5]

The MMPI was designed to provide a broad analysis of an individual's personality structure in a single test instrument. As the test was developed it was administered to persons in mental institutions and to others with known mental or emotional problems. The various scales on the test were formulated on the basis of patterns of answers that correlated well with known and diagnosable mental problems. The instrument consists of 550 statements, all in the first person, covering a range from the physical condition of the person being tested to his moral and social attitudes. Each question is to be answered "yes", "no" or "can't say". After the subject has answered all the questions about himself, his test is scored so as to routinely yield scores on three or four validity scales and ten or more clinical scales. The number of scales scored will depend on the reason for giving the test. For the School of Police Administration and Public Safety scoring was based on three validity scales and nine clinical scales.

The L Score is based on a small group of items that will yield a score tending to indicate any naive attempt by the subject to put himself in a good light. High scores are obtained by persons who try (often unconsciously) to answer all questions in ways that will fit their conception of the moral code regardless of their own knowledge

---

[5]Hathaway, Starke R., and McKinley, Charnley J., Booklet for the Minnesota Multiphasic Personality Inventory. (The Psychological Corporation, New York), 1953.

of themselves. The questions usually dealt with undesirable conduct or moral "failures" that may seem bad, but which are common in virtually all people to one degree or another. The assumption is that a normal person, answering truthfully, will admit yes to most of these items. Attempts at concealment here are called defensiveness, or even "faking good" in extreme cases.

The F Score is just the opposite of the "L" score. Persons scoring high on this scale are often attempting to show themselves in a bad light. This could be because the subject is overly candid or because he is generally maladjusted.

The K score is similar to the "L" score, but much more subtle. It is actually mixed in with five of the clinical scales which when high indicate defensiveness and a lack of candor. Low scores, like a high "F" score, indicate frankness and self criticism.

The Hs scale is based on a comparison of normal persons with patients having numerous physical complaints, many of them psychological in origin. These complaints may be organically based, but are more often classed as psychosomatic.

The D scale is based on the scores of emotionally depressed persons. A high score here is often associated with persons who are unsure of themselves and of the future. They frequently tend to be sad or moody. Higher scores here frequently occur when a subject is in trouble.

The Hy scale is similar to the Hs scale, but gives evidence of more elaborate and serious symptoms. High scores here are indicative of persons who are more likely to avoid or run from crises and when pushed too far are likely to develop physical or psychological

illness as the means of escape from a difficult situation.

The Pd scale deals with what is termed psychopathic deviancy. This is characterized by a person's failure to be controlled by the mores of society and an apparent inability to feel guilt or remorse or to respond to punishment. These people frequently commit asocial acts with no apparent motive.

The Mf scale is a measure of masculine or feminine interests. High scores here tend to indicate that a person's interests are more in line with those of the opposite sex than those of his or her own sex. High scores here can also be indicative of homosexual tendencies, especially for men.

The Pa scale is a measure of interpersonal sensitivity, and in extreme cases, paranoia. High scores here may indicate a person who feels he is being persecuted or mistreated by other people.

The Pt scale is related to compulsions and obsessions. High scores here may indicate neurotic compulsions in selected areas, excessive worrying or a tendency to be overly introspective and self critical. Many asocial acts are not a result of psychopathic deviancy, but of persons driven by compulsions or obsessions.

The Sc scale is concerned with how much the subject reacts and and thinks like others around him. Extreme scores may be indicative of schizophrenia, but more generally the scale measures a person's tendency to distort what he sees and hears and to react in unusual ways.

The Ma scale attempts to measure enthusiasm and energy. High scorers become readily interested in new projects and approach problems with animation. In extreme cases this may result in antisocial

acts or irrational manic behavior.

It should be emphasized that the MMPI is not a complete diagnostic tool, but rather, that it will tend to indicate where a person may have emotional problems. Further testing and interviewing is usually necessary before a diagnosis can be made.

This testing program was comprehensive in scope. As was already mentioned, it dealt not only with academic ability, but with several aspects of the student's emotional and mental makeup. The pertinent question now is the degree to which any or all of these tests might correlate with the subject student's actual college performance. Since any one of the above tests or scales could theoretically correlate well with academic performance, it is necessary to include instruments that do not, on the surface, seem to be related to academic ability. The MMPI is a good case in point.

CHAPTER THREE

A REVIEW OF CURRENT ATTEMPTS TO PREDICT

ACADEMIC PERFORMANCE

The professional literature dealing with attempts to predict academic success bears a strong resemblance to a literary Irish stew. It consists of a heterogeneous mixture of predictive variables, success criteria and methods of operation with little or no over-all pattern or organization discernible. There seems to be a tendency for many of the writers in this area simply to take whatever materials are at hand and make a study of them. More often than not these studies are attempts to salvage something from educational testing programs that have been operating for years, but for which no data on their predictive validity has ever been gathered.

Virtually all colleges and universities administer some type of admissions testing program for entering students. These test batteries are generally designed to measure academic ability in such areas as Mathematics, Reading, Grammar, Spelling and general verbal usage. In some instances highly specific tests may be used for students interested in technical areas such as Engineering and the sciences. The assumption behind these tests is that in some way they separate potentially good from potentially poor students. This assumption has in some instances been proven totally incorrect and

at best only partially correct.

Attempts to predict academic success are being made on all educational levels, from pre-school to graduate school. The results obtained seem to be roughly similar on all levels, so a review of a narrow segment of this literature should provide a representative picture of the general field. Since this particular study deals with prediction on the college level, the literature reviewed dealt with studies on this level.

The literature research was conducted during the summer of 1965 in the Ohio State University Libraries. The basis of this research consisted of a review of articles on academic prediction listed in the Psychological Abstracts from late 1960 through early 1965. Of the total number of articles reviewed, forty five were relevant to this study. The articles and studies cited in this chapter have been selected on the basis of how representative they are of certain areas of academic prediction.

It should be noted that studies in this field can approach their material from several different directions, and each direction places a slightly different emphasis on the materials involved. Perhaps the most common form of study simply attempts to find statistically significant correlations between test scores or other variables and selected criteria of academic success.

A good example of this type of approach is a 1964 study by Michael, Baker and Jones.[6] Using a sample of Freshmen men and women

---

[6] William B. Michael, David Baker, and Robert A. Jones, "A Note Concerning the Predictive Validities of Selected Cognitives and Non-Cognitive Measures for Freshman Students in a Liberal Arts College," Educational and Psychological Measurement, Vol. XXIV, No. 2, (1964) 373-375.

in a liberal arts college they administered a battery of tests and worked out correlations with grade point averages (GPA) for their Freshman year. The tests employed in this study were:

1. California Study Methods Survey (CSMS) with four scales – Attitudes toward school, mechanics of study, planning and system, and verification;

2. Type Indicator (TI) with four scales – Introversion, Intuition, Thinking, and Perception;

3. High school GPA's in Algebra and Chemistry;

4. University of Southern California English Classification Test (ECT);

5. Scholastic Aptitude Test (SAT) with three scales – Verbal, Quantitative, and Total.

Of all these scales, CSMS Attitudes toward school yielded the highest correlations for men (r = .42) while the SAT Total was the highest correlator for women (r = .48). An examination of Table I reveals three interesting characteristics. One is that with the exception of the Introversion scale of the Type Indicator test, correlations for men and women are either both significant or not significant. Second, with the exception of the English Classification Test, correlations for men and women are within .15 points of each other and often much closer. Third, the scores for women seem to correlate better with academic performance than do the scores for men.

TABLE I[a]

CORRELATION OF SELECTED VARIABLES WITH THE GPA's
OF FRESHMEN IN A LIBERAL ARTS COLLEGE

Test given	Men		Women	
	Sample size	r	Sample size	r
SAT Verbal	210	.35[b]	177	.44[c]
Quantitative	210	.24	177	.39
Total	210	.35	177	.48
CSMS Attitudes toward school	155	.42	141	.39
Mechanics of study	155	.36	141	.22
Planning and system	155	.33	141	.18
Verification	155	.09	141	.10
Type Indicator Introversion	140	.16	131	.24
Intuition	140	.06	131	.07
Thinking	140	.02	131	.02
Perception	140	.08	131	.00
High School GPA's in Algebra	161	.06	69	.06
Chemistry	173	.03	113	.16
English Classification Test	210	.23	177	.47

[a] After Michael, Baker and Jones (1964).

[b] r's greater than .22 and .17 are significant to the
01% and 05% levels respectively (when n = 140).

[c] r's greater than .302 and .232 are significant to
the 01% and 05% levels (when n = 70).

Another prediction study was made by Arvo E. Joala[7] in 1963

[7] Arvo E. Juola, "The Development of an Academic Predictor
Scale Based on Students' Attitudes Toward Education," The Personnel
Guidance Journal, Vol. XLII (December, 1963) 381-386.

based on both ability and attitude measuring instruments. He was

exploring the possibility of predicting academic success on the

basis of attitudes toward education rather than using just ability

measures. In addition to the College Qualification Test (CQT),

Juola employed the Academic Attitude Preference Inventory (AAPI)

wherein the student rates activities or situations likely to be

encountered in college on the basis of how much he likes or dis-

likes them.

TABLE 2[a]

COMPARISON OF CORRELATIONS BASED ON ABILITY AND ATTITUDES
WITH FIRST TERM AND FIRST FULL YEAR COLLEGE SUCCESS

Tests given		Test Correlations with GPA's			
		First term GPA		Freshman year GPA	
		men[b]	women	men	women
AAPI		.52[c]	.52	.48	.48
CQT	Verbal	.50	.47	.48	.43
	Information	.53	.58	.49	.56
	Numerical	.49	.48	.46	.52
	Total	.63	.65	.59	.64

[a] After Juola, Arvo E. (1963).

[b] Based on samples of 212 men and 210 women.

[c] All r's significant to the .01% level.

Examination of Table 2 seems to indicate that the AAPI may

have a certain consistency between men and women that the CQT lacks,

at least in Juola's study. However, since very few instruments

yield identical correlations for both sexes, this may well have been

a statistical "accident". In degree of correlation, the AAPI seems

to be inferior to both the Information and Total scores of the CQT. Two other points of interest in the preceding table are that correlations appear to be consistently higher for first term GPA's than for the full year, and that correlations with womens' grades again seem higher than for men. (Could it be that women really are more predictable?)

An especially relevant study by King, Norrell and Erlandson[8] dealt with 375 male Police Administration students here at Michigan State University and used several of the instruments that are employed in this study. Table 3 summarizes the major results of this study between test scores and first term grades. In general, correlations seem uniformly low in this study, with only the total score for the MSU Reading Test and the Language scale of the California Test of Mental Maturity even rising as high as .50. Twenty eight of the thirty seven correlations are below .25. It is interesting to compare the results of this study to what Michael, Baker and Jones discovered about the Type Indicator. Both studies seem to support the contention that tests designed to measure aspects of the personality (Type Indicator, Minnesota Multiphasic Personality Inventory and the Thurstone Temperament Schedule) seem to bear little relation to academic performance. It is also interesting to note that vocational preference, as measured by the Kuder test, also seems to have little bearing on the problem.

---

[8]Paul King, Gwen Norrell and F. L. Erlandson, "The Prediction of Academic Success in A Police Administration Curriculum," Educational and Psychological Measurement, Vol. XIX, No. 4 (1958), 649-651.

TABLE 3[a]

CORRELATION OF CERTAIN TEST AND SUB-TEST SCORES WITH THE FIRST TERM GRADES OF POLICE ADMINISTRATION FRESHMEN

Variable	Correlation
American Council on Education	
Quantitative	.22[b]
Linguistic	.43
Total	.38
M.S.U. Reading Test	
Vocabulary	.46
Comprehension	.44
Total	.50
M.S.U. English Test	.47
M.S.U. Arithmetic Test	.26
California Test of Mental Maturity	
Language	.50
Non Language	.17
Total	.46
Kuder Preference Record C Form	
Outdoor	.02
Mechanical	.05
Computational	-.14
Scientific	.18
Persuasive	.01
Artistic	.04
Literary	.08
Musical	-.11
Social Service	.04
Clerical	-.11
Minnesota Multiphasic Personality Inventory	
Hypochondriasis	.08
Depression	.13
Hysteria	.05
Psychopathic Deviate	.13
Masculine-Feminine	-.01
Paranoia	.23
Schizophrenia	.06
Mania	.04

Table 3 -Continued

Variable	Correlation

Thurstone Temperament Schedule

Active	-.05
Vigorous	-.08
Impulsive	-.11
Dominant	.00
Stable	.04
Sociable	.02
Reflective	.20

[a] After King, Norrell and Erlandson (1959)

[b] r's greater than .148 and .113 are significant to the .01% and .05% levels respectively (when N = 300)

Single test or sub-test correlations with college grades are generally poor, at best ranging around .50 and with large differences between tests. It is also interesting to note that different researchers often get entirely different results even when using the same tests as predictor variables. Table 4 shows, in summary form, the results of three studies correlating the verbal and

TABLE 4

CORRELATIONS OF FRESHMEN GPA's WITH TWO SCALES
OF THE SCHOLASTIC APTITUDE TEST

Authors of study	Correlations with GPA	
	SAT verbal	SAT quantitative
Griffen and Flaugherty[9]	.63**	.59**
Watley and Merwin[10]	.28***	.45***
Michael, Baker and Jones[11]	.35** (men) .44** (women)	.24* (men) .39** (women)

* Significant to the .05 level

** Significant to the .01 level

quantitative scales of the Scholastic Aptitude Test with grades for college Freshmen. This table shows the same predictor variables yielding significantly different correlations when used with different samples of students. This is rather typical of studies in this area and would seem to indicate that test scores are not the only variables involved in academic performance.

SPECIALIZED TEST BATTERIES

One line of thought that has involved several researchers is that perhaps specialized tests should be designed for each distinct academic area. Underlying this thought is a feeling that trying to predict the performance of a group of college students majoring in different areas may tend to cancel out predictive validity for some areas. That is, if tests are given to students in several academic areas, very high scores for some groups of students and very low ones for others could cancel each other out and yield nothing but an average figure incapable of predicting anything.

Providing at least partial support for the idea of giving different tests in different academic areas is a 1964 study by Lewis,

---

*** No data given on significance.

[9] Mary Louise Griffin and Sistem M. Rita Flaherty, "Correlation of CPI Traits with Academic Achievement," Educational and Psychological Measurement, Vol. XXIV, No. 2 (1964) 369-372.

[10] Donivan J. Watley and Jack C. Merwin, "The Effectiveness of Variables for Predicting Academic Achievement for Business Students", The Journal of Experimental Education, (Winter, 1964) 189-192.

[11] Michael, Baker and Jones, loc. cit.

Braskamp and Statler.[12] In this study they administered a speci-

ally designed Law School Admissions Test (LSAT) for correlation

TABLE 5[a]

GPA CORRELATIONS WITH A SPECIALLY DESIGNED TEST FOR
PREDICTING SUCCESS IN A COLLEGE OF LAW

Test given	Correlation
Law School Admissions Test	
Total	.39[b]
Writing ability	.28
General background	.18
Undergraduate GPA	.26

[a]After Lewis, Braskamp and Statler

[b]All r's over .20 significant to 1%
level (when n = 180)

with first semester grades in a Law school. (See Table 5). Of

the four measures employed, the LSAT total did yield the highest

correlation (r = .39), and may indicate a relationship between the

test and Law School performance, but a tenuous one at best.

A 1964 study by Boe[13] also gives some support to this idea,

but in a rather indirect manner. He administered a test battery to

Engineering students and then computed correlations with grades

[12]John W. Lewis, Larry Braskamp and Charles Statler, "Pre-
dicting Achievement in a College of Law," Educational and Psycho-
logical Measurement, Vol. XXIV, No. 4 (1964) 947-949.

[13]Erling E. Boe, "The Prediction of Academic Performance of
Engineering Students," Educational and Psychological Measurement,
Vol. XXIV, No. 2 (1964) 377-383.

earned in Engineering courses, in non-Engineering courses, and in all courses combined. Table 6 shows that when the correlations were broken into two groups they appeared to increase for one group (non-Engineering courses) and decreased for the other group (Engineering courses).

TABLE 6[a]

TEST CORRELATIONS WITH GRADES IN ENGINEERING
AND NON-ENGINEERING COURSES FOR
ENGINEERING MAJORS

	Correlations		
Tests given	Engineering courses	Other courses	Combined
American Council on Education			
Linguistic	.209*	.626**	.358**
Quantitative	.037	.184*	.107
Total	.162	.450**	.298
Cooperative English Test	.282**	.561**	.430
Wisconsin State College Math Placement Test	.344**	.407**	.418**
High School GPA	.327**	.556**	.451**

*Significant to .05 level

**Significant to .01 level

[a] After Boe (1964)

The greatest change was with the Linguistic scale of the ACE which appeared to increase by .168 (from .358 to .526) for non-Engineering courses and decrease by .149 (to .209) for Engineering courses. Two other facts stand out in this table: (1) co-

rrelations for non-Engineering courses were consistently more significant, and (2) the lowest correlation with Engineering GPA was, surprisingly, the Quantitative scale of the ACE. English and Linguistic measures seemed to provide the best correlations for non-Engineering courses and the specially designed WSC Math Placement Test was the highest for Engineering areas. These results would seem to indicate that different variables are influencing grades in the different areas, and that special tests may be appropriate.

However, since the operating variables are so difficult to isolate, not all attempts to predict with specialized batteries have succeeded. In a 1961 study, Ellis[14] arrived at the conclusion that specialized test batteries did not seem to be worth the extra effort involved in their preparation and administration. He selected a sample of students majoring in Liberal Arts, Engineering and Commerce; administered a specialized test battery to each group; and then administered a general test battery to all three groups.

TABLE 7

DIFFERENTIAL PREDICTION IN THREE TYPES OF COLLEGE CURRICULA

Curriculum	GPA correlations with		Change
	General battery	Specific battery	
Liberal Arts	.58*	.61	+.03
Engineering	.42	.47	+.05
Commerce	.53	.55	+.02

*r's significant to .01% levels

[14] Kenneth Ellis, "How Effective is Differential Prediction in Three Types of College Curricula", Educational and Psychological Measurement, Vol XXI, No. 2 (1961).

Table 7 shows that the specialized batteries did yield slightly higher correlations, but at best only increased by .05 for students in Engineering. Considering these small changes in correlation values, Ellis would seem to have a valid point in his contention that specially designed test batteries are not justifiable in the light of their limited success.

HIGH SCHOOL PERFORMANCE AS AN INDICATOR

In spite of all the work that has been done with psychological inventories and tests of academic abilities, the best single predictor of college success at the present time seems to be high school performance. The performance criteria may be stated in terms of high school grades earned or on the student's standing in his class, but in both instances correlations with college performance are as high or higher than those obtained through the use of tests and inventories. Table 8 summarizes the results of four studies that compared the predictive potential of high school factors and test results. In each of these studies, high school performance seemed to provide the highest correlations with college GPA's. This is not surprising, as a measure of over-all high school performance would include the effects of many non-academic factors such as family attitude toward education, other interests, etc. that would also effect college performance.

MULTIPLE FACTOR CORRELATIONS

Fortunately, it is not necessary to rely exclusively on single factors when attempting to predict academic success. By combining scores from several measures and assigning different

HIGH SCHOOL RANK AND GPA COMPARED WITH OTHER PREDICTOR VARIABLES

High School measure	Correlation with College performance			Other measures employed in study	Correlation College performance		
	male	female	both		male	female	both
GPA[15]	.49**	.84**		Sequential test of Educational Development (reading)	.35*	.55**	
GPA[16]			.45**	ACE Linguistic			.36**
				Quantitative			.11
				total			.30**
				Cooperative English Test			.43
				Wisconsin State College Math Test			.42**
Rank[17]			.64	Scholastic Aptitude Test Verbal			.56
				Mathematics			.48
				CEEB English Composition Test			.46
				Life Expectency Inventory			.50
Rank[18]			.57	Guilford-Zimmerman Temperament Scale Restraint			.36
				Thoughtfulness			.33
				Scholastic Aptitude Test Mathematics			.45
				Verbal			.28

* Significant to .05 level   ** Significant to .01 level

[15] Boe, loc cit.

[16] John M. Ivanoff, John P. Malloy and Janet R. Rase, "Achievement, Aptitude and Biographical Measures as Predictors of Success in Nursing Training", Educational and Psychological Measurement, Vol. XXIV, No. 2 (1964)

[17] Watley, loc cit.

[18] King, loc cit.

! Compare r's for high school performance and other measures within each study. Read across page.

weights to these scores one can usually arrive at higher correlations than when single factors are used. The study cited earlier by King, Norrell and Erlandson utilized this technique with a slight improvement in correlations. They combined the language scale of the California Test of Mental Maturity (r = .50) with the total score for the M. S. U. Reading Test (r = .50) to obtain an over-all correlation of .56. The study by Lewis, Braskamp and Statler (also cited earlier), achieved similar results. In this instance they combined the composite score for the Law School Admission Test (r = .39) and undergraduate college GPA (r - .26) for a correlation of .43 with grades earned in the first semester of law school. Most other studies being reported today also attempt to combine a number of factors as the basis for their predictions of academic success.

GENERAL COMMENTS

Efforts to seriously predict academic performance are relatively recent, and are plagued with many of the problems common to new disciplines. One big problem area here seems to be a lack of guidelines as to what constitutes sound study design. Many of the researchers in this area seem to have a tendency to study whatever material comes to hand, regardless of how well this material may be suited to quantitative study. Another problem lies in the reporting of these studies in that guidelines are again lacking and that some writers seem to omit almost as much data as they include. Many of their reports are sketchy at best, often leaving out important information on sample size, statistical significance of find-

ings and research methodology. A writer may claim a correlation of .85 or more between some test and academic achievement, but if he fails to explain how he arrived at a particular figure his results can not be taken too seriously.

In an area as new and as lacking in background knowledge as the field of psychological testing, even negative results are valuable. They at least say something about what is not so, provided, of course, that they are reported in such a manner as to inspire confidence. In contrast, the poorly reported or documented study really says very little, no matter what it may claim to say. Fortunately there have been enough well documented studies so that some general comments can be made concerning the "state of the art".

If we assume that the published literature reflects developments in the field of academic prediction, several general conclusions can be drawn.

1. Tests measuring academic ability seem to give higher correlations than do psychological inventories.

2. Those tests that measure a student's linguistic proficiency are better predictors than those based on quantitative skills. This is often even true for areas such as engineering and the sciences that tend to rely heavily on quantitative skills.

3. Specialized test batteries seem to do a slightly better job of predicting college success than do general batteries when given to homogeneous groups.

4. High school performance is as valid an indicator of

academic potential as most test batteries being administered today. This remains true even when high school performance is compared to scores on tests specifically designed to predict academic success.

5. Women may be more easily predictable than men (in relation to college performance).

6. If two different people or groups study the same test or battery of tests they are likely to get significantly different results.

7. We really do not know enough about the prediction of academic success to be thoroughly satisfied with the outcome of much of the work that has been done in this field. The best that can be said is that high school performance and some test scores do, in some way, seem to be related to academic performance. However, the relationship is far from perfect.

In consideration of these shortcomings, one must be extremely cautious in attempting to predict academic success. Actually, simply trying to predict success presents about as many problems as it does results. It is in the use of these correlations that caution is necessary. Since correlations between test variables and academic performance are generally weak, there will be numerous exceptions. With our present knowledge, any attempt to base college admission procedures solely on admissions tests or high school performance would result in many good students being rejected and many poor ones admitted. There are simply too many variables operating that cannot even be identified.

Watley[19] has a particularly graphic way of illustrating the problem involved in this imperfect relationship. He took 109 successful (above a 2.00) and 39 unsuccessful (below a 2.00) first term college freshmen and plotted cumulative percentage curves for their scores on the Guilford-Zimmerman Restraint Scale. Table 9 shows these curves. Each point on a curve shows the percent of students in that category with scores below the corresponding percentile on the Restraint Scale. The reason for constructing these curves was to illustrate visually the effects of establishing an arbitrary admissions cut-off on the basis of such a test. For example, if the thirtieth percentile was chosen as the lowest acceptable score, the dotted lines at the lower left of Table 9 show that almost thirty percent of the unsuccessful students would have been rejected. However, this same cut-off point would also have rejected almost ten percent of the successful students. Since the curves are so close to parallel, any arbitrarily established cut-off point will eliminate some of both types of students and allow some of both types to remain.

While correlation of the trait with GPA was only .36, it still illustrates why caution must be exercised in the area of academic prediction. Any correlation below a perfect 1.00 would also have this tendency to eliminate some good students and admit poor ones.

In spite of the apparent confusion in this field of prediction, we do have many good indications of what is or is not so,

---

[19] Watley, loc. cit.

TABLE 9

CUMULATIVE PERCENT CURVES FOR SCORES OF SUCCESSFUL
AND UNSUCCESSFUL STUDENTS ON THE RESTRAINT SCALE

Percentile on Restraint trait
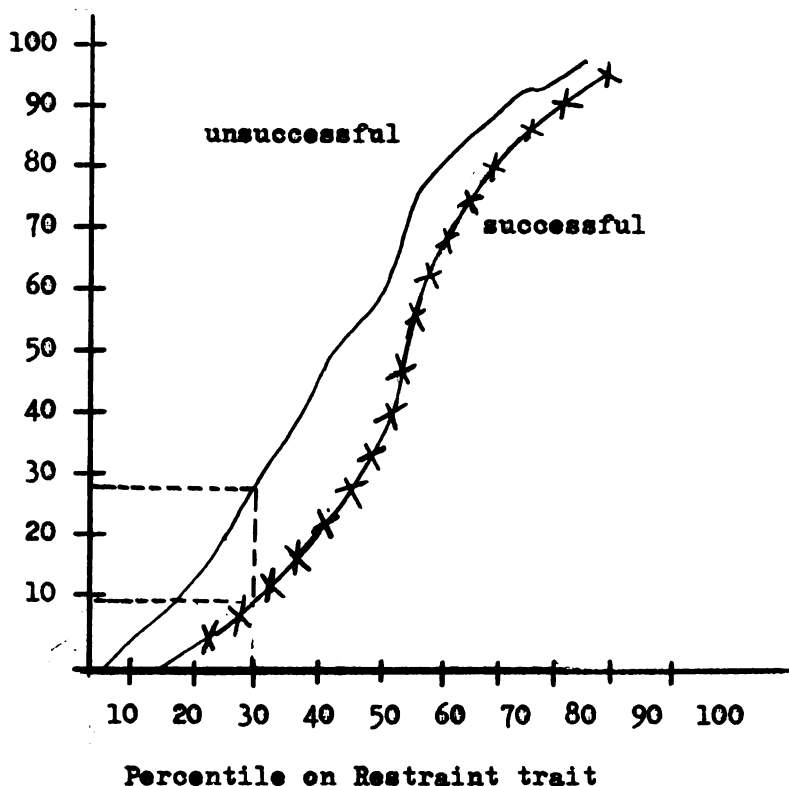
and have established some guidelines for continuing research.
This review of the literature does, however, point out several
areas that are very much in need of improvement. One of these
areas is what one might term the spontaneous study where the "re-
searcher" seems to study whatever lies at hand simply for the sake
of studying it; and perhaps to get his name in print. There even
seems to be a standard recipe for this type of study.

> "Start with an existing batch of test results given and
> scored by someone else; fold in an equal amount of suc-
> cess criteria; strain through an IBM sorting machine;
> and then bake in the largest and most impressive com-
> puter available. Finally, analyze results. Caution:

do not analyze too critically."[20]

The need is for more studies that are carefully planned from the beginning of the testing program. In this way research controls can be placed on any factors that may tend to bias or invalidate the results. In addition, there should be more unified and coordinated research programs in this area. "Hit or miss" research programs may well add to our knowledge, but they do so in such a disorganized manner that the value of their contributions is minimized.

Finally, when research is reported, it should be reported completely. A statement of "results" or "findings" is virtually meaningless unless it is accompanies by other information supporting the findings in such a manner that they can be accepted by other persons working in the same area.

---

[20]Ye Olde Recipe Book for Graduate Student Research Projects, (formerly titled: Forty Fast Ways into Print and Higher Paying Positions).

CHAPTER FOUR

STUDY DESIGN

PRELIMINARY PROBLEMS

This particular study, like so many in the area of academic

testing, is something of a compromise measure. The testing program

concerned was not, apparently, established with this type of study

in mind. The full length of the testing period was from the fall

term of 1953 through the fall of 1964. During this period about

1,200 students were actually tested. The final sample for this study,

however, consists of only 277 students representing a four year span

of time.

There were several reasons for this selection of usable test

records. The primary one was a major change in the composition of

the test battery. From its beginning in 1953 through 1958 the bat-

tery consisted of the following tests:

    Michigan State University Engrance Examination (American
    Council on Education Psychological Examination)

    Kuder Preference Record - Occupational C Form

    Minnesota Multiphasic Personality Inventory

    California Test of Mental Maturity

580 students took at least part of this test battery, and 466 of

them had complete sets of scores. A significant complication was

that many of the available scores were not reported in a manner

suited to correlation analysis. To obtain a meaningful coefficient

of correlation between two variables, there must be a fairly large range within which each variable can fluctuate. For example, grade point average could fluctuate from 0.00 (failed all courses) to 4.00 (perfect), while a test score reported in percentiles could vary from 00 (bottom) to 99 (top).

Unfortunately, many of the scores during this period were reported in terms of rather narrow categories. For example, scores for the California Test of Mental Maturity were given only as "high", "medium high", "medium", "medium low", and "low". The Kuder Preference Record was also, unfortunately, scored in terms of "high", "medium", and "low" for each scale. Since the actual scores that the students obtained on these tests were in many instances no longer available, this first test period had to be disregarded.

During the second testing period the following tests and scales were used:

Michigan State University Entrance Examinations

College Qualification Tests with verbal, numerical, general information and total score scales

M. S. U. Reading Test

M. S. U. English Placement Test

Otis Test of Mental Ability

Minnesota Multiphasic Personality Inventory (3 validity scales and 10 clinical scales)

Kuder Preference Record D Form

Strong Vocational Interest Inventory

A total of about 600 students took at least part of this battery of tests but several considerations dictated a reduction in the number

of records used. One reason for this was that the records for each student had to be complete for every test, so any records with missing scores were discarded. Another reason was that the study required full data on the subjects' performance in college. Because of this requirement no tests given after the fall of 1961 could be used since the later groups would not have completed their University program when this study began. In addition to a reduction in sample size, the Kuder Preference Record and the Strong Vocational Interest Inventory had to be eliminated because of the same qualitative scoring that necessitated the rejection of the earlier half of the testing program. Finally, since this study deals with academic prediction in a Police Administration curriculum, all test records for students not in Police Administration were rejected.

THE FINAL SAMPLE

The final sample consisted of test records that met the following criteria

1. The scores for the various tests had to be reported in a quantitative manner, as raw scores or percentiles.

2. Each record used had to have scores for the M. S. U. Entrance examinations, the MMPI and the Otis Test of Mental Ability.

3. The student must have taken the test battery no later than the fall term of 1961 to allow him to graduate prior to the beginning of this study.

4. The student must have graduated with a degree in Police Administration and Public Safety.

The group of test scores finally selected for the study consisted of 277 sets of records spread out over the following terms:

TABLE 10

NUMBER OF SUBJECTS WITH COMPLETE
TEST SCORES IN EACH TERM

Term	Year	Number of Subjects
Fall	1958	61
Winter	1959	15
Spring	1959	10
Fall	1959	90
Winter	1960	16
Spring	1960	15
Winter	1961	16
Spring	1961	11
Fall	1961	43
Total		277

This represents a sharp reduction from the ideal of including the records of every student taking the tests, but is the largest sample that can be employed in this study. It should be noted that this sample includes all students in the School of Police Administration and Public Safety for which sufficiently detailed records were available. The sample was not limited to students who successfully completed college. In fact, some of the students failed to complete more than one term.

CRITERIA OF ACADEMIC SUCCESS

Selection of measures of academic success was limited by two major factors. The first of these was inherent in the design of the study. Those measures utilized had to be suitable for correlation analysis in that they would have to be capable of varying over a fairly wide range of scores. They also had to be of such a nature as to not require too much office staff time to obtain from the student records.

The success criteria finally settled upon were three:

1. All university grade point average (GPA). This consists of the student's total number of honor points divided by credits earned. An "A" is worth four honor points, a "B" is worth 3, "C" is worth two, "D" is worth one, and "F" is worth zero. In this manner a four credit "A" would be worth 16 honor points (four credits x four points = 16). Similarly, two three credit "B's" and two three credit "C's" would yield a GPA of 2.50.

2. Grade point average in Police Administration courses. This figure is computed in the same manner as the all-university GPA except that only courses taught in the school of Police Administration and Public Safety were counted. It was hoped that this figure would provide a measure of just how well the student performed in course work directly related to his chosen profession.

3. Average number of credits carried per term. This figure was obtained by dividing the total number of credits earned while in the university by the number of terms

spent earning those credits. The reason for including
this figure was to provide added meaning to the all-
university GPA by making allowances for below-average
students who could only maintain passing grades by
taking light credit loads or above average students
who carried exceptionally heavy study loads.

RECORDING PROCEDURES

The recording of information relative to academic success was
performed by the staff in the office of The School of Police Adminis-
tration and Public Safety. The data were computed and entered on the
reverse sides of the test score report forms in the following manner:

1. The student's all-university GPA is an official figure com-
   puted by the university. The source for this figure was
   the "Report of Standings" issued by the university for
   each student's final term in college.[22]

2. GPA in Police Administration courses was computed by list-
   ing all Police Administration courses taken along with the
   credit values and grades for each course. Dividing total
   honor points in these courses into the total number of
   Police Administration credits carried gave the desired
   measure.

3. Computing a student's average credit load per term was
   simply a matter of dividing the total number of credits
   taken by the number of terms spent as a full or part-time

---

[22]The "Report of Standings" is issued by the University for each
student at the close of each term and shows courses taken, grades
earned, and cumulative totals, including all university GPA.

student here. This figure does not include any credits

for courses taken at other schools.

As the material was recorded it was periodically checked against the

files by the writer and one of the office staff.

The final recording step was to enter all data, test scores

and success criteria on IBM forms to serve as a guide for the card-

punch operators at the Michigan State University Computer Center.

The information entered on the IBM forms was checked for accuracy by

the writer before it was sent to the computer center.

STUDY DESIGN

The study was divided into three major steps, each involving

about one-third of the total sample (see Table 11) and each step

requiring processing:

1. Run a correlation analysis between the test scores for

   the first group (N - 1/3 of total population) and the suc-

   cess criteria. Then identify those test or sub-test vari-

   ables in the test battery that seem to correlate well with

   the criteria. Eliminate items with relatively low corre-

   lations (r's).

2. Using the surviving items from step one, and the second

   one-third of the population, develop a multiple regression

   equation for predicting academic performance from the

   scores.

3. On the basis of the regression equation obtained in step

   two, and using the test scores of the third group, predict

   how well the third group will do academically. Run a

   correlation analysis between predicted and actual per-

formance for the third group.

If the correlation between predicted and achieved performance for the third group is of the same magnitude as the multiple correlation from the regression equation in step two, the test scores can be thought of as having a consistent relationship with academic performance. The size or importance of this relationship will be indicated by the size of the relative multiple correlations for the two groups.

IMPLEMENTING THE DESIGN

Once the general outline for the study had been decided upon and the sample selected, the actual implementation was a fairly straightforward matter. The first step was to separate the sample into three equal parts, using a table of random digits. This was done by numbering the test scores in each term from one to n. In the first term there were sixty sets of test scores. The first thirty sets of test scores with serial numbers corresponding to the first thirty random digits were withdrawn and labeled Group A. The second thirty sets corresponding with the second thirty random digits was labeled Group B. Group C consisted of the remaining thirty sets of test scores. This same procedure was applied to the remaining eight terms. To further insure random selection of the scores used in the three computational steps, the table of random digits was also used to select the relative order in which Groups A, B and C would be taken from each term. The final arrangement of the samples is shown in Table 11.

TABLE 11

DIVISION OF TERM SUB-GROUPS INTO THE
THREE COMPUTATIONAL STEPS

Term	N	Number of Test Records And Relative Ordering in Each Step		
		Step 1	Step 2	Step 3
Fall, 1958	61	19 (B)	23 (C)	19 (A)
Winter, 1959	15	4 (C)	6 (B)	5 (A)
Spring, 1959	10	3 (C)	4 (A)	3 (B)
Fall, 1959	90	29 (C)	31 (A)	30 (B)
Winter, 1960	16	6 (A)	6 (C)	4 (B)
Spring, 1960	15	5 (C)	6 (A)	4 (B)
Winter, 1961	16	8 (B)	4 (C)	4 (A)
Spring, 1961	11	1 (A)	6 (B)	4 (C)
Fall, 1961	43	11 (A)	16 (B)	16 (C)
Total	277	86	102	89

The next step was to obtain the success criteria from each student's records and enter them on the forms reporting the test results. Once this was finished the data could then be arranged and itemized for computer processing. The variation of totals in the three steps in Table 11 is a result of missing information in the office files.

COMPUTER PROGRAM

The computation of correlation coefficients was done on the University's Control Data 3600 computer as part of a standard "pack-

age" of services available for student research projects. In addition to showing correlations between all variables on the first two runs, the computer output also provided means, standard deviations and the sum of each variable over all observations. The third computer run supplied single and multiple correlation coefficients and certain information relative to statistical significance.

Table 12 - continued

	Variable	Mean	With All-Univ. GPA		With PLA GPA	
			r	sig.	r.	sig.
6.	Information	45.9*	.21	5%	.14	
7.	Numerical	40.4*	.09		.09	
8.	Total	44.0*	.22	5%	.16	
9.	MSU English	40.6*	.24	5%	.18	
10.	MSU Reading	45.5	.41	1%	.35	1%
11.	MMPI L	4.0	-.11		-.05	
12.	F	3.9	-.12		-.12	
13.	K	17.0	.12		.04	
14.	Hs	12.5	-.02		-.13	
15.	D	17.2	-.04		-.10	
16.	Hy	20.1	-.01		-.06	
17.	Pd	22.5	.10		.05	
18.	Mf	23.2	-.11		-.08	
19.	Pa	9.0	-.20		-.17	
20.	Pt	26.3	-.14		-.23	5%
21.	Sc	25.6	-.09		-.16	
22.	Ma	20.1	-.13		-.04	

[a]Space left blank indicates significance level below 5%.

*Percentile scores based on total number of entering students who took the M. S. U. entrance battery in a given term.

The most obvious thing about Table 12 is that none of the independent variables seem to correlate particularly well with academic performance. For prediction purposes, the MMPI was a dismal

CHAPTER FIVE

RESULTS OF STUDY

THE FIRST COMPUTER RUN

The first computer run was based on data for eighty six sub-
jects and compared twenty two variables for each subject. The
dependent variables were the two grade point averages while the
independent variables consisted of the various test scores and the
average number of credits carried per term. As was mentioned in
the preceding chapter, the primary purpose of the first run was to
reduce the number of variables by eliminating those bearing no
strong relationship with academic performance. Table 12 shows the
computer output relevant to this phase of the study.

TABLE 12

AVERAGES OF VARIABLES AND CORRELATIONS WITH ALL-UNIVERSITY
AND POLICE ADMINISTRATION GRADE POINT AVERAGES

Variable	Mean	With All-Univ. GPA r	With All-Univ. GPA sig.	With PLA GPA r	With PLA GPA Sig.
1. All University GPA	2.37			.89	1%
2. PLA GPA	2.48	.89	1%		
3. Credits per term	14.6	.30	1%	.25	5%
4. Otis IQ score	113.	.20	a	.17	
5. CQT Verbal	47.1*	.18		.12	

CHAPTER FIVE

RESULTS OF STUDY

THE FIRST COMPUTER RUN

The first computer run was based on data for eighty six subjects and compared twenty two variables for each subject. The dependent variables were the two grade point averages while the independent variables consisted of the various test scores and the average number of credits carried per term. As was mentioned in the preceding chapter, the primary purpose of the first run was to reduce the number of variables by eliminating those bearing no strong relationship with academic performance. Table 12 shows the computer output relevant to this phase of the study.

TABLE 12

AVERAGES OF VARIABLES AND CORRELATIONS WITH ALL-UNIVERSITY AND POLICE ADMINISTRATION GRADE POINT AVERAGES

Variable	Mean	With All-Univ. GPA		With PLA GPA	
		r	sig.	r	Sig.
1. All University GPA	2.37			.89	1%
2. PLA GPA	2.48	.89	1%		
3. Credits per term	14.6	.30	1%	.25	5%
4. Otis IQ score	113.	.20	a	.17	
5. CQT Verbal	47.1*	.18		.12	

failure with most of the correlations clustering around zero. The large number of negative correlations between the various scales of this instrument and academic success is to be expected since high scores on most MMPI scales could be indicative of mental or emotional problems. Most of these problems would probably reduce a student's effectiveness. It should be noted that, aside from a larger number of negative correlations, these findings on the MMPI tend to correspond with the findings of King, Norrel and Erlandson. (See Table 3). On the basis of the results of this first computer run, the MMPI was dropped from further consideration.

The highest correlation in the first run was .89, but this simply deals with the relationship between the two grade point criteria and has no predictive value. A student who does well in one area of academic study can generally be expected to do well in other areas also. Average credits per term show a significant positive relationship with academic performance, but being an "after the fact" calculation, cannot be used for prediction. In addition, neither of these correlations represents a cause and effect relationship. Rather, the pairs of items show a somewhat consistent relationship because they are influenced by the same outside factors. For example, while the better students will tend to carry heavier course loads, these course loads do not determine their grades. Their grades are determined by generally superior scholarship.

According to the results of Table 12, the most promising instruments seem to be the specially designed MSU English and Reading tests. Their correlations with the criteria were .24 and .41

respectively, for all-University Grade Point Averages, and .16 and .35 for Police Administration Grade Point Averages. Only the MSU Reading Test was statistically significant for both GPA's. Both the CQT and the Otis Test of Mental Ability yielded disappointingly low correlations with the criteria. They do seem to bear some relationship to success, but only the CQT Information and Total scores yield statistically significant correlations with the success criteria; and then only in relation to all-University grade point average.

The correlation matrix shown in Table 13 gives the intercorrelations between all variables except the various MMPI scales. These were eliminated from this table because they fail to correlate significantly with anything but themselves. It should be noted that, on the whole, the tests correlate better with themselves than with either of the two success criteria. For example, variable number 8, the CQT Total, shows correlations of only .22 and .16 with the GPA's. However, when the total is compared to the sub scores of the CQT; namely Verbal, Information and Numerical, the correlations are .80, .78 and .63, respectively. A comparison of the CQT Total and the Otis IQ Test shows the same type of fairly high relationship. The correlation between two measures is .66, much higher than either measures correlation with the success criteria. From this it would seem that the tests are doing a better job of measuring some other factor than they are of measuring academic performance. It appears that the tests are measuring academic mental ability and achievement, but that there is more to succeeding in college than simply possessing mental ability and prior academic achievement.

TABLE 13

CORRELATIONS OF SELECTED VARIABLES WITH THE TWO GRADE POINT AVERAGES AND WITH EACH OTHER FIRST COMPUTER RUN

Variable	1	2	3	4	5	6	7	8	9	10
1. All University GPA	1.00									
2. PLA GPA	.89	1.00								
3. Average credits	.30	.25	1.00							
4. Otis IQ test	.20	.18	.17	1.00						
5. CQT Verbal	.18	.12	.18	.54	1.00					
6. Information	.21	.14	.07	.48	.48·	1.00				
7. Numerical	.09	.09	.12	.47	.17	.36	1.00			
8. Total	.22	.16	.20	.66	.80	.78	.63	1.00		
9. MSU English	.24	.18	.04	.53	.56	.35	.41	.59	1.00	
10. MSU Reading	.41	.35	.19	.48	.56	.50	.43	.67	.54	1.00

r's greater than .217 and .283 are significant to the .05 and .01 levels, respectively.

THE SECOND COMPUTER RUN

It was decided to eliminate the MMPI and average credits per
term from further consideration for reasons indicated above, and to
base the second run on the Otis Test, the CQT scores and the M.S.U.
English and Reading tests. This second phase of computer work was
based on a sample of 102 students. In addition to developing re-
gression equations, the second-phase computer output also included
single item correlations between all variables as in the first run.

Unfortunately, the second phase results did not match the wri-
ter's expectation of what they should have been. The idea behind
random selection of the three groups was to insure, insofar as pos-
sible, equal distribution of talent among the three groups. If the
equal distribution of talent and other relevant factors could be
achieved, one would expect the correlations within the three groups
to be stable from one computer run to the next. Table 14 clearly
shows that the correlations between test scores and success cri-
teria within the first and second groups were anything but stable.
All correlations between test scores and criteria in the second
group were higher than in the first group. In some instances the
increases were minor, but many correlations more than doubled in
numerical value from the first to the second run. While the right
hand column in Table 14 shows that only two of these increases are
significant to the .05 level, the fact still remains that all of
the correlations increased by at least a few points. (Significance
data were derived from a table of Fisher's Z.)

While it is remotely possible that these increases may not be
significant, a more likely assumption is that bias was somehow

TABLE 14

A COMPARISON OF SELECTED PORTIONS OF
THE RESULTS OF THE FIRST AND
SECOND COMPUTER RUNS

	Correlations with all-University GPA			
Variable	First run	Second run	Increase	Z*
Otis IQ Test	.20	.30	.10	.718
CQT Verbal	.18	.41	.23	1.705
Information	.21	.45	.24	1.825
Numerical	.09	.39	.30	2.16 *
Total	.22	.53	.31	2.464*
M.S.U. English	.23	.30	.07	.506
M.S.U. Reading	.41	.47	.06	.501
	Correlations with PLA GPA			
Otis IQ Test	.17	.22	.05	.350
CQT Verbal	.12	.33	.21	1.490
Information	.14	.35	.21	1.500
Numerical	.09	.30	.21	1.470
Total	.16	.41	.25	1.84
M.S.U. English	.18	.23	.05	.350
M.S.U. Reading	.35	.38	.03	.23

*When Z is equal to or greater than 1.96, difference is
significant to the .05 level.

introduced into the sampling and sorting procedures early in the

study. A comparison of the correlation matrices for the two runs

(Tables 13 and 15) supports this assumption. The reader will note

TABLE 15

CORRELATIONS OF SELECTED VARIABLES WITH THE TWO GRADE POINT AVERAGES AND WITH EACH OTHER
SECOND COMPUTER RUN

Variable	1	2	3	4	5	6	7	8	9
1. All University GPA	1.00								
2. PLA GPA	.90	1.00							
3. Otis IQ Test	.30	.22	1.00						
4. CQT Verbal	.41	.33	.57	1.00					
5. Information	.45	.35	.52	.49	1.00				
6. Numerical	.39	.30	.45	.20	.57	1.00			
7. Total	.54	.41	.66	.79	.84	.67	1.00		
8. M.S.U. English	.30	.23	.48	.46	.39	.42	.55	1.00	
9. M.S.U. Reading	.47	.38	.57	.77	.55	.34	.75	.59	1.00

r's greater than .195 and .254 are significant to the .05 and .01 levels, respectively.

that while test score-GPA correlations increased, the inter-correlations between test variables do not vary to any appreciable degree. While some of them do vary slightly, the variations seem to be small and random. This would seem to indicate that whatever bias may exist in the samples is primarily related to success factors, and not to the test scores. Exactly what this bias is, or how it operates, is unclear. In fact, there is no assurance than any bias even exists. However, the likelihood of accidental bias was such that certain portions of this study were repeated. Before examining the results of this replication, we would do well to return to the second and third stage results of the original study plan.

The reader will recall from Chapter Three that the primary purpose of the second run was to develop regression equations for predicting academic success on the basis of test scores. These equations were developed by an automatic computer process that screened out test variables that failed to contribute significantly to the equation, and retained those that did. In this particular instance, the criteria for including any given test variable in the equation was the statistical significance of the variable when compared to one of the success criteria. Theoretically, the regression equations could have included any number of the seven test variables found to bear a significant relationship to the success criteria on any reasonable level of statistical significance. (The .001, .005, .01 and .05 levels of significance were tried for all seven variables against both success criteria.)

Since only the CQT Total seemed to be significant, the regression equations were calculated for just that one variable.

The formula used was:

$$Y = r_{xy} \left(\frac{Sy}{Sx}\right) \ X - r_{xy} \left(\frac{Sy}{Sx}\right) \ \overline{X} + \overline{Y}$$

where $Sx = \sqrt{\dfrac{\Sigma x^2}{Nx-1}}$ $Sy = \sqrt{\dfrac{\Sigma y^2}{Ny-1}}$

and $\Sigma x^2 = \Sigma X^2 - \dfrac{(\Sigma x)^2}{N}$ $\Sigma y^2 = \Sigma Y^2 - \dfrac{(\Sigma Y)^2}{N}$

Y refers to the predicted GPA

r_{xy} is the correlation between CQT Total

and one of the success criteria

When the actual values were inserted into the formulas, the equations evolved as follows:

for predicting all-University GPA

Y = .016 (CQT Total) + 1.49

for predicting PLA GPA

Y = .01 (CQT Total) + 1.89

CROSS VALIDATION

The second computer run and related statistical manipulations yielded a regression equation for predicting academic success. The purpose of cross validation was to see if this equation, derived from one sample of students, will successfully predict the performance of another sample on the basis of their own test scores. The first step in this process was to use the regression equations

derived above to predict all-University and PLA GPA's for the third group.

After computing the predicted GPA for each student in the third group, the following formula was used to compute the correlation between GPA's predicted and actually achieved.

$$r = \frac{N\Sigma(XY) - (\Sigma X)(\Sigma Y)}{(N\Sigma X^2 - [\Sigma X]^2)(N\Sigma Y^2 - [\Sigma Y]^2)}$$

Table 16 shows the results of this cross validation in summary form.

TABLE 16

CROSS VALIDATION OF REGRESSION EQUATIONS:
R's BETWEEN PREDICTED AND ACHIEVED
GPA's AND Z VALUES OF CHANGES

	All University GPA	PLA GPA
Group 2	.534	.412
Group 3	.427	.273
Z	0.89	1.08

Z of less than 1.96 indicates
changes are not significant.

When subjected to cross-validation, the predictive efficiency of the regression equations seemed to be reduced for both GPA's. The correlations between predicted and achieved all-University GPA dropped from .534 to .427, while the PLA correlation dropped from .412 to .273.

From this it would seem that a student's total performance on the CQT bears a consistently significant relationship with both his over-all academic performance and his performance in Police Administration courses. The relationship seems to be stronger for over-all university performance than for the portion of this performance represented by the student's grade point average in Police Administration courses. The actual importance of this relationship will be explored in detail in Chapter VI, after we have covered the additional findings resulting from re-running portions of this study.

CROSS CHECKING FOR BIAS

Since it seemed highly possible that the original sampling procedures in this study may have been biased, it was decided to run some cross checks on the extent of any possible bias. The first step was to recombine all the usable test records for a new division into two groups. The procedure here was to assign random numbers to the test records in each term, and then to pick half of each term's records through the use of a table of random numbers. This process resulted in two groups of test records (N = 140 and 141) with an approximately equal number of test records from each term.

For the first group (N = 140), correlations were computed between CQT Total scores and the two grade point criteria. These correlations turned out to be .651 for all-University GPA and .307 for PLA GPA. The formula shown on page V-10 was then used to compute regression equations for predicting the two GPA's for students in the second group (N = 141). The final equations were:

for all-University GPA

$$Y = .053 \ (CQT \ Total) + .08$$

for PLA GPA

$$Y = .037 \ (\text{CQT Total}) = .86$$

The results of this cross checking are shown, in summary form, in

Table 17

TABLE 17

SIMPLE CORRELATIONS BETWEEN CQT TOTAL AND THE TWO GPA's
AND RESULTS OF CROSS VALIDATION OF
THE REGRESSION EQUATION

	Correlations with	
	All University GPA	PLA GPA
CQT total and Academic Performance First Group	.651	.307
Predicted and Achieved Performance Second Group	.545	.281
Z	1.13	0.27

As in Table 16, the correlations dropped somewhat when subjected to

cross validation. The Z values of these changes are low enough to

warrant the assumption that these drops are not statistically sig-

nificant. Again, it appears as if a student's total score on the

College Qualification Test is at least partially indicative of how

he will perform academically in college.

CHAPTER SIX

ANALYSIS OF RESULTS AND CONCLUSIONS

COMPARISON WITH OTHER STUDIES

By now the reader may have decided that the results of this study are little different from the results of most of the studies cited in Chapter III. In one respect at least, this is quite true. The relationships revealed by this study are far from strong and only explain or account for a small portion of the factors influencing academic success, as discussed below. However, the cross validation built into this study insures that a greater degree of confidence can be placed in these findings than in the results of most correlation studies concerning academic success reported in the literature. Most of the writers in this area seem content simply to report correlations between variables and, evidently assume that their discovered relationships will hold constant. Very few writers seem to make any effort at cross validation.

In this study we have seen that many of the correlations between test scores and success criteria do not hold constant. However, we have also seen that this fluctuation does not destroy the predictive capabilities of the test measure studied most closely - the CQT Total score. And, finally, the cross validation procedures in this study (done twice, actually) leave one with a feeling of relative confidence in these results.

THE MEANING OF THE STATISTICAL RESULTS

Several of the test scores studied seem to bear some relationship to the academic success of Police Administration majors at

Michigan State University. The most notable among these are the specially designed M. S. U. English and Reading tests, and the total score for the College Qualification Test. Other test measures, such as the MMPI, seem to have no significant bearing on the subject.

After studying the CQT so thoroughly, the question now arises as to the practical significance of these findings. During the course of this study, correlations between the CQT Total and all-University GPA ranged from a low of .22 to a high of .65. One way of looking at the meaning of a correlation is to square the figure. This new figure is an estimate of how much of the variation in the dependent variable is accounted for, or associated with variation in independent variable. For example, the above correlations of .22 and .65 indicate that variation in the CQT Total score (or more precisely, the academic ability measured by this score) accounts for, or is associated with, from 5% to 42% of the variation in all-University grade point averages.

In summary, it would seem that the level of intellectual ability indicated by a student's total score on the College Qualification Test consistently bears a small but fluctuating relationship to his academic performance. This relationship is stronger for all-University grade point average than for Police Administration grade point average. That this relationship is small is not surprising when one considers the multitude of variables that can effect academic success. Just a few of these are intelligence, students' attitude toward education, family attitudes toward education, health and personal problems. Unfortunately, we are in no position to say

how or to what extent these other variables may be operating in any
given student. And, finally, as was discussed at the close of Chap-
ter III, any attempts to use test scores such as those studied here
as screening devices for admission should proceed with extreme cau-
tion. There are too many other variables operating about which we
know all too little.

EVALUATION OF THE TESTING PROGRAM

In general, the idea of gathering and studying data relative
to academic success in college is a good one. However, if these
studies are to be effective, to yield meaningful results, the stud-
ies must be well planned and carefully implemented. The former
testing program in the School of Police Administration and Public
Safety fails to meet these requirements. There are several reasons
for this, and the most important of them lies in the area of long
range planning. When the study was first prepared, its originators
did have certain ideas on what the study should accomplish. They
were particularly interested in exploring the personality and inter-
est makeup of entering Police Administration students. There was
also, apparently, some thought given to isolating tests with pre-
dictive value. Unfortunately, and partly because of personnel turn-
over, no single person was ever given continuing responsibility for
supervising the program. As a result of this, the program responsi-
bility passed through the hands of a series of disinterested, or at
best moderately interested but untrained, custodians. Some students
never took the tests, some scores were misplaced, and even the scor-
ing procedures were changed several times. There was some short-run
use made of the individual results for counseling and advising pur-

poses, but there seems to have been no plan for over-all analysis. This made any study of the whole program virtually impossible.

This is not to say that exploratory testing programs should not be conducted in the future by the School of Police Administration and Public Safety. In fact, they should be; but according to some sort of over-all plan. There is no reason that such studies could not serve short term advising and counseling needs as well as accumulating data for intensive analysis.

SUGGESTIONS FOR FUTURE STUDIES

One would hope that the School of Police Administration and Public Safety would extend its leadership role in law enforcement education into the study of the effects of this education and ways to improve its quality. With careful planning and implementation, there is every reason to believe that such studies would make significant contributions to our knowledge in this area.

However, continued research into the relationship between academic ability and academic success may contribute less than studies will in certain other areas. It may well be that studies of this area have reached a temporary "dead end" until such time as we learn more about some of the other factors effecting academic success. Some of these "other factors" of particular interest are the student's emotional stability and maturity, his attitudes toward education, his occupational and vocational interests and his family-educational-cultural background. An area that might be of particular interest to the School of Police Administration and Public Safety is the personality make-up of students in this field. We ought to be concerned with exploring the nature of the people

who are attracted to careers in law enforcement and related areas. There are presently in existence test devices or procedures that will at least estimate the extent of these various factors in a student. Studies in these areas would require time and money, but would not be of prohibitive difficulty. This would hold especially true if the School of Police Administration and Public Safety availed itself of the consulting services available from both the Counseling Center and the Psychology Department prior to initiating further studies.

Finally, there is one area that may well be highly important, but is seldom studied That is the relationship between academic ability, academic success, the "other factors" mentioned and the success and job satisfaction of the student after graduation. This type of study would be very long range, perhaps covering ten to twenty years in all. It would require up-to-date record maintenance the like of which does not exist now. And it would require significant amounts of time and money. However, until information of this nature is gathered and carefully studied, we really cannot assess the effectiveness of the educational offerings of the Michigan State University School of Police Administration and Public Safety. We do not, at the present time, have much accurate information on how college success relates to job success; nor do we know how academic ability, when considered apart from academic success, relates to job performance. And finally, apart from some scores on academic tests, we do not have any accurate estimation of the type of person who is attracted to the field of law enforcement.

In summary, there is a definite need for further study in the

areas mentioned above and many reasons why the Michigan State University School of Police Administration and Public Safety should assist in these efforts. Admittedly, the first attempt was not particularly successful, but it can serve as a beginning. By learning from earlier mistakes, and by utilizing available information sources and consulting services, the Michigan State University School of Police Administration and Public Safety can become a leader in the study and analysis of educational offerings in the various areas of law enforcement and how much these educational offerings actually accomplish. As a result of these studies, the School will be able to continually improve its course offerings, outside services, and general value to the community.

BIBLIOGRAPHY

Books

Buras, Oscar K., The Fifth Mental Measurement Yearbook.

Cronback, Lee J., Essentials of Psychological Testing. 2nd ed. New York: Harper and Brothers, 1960.

Freeman, Harold, Introduction to Statistical Inference. Reading, Pa.: Wesley Publishing Co., Inc., 1963.

Hathaway, Starke R., and Monachesi, Elio D., Analyzing and Predicting Juvenile Delinquency with the MMPI. Minneapolis: Calwell Press, Inc., 1953.

Ostle, Bernard, Statistics in Research. Ames, Iowa: The Iowa State College Press, 1954.

Rand Corporation, A Million Random Digits with 100,000 Normal Deviates. Glencoe, Ill.: The Free Press, 1955.

Journal Articles

Aiken, Lewis R. Jr., "Some Nomographs for Academic Prediction Work," Educational and Psychological Measurement, XXIV, (Fall, 1964), 913-920.

_____. "The Prediction of Academic Success and Early Attrition by Means of a Multiple-Choice Biographical Inventory," American Educational Research Journal, (March, 1964), 127-135.

Bachman, Gerald G., "Prediction of Academic Achievement Using the Edwards Need Achievement Scale," Journal of Applied Psychology, XLVII, (February, 1964), 16-19.

Barger, Ben, and Hall, Everette, "Personality Patterns and Achievement in College," Educational and Psychological Measurement, XXIV, (Spring, 1964), 339-346.

Barnette, Leslie W. Jr., "A Structured and Semi-Structured Achievement Test Applied to a College Sample," Educational and Psychological Measurement, XXI, (Summer, 1961), 647-656.

Bernhardt, Harold E. Jr., "Intraception Test Score and Psychiatry Grade as A Freshman and as A Sophomore Medical Student: A Validation Study of A Subscale of the Edwards Personal Preference Schedule," Educational and Psychological Measurement, XX, (Summer, 1960), 365-379.

Boe, Erling E., "The Prediction of Academic Performance of Engineering Students," Educational and Psychological Measurement, XXIV, (Spring, 1964), 377-383.

Boyce, Richard W., "The Prediction of Achievement in College Algebra," Educational and Psychological Measurement, XXIV, (Spring, 1964), 419-420.

Brown, Frederick G. and Wolins, Leroy, "An Empirical Evaluation of the American College Testing Program," The Personnel and Guidance Journal, (January, 1965), 451-456.

Cash, W. L. Jr., "Predictive Efficiency of Freshman Entrance Tests," Journal of Psychological Studies, (June, 1962), 111-116.

_____."Relationship of Personality Traits and Scholastic Aptitude to Academic Achievement in Theological Studies," Journal of Psychological Studies, (June, 1962), 105-110.

Chahbazi, Powiz, "Use of Projective Tests in Predicting College Achievement," Educational and Psychological Measurement, XX, (Winter, 1960), 339-342.

Demas, George D., and Spolyar, Ludwig J., "Academic Achievement of College Freshmen in Relation to Edwards Personal Preference Schedule," Educational and Psychological Measurement, XXI, (Spring, 1961), 473-479.

Domino, George, "Comparison of the D48, Cattell Culture Fair and Army Beta Tests in A Sample of College Males," Journal of Consulting Psychology, (October, 1964), 468-469.

Ellis, Kenneth E., "How Effective is Differential Prediction in Three Types of College Curricula," Educational and Psychological Measurement, XXI, (Spring, 1961), 459-471.

Endler, Norman S. and Steinberg, Danny, "Prediction of Academic Achievement at the University Level," Personal and Guidance Journal, VIII, (1963), 694-699

French, John W., "Aptitude and Interest Score Patterns Related to Satisfaction with College Major Field," Educational and Psychological Measurement, XXI, (Spring, 1961), 287-294.

Froelich, H. P. and Mayo, G. D., "A Note on Under- and Overachievement Measure," Personnel and Guidance Journal, (March, 1963), 621-623.

Gallagher, Joseph W. and Spencer, Richard E., "Prediction of Success in Basic German at the College Level," Educational and Psychological Measurement, XXIV, (Fall, 1964), 955-960.

Goodstein, Leonard D. and Heilbrun, Alfred B. Jr., "Prediction of College Achievement from the Edwards Personal Preference Schedule at Three Levels of Intellectual Ability," Journal of Applied Psychology, XLVI, (May, 1962), 317-320.

Griffen, Mary Louise, and Flaherty, Sister M. Rita, "Correlation of CPI Traits with Academic Achievement," Educational and Psychological Measurement, XXIV, (Spring, 1964), 369-372.

Hansmeier, Thomas W., "The Iowa Tests of Educational Development as Predictors of College Achievement," Educational and Psychological Measurement, XX, (Winter, 1960), 843-845.

Hewitt, John H. and Rosenberg, Leon A., "The MMPI as A Screening Device in an Academic Setting," Educational and Psychological Measurement, XXII, (Winter, 1962), 129-137.

Ivanoff, John M., Malloy, John P., and Rase, Janet R., "Achievement, Aptitude and Biographical Measures as Predictors of Success in Nursing Training," Educational and Psychological Measurement, XXIV, (Spring, 1964), 389-391.

Juola, Arvo E., "The Development of an Academic Predictor Scale Based on Students' Attitudes Toward Education," The Personnel and Guidance Journal, XLII, (December, 1963), 381-386.

King, Paul, Norrell, Gwen, and Erlandson, F. L., "The Prediction of Academic Success in A Police Administration Curriculum," Educational and Psychological Measurement, XIX, (Winter, 1959), 649-651.

Klugh, Henry E. and Bierley, Robert, "The High School and College Ability Test and High School Grades as Predictors of College Achievement," Educational and Psychological Measurement, XIX, (Winter, 1959), 625-626.

Lewis, John W., "Correlates of Persistence in A College of Pharmacy," Educational and Psychological Measurement, XXIV, (Spring, 1964), 385-388.

_____. "Pre-College Variables as Predictors of Freshman, Sophomore and Junior Achievement," Educational and Psychological Measurement, XXIV, (Spring, 1964), 353-356.

_____. "The Relationship of Selected Variables to Achievement and Persistence in A Masters Program in Business Administration," _Educational and Psychological Measurement_, XXIV, (Winter, 1964), 951-954.

Lewis, John W., Braskamp, Larry and Statler, Charles, "Predicting Achievement in A College of Law," _Educational and Psychological Measurement_, XXIV, (Fall, 1964), 947-949.

Mann, Sister M. Jacinta, S. C., "The Prediction of Achievement in a Liberal Arts College," _Educational and Psychological Measurement_, XXI, (Spring, 1961), 481-483.

Michael, William B., et al., "A Note Concerning the Predictive Validities of Selected Cognitive and Non-Cognitive Measures for Freshman Students in A Liberal Arts College," _Educational and Psychological Measurement_, XXIV, (Spring, 1964), 373-375.

Michael, William B., et al., "High School Record and College Board Scores as Predictors of Success in A Liberal Arts Program During the Freshman Year of College," _Educational and Psychological Measurement_, XXII, (Spring, 1962), 399-400.

Miller, Adam, "General Ability and Interest Measure as Differential Predictors of Academic Achievement," _Educational and Psychological Measurement_, XXIV, (Spring, 1964), 357-362.

Pippert, Ralph, and Archer, Sidney N., "A Comparison of Two Methods for Classifying Underachievers with Respect to Selected Criteria," _The Personnel and Guidance Journal_, (May, 1963), 788-891.

Robertson, Malcolm, and Hall, Everett, "Predicting Success in Graduate Study," _The Journal of General Psychology_, (October, 1964), 359-365.

Schultz, Richard E., and Baker, Robert L., "A Factor Analysis of the Kuder Preference Record - Occupational, Form D," _Educational and Psychological Measurement_, XXII, (Spring, 1962), 97-104.

Sgan, Matthew R., "An Alternative Approach to Scholastic Aptitude Tests as Predictors of Graduation Rank at Selected Colleges," _Educational and Psychological Measurement_, XXIV, (Spring, 1964), 347-352.

Sharp, Bert L., "College Achievement: Its Relationship to High School Achievement Experiences and Test Scores," _The Personnel and Guidance Journal_, XLI, (Summer, 1962), 247-250.

Spaulding, Helen, "The Prediction of First Year Grade Averages in A Private Junior College," Educational and Psychological Measurement, XIX, (Winter, 1959), 627-628.

Vick, Mary Catherine, "Predicting GPA at A Small Southern College," Educational and Psychological Measurement, XXII (Winter, 1962), 795-798.

Warwick, Charles E., "Relationship of Scholastic Aspiration and Group Cohesiveness to the Academic Achievement of Male Freshmen at Cornell University," Human Relations, XVII, (No. 2), 155-168.

Watley, Donivan J., "Type, Location, and Size of High School and Prediction of Achievement in an Institute of Technology," Educational and Psychological Measurement, XXIV, (Spring, 1964), 331-338.

Watley, Donivan J. and Merwin, Jack C., "The Effectiveness of Variables for Predicting Academic Achievement of Business Students," The Journal of Experimental Education, (Winter, 1964), 189-192.

OTHER SOURCES

Hathaway, Starke R., and McKinley, Charnley J., Booklet for the Minnesota Multiphasic Personality Inventory, New York: The Psychological Corporation.

Hickok, Benjamin J., English Placement Test, Form A, East Lansing: Michigan State University.

Juola, Arvo E., Comparative Standings of the Various College and Curriculum Groups on the Orientation Week Examinations, East Lansing: Michigan State University, Fall Terms, 1959, 1960, 1961.

Michigan State University, Reading Test, Form A62, Prepared by The Office of Evaluation Services, East Lansing: Michigan State University.

Michigan State University, The Use of Orientation Test Data, A Report Prepared by the Office of Evaluation Services, East Lansing: Michigan State University, 1960.

Otis, Arthur S., Otis Quick-Scoring Mental Ability Tests, Gamma Test, Form EM, New York: The World Book Company, 1954.

Psychological Corporation, College Qualification Tests, Combined Booklet Edition, Form C. New York: The Psychological Corporation, 1959.

APPENDIX

POLICE ADMINISTRATION TESTING

NAME:_____   DATE:_____

Orientation Test Scores:

   CQT   V___   I___   N___   T___: E___   A___   R___

Otis Mental Ability Test: IQ score__ Below Av._ Av.__ Above Av.__

Kuder Occup. Int., D   (verif. S__D__U__)   MMPI

Title	D.S.	Title	D.S.	Trait	Raw Score
Agri. Agent		Journ		L	
Elect. Eng		Archit.		F	
Farmer		Lawyer		K	
Forester		Ret. Cloth.		Hs	
Minister		Ins. Agent		D	
Editor		Dentist		Hy	
Physician		Vetern.		Pd	
Clin. Psych.		Indust. Eng.		Mf	
Ind. Psych.		Pediatr.		Pa	
Y.M.C.A. Sect.		Psychiatr.		Pt	
School Sup.		Rad. Mngr.		Sc	
Acct.		Int. Decor.		Ma	
Meteor.		H.S. Counsel			
Pers. Mngr.		H.S. Sci. Tea.			
Sales		H.S. Math. Tea			
Prof. Psych.		Chemist			
Civil Eng.		Min.-Met. Eng.			
Mech. Eng.		Druggist			
Couns. Psych.		Job Print.			

Prob. Area____

No Prob. Area____

Strong Vocational Interest Test

	B+ Area	A Area
Artist		
Psychol.		
Architect		
Physician		
Osteopath		
Dentist		
Veterinarian		
Mathematician		
Physicist		
Engineer		
Chemist		
Produc. Mngr.		
Farmer		
Aviator		
Carpenter		
Printer		
Math. Phys. Sci. Teach.		
Ind. Arts Teach.		
Voc. Ag. Teach.		
Policeman		
Forest Serv. Man		
Y.M.C.A. Phys. Dir.		
Pers. Director		
Public Adminis.		
Y.M.C.A. Sect.		
Soc. Sci. H.S. Teach.		
City School Supt.		
Social Worker		
Minister		
Musician		
C.P.A.		
Senior C.P.A.		
Accountant		
Office Man.		
Purch. Agent		
Banker		
Mortician		
Pharmacist		
Sales Mngr.		
Real Est. Sales		
Life Insur. Sales		
Adver. Man		
Lawyer		
Author-Journ.		
Pres., Mfg. Concern		