FAIRCHILD, Patricia Carlette, 1940-
GRADE POINT AVERAGE AND VARIANCE AS CRITERIA OF COLLEGE ACADEMIC PERFORMANCE.

The University of Oklahoma, Ph.D., 1969
Education, psychology

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

GRADE POINT AVERAGE AND VARIANCE AS CRITERIA OF COLLEGE ACADEMIC PERFORMANCE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

PATRICIA CARLETTE FAIRCHILD

Norman, Oklahoma

1969

GRADE POINT AVERAGE AND VARIANCE AS CRITERIA

OF COLLEGE ACADEMIC PERFORMANCE

APPROVED BY

Gerald T. Kowitz

William F. Eck

Herbert R. Hengst

Donald Reynolds

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

The writer wishes to express her sincere appreciation to Dr. Gerald T. Kowitz, her advisor and committee chairman, for his continued guidance and expert advice during her course work at the University of Oklahoma and during the writing of this dissertation. The suggestions offered by Dr. Donald Reynolds, Dr. William Eick, and Dr. Herbert Hengst were greatly appreciated.

Special acknowledgment is extended to Jerry Prather for his extremely helpful and appreciated suggestions concerning the theoretical and statistical design of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

GRADE POINT AVERAGE AND VARIANCE AS CRITERIA

OF COLLEGE ACADEMIC PERFORMANCE

CHAPTER I

INTRODUCTION

In recent years it has become increasingly important to explain or, at least, to account for the academic performance of students, particularly at the college and university level. Due to an increased student population at the college level, facilities and resources of the university have become strained. Because of these limitations, higher standards for admission to colleges and universities are being enforced, and a greater number of students are being refused admittance to college. As a result, the competition for admission to the colleges and universities has reached a new level of intensity. The selection of the most promising students has become a much more difficult task for the institutions of higher learning than it was in previous years. This increased interest in the prediction of academic performance also has stemmed from the increase of national programs, such as the National Merit Scholarship Corporation, which identify and provide advanced training for outstanding students. Naturally, these programs do not want to misspend their finances on unproductive students who will not meet the high caliber performance levels required in such study programs. There are many superior students attempting to enter these scholarship programs, but only the more promising ones are selected.

Fishman (1957) in discussing the factors related to college success, classified them into two large groups. The first group included those "factors predictive of college success," those characteristics of the individual measured before his entrance into college. These included high school performance, general academic ability, general intelligence, achievement in specific subject-matter fields, personality characteristics, personal background and social experience, and study habits. This group of factors would serve the purpose of "selective or guided admission." The second group of factors, "factors mediating college success," was concerned with such phenomena as quality of instruction, personality of the instructor, recognition given for outstanding performance, close friends being on campus or in specific classes, stability of student interests, conflicting stimuli on campus, and physical and emotional health. Many different interrelated factors are involved with whether a student will or will not succeed in college. Much time and effort has been devoted to the investigation of academic performance of students at all levels of the educational system.

Particular emphasis has been given to the study of academic performance of college students and the prediction of this performance. Intellective and ability factors as predictors were the primary concerns of earlier research on academic performance. Intellective measures are those concerned with grades or the measurement of intellectual "products" or "levels," which would include aptitude scores, achievement test scores, intelligence test scores, grades, and academic honors. However, due to the relative inefficiency of intellective factors as predictors of future academic performance, more recently investigators have begun to focus their attention on nonintellective or personality characteristics as predictors. These factors, referring to intellectual "interests" or "dispositions," include those measured by personality tests and inventories, personal ratings, biographical information, and study-habit inventories. By using the nonintellective factors in addition to the intellective factors, however, little predictive efficiency has been added. A third major concentration of investigation has been that of the interaction between the student's personality and his social environment.

Traditionally, academic performance, at the college level, has been measured by the grades which the student receives in his courses. The grades are converted to a quantitative scale, usually from four to zero, and then averaged to obtain a grade point average (GPA), which then supposedly represents the student's academic performance. This GPA is widely used as a criterion measure for academic success. It thus becomes extremely important to the student as he pursues his academic career as well as to the college administrator, who must decide who is to be admitted to and allowed to remain in the university.

In discussing the problem of "what is success?," Fishman (1958) stated that the most common answer related to grades, "with good grades equivalent to good success and poor grades equivalent to poor success." The first problem arises with the mechanical quantification of grades. He described the practice of converting letter grades of A, B, C, D, and F into 5, 4, 3, 2, and 1 equal interval scales as a "makeshift affair." The second problem related to the unreliability of grades within departments and the variability of grading standards across departments. A solution to this problem would need to begin with the way the instructors in the various departments grade. "Do they have absolute standards or do they grade relative to the achievement of the group or even relative to the ability of the individual? Is their grading influenced by student abilities in areas unrelated to course content (such as personality or physical attractiveness)?" (Fishman, 1958, p. 341).

Most investigations concerned with the prediction of college academic success have selected the grade point average, generally that of the freshman year, as their criterion measure. However, the relationships between the predictors, usually intellectual factors, and the performance criterion, grade point average, are not strong, generally around +0.50.

Lavin (1965) pointed out that low correlations between predictors and performance criteria may be due to uncontrolled sources of variance in academic grades. Until these sources of variations are controlled, predictability of academic performance may not be improved to any great extent. He classified these sources into two categories. Since students major in different curricular areas, of varying degrees of difficulty, they will not enroll in the same courses. When analyzing total grade point averages, it must be remembered that some have been obtained in physics while others have been accumulated in physical education. He suggested the idea that comparing grade point averages across majors and academic disciplines may prove questionable and unprofitable; rather performance should be studied within curricular groupings.

Another source of variance is the use of different criteria in assigning grades by various teachers. Some instructors characteristically inflate or deflate the value of the grade. As a result, students' grades do not possess a sufficient degree of comparability. Instructors of the same subject exhibit great variability in their grading practices (Thompson, 1958). Morris (1953) found that there is great variability among the instructors and departments combined across years. Bass (1951) showed that there is great variability in grading practices among departments and subject matter areas. Different criteria are used in the assignment of grades by the various instructors within a single discipline as well as by the instructors in different disciplines. While many examinations are objective, some are subjective, involving the writing of essays, and students often perform differently on these two types of examinations. Again a differential factor is involved when major term papers are requirements and when oral presentations are made. These various types of assignments are weighted differently by different instructors. Therefore, it becomes difficult to compare accurately grades from one course to another or from one instructor to another.

Aiken found, after analyzing the grades given by twenty-two college departments for six semesters, the departmental grade ranks as "rather consistent across grading periods and between divisions" (Aiken, 1964, p. 828). He also found that departments with lower grade ranks tended to be those that had required courses and therefore enrolled more students. The previous year, Aiken (1963) had studied the grading practices at Women's College at the University of North Carolina. He concluded that this faculty did not base its grading on standards that were consistent over the years.

In a study concerned with college grading practices of faculty members in the areas of education, biological science, physical science, social science, and language and literature, Travers and Gronlund (1950) found that there was no single accepted concept as to the meaning of a mark. There were at least fourteen components found to determine grades, the greatest number of instructors employing the application of logical criticism as a criterion. This was followed by effort in the course, writing skill, and class discussion. Other factors considered were interest, attitude, promptness in turning in work, attendance, and attentiveness.

Part of the error variance involved in total grade point average results from the poor reliability of grades as a measure of performance across instructors and academic areas. The error variance in over-all grade point average, resulting from grading differences, tend to be self-correcting, and thus involve some compensation. It is for this very reason that total grade point average has been used in many studies of prediction of college achievement. However, the use of grade point average in the prediction of college achievement was challenged by Stern, Stein, and Bloom (1956) since in calculating an average, the grade point average being an arithmetic mean or average, high grades do compensate for low grades.

Theoretical Model

Hays (1963, p. 7) discussed what he calls the central problem of theoretical statistics--that is, "mathematical statistics is a theory about uncertainty, the tendency of outcomes to vary when repeated observations are made under identical conditions." Deductions can be made about the probability of various outcomes from theoretical statistics if the mathematical conditions are met. Mathematical statistics, a formal mathematical system, consists of three parts: undefined abstract "things" or "elements," undefined operations, and postulates and definitions, describing specific relations among the elements. It must be remembered that a mathematical system is abstract and does not deal with "anything" in particular. When the mathematical system is employed in an attempt to describe real objects or events, it becomes a mathematical model for those objects or events. The undefined, abstract properties of the mathematical system become identified with specific properties of objects and events. If the "formal" characteristics of the system parallel the "real" characteristics of the objects, "then anything that is a logical consequence in the system is a true statement about the objects in the model" (Hays, 1963, p. 9). It becomes very important that the mathematical conditions be met for valid results to be claimed. As Guilford (1965) pointed out there is never absolute isomorphism between mathematical ideas and phenomena of nature, but there is enough similarity so that the mathematical logic can be applied to nature. The Gaussian curve is such a mathematical model so that it is often used to make inferences about groups of natural events.

Statistics serves two purposes--that of description and that of inference. Descriptive statistics allows the investigator to organize and summarize the data in his distribution. The values used in descriptive statistics include the moments of the distribution which are the "expectations" of these values. Efficiency is generally attained only from the first two moments, the first moment being the mean, the "location" of the distribution, and the second moment, the variability (expressed as the standard deviation), the dispersion of the distribution. The third moment is skewness, which will be zero for a normal distribution, and the fourth moment is Kurtosis, or "peakedness."

Of the two, inferential statistics is more important as it allows for the prediction of events beyond the data in the immediate distribution. Small samples, or even one sample, may be taken from a larger population for the purpose of making inferences about that population. It is this sampling that serves as the basis for inference. This procedure is extremely important in applying the theoretical model to the academic performance of individuals since only a small subset of the student's academic performance is taken to predict his future academic achievement. The statistic, a value computed from a subset of the population, is merely an approximation of the parameter, which is the value of the population, the common parameters and statistics being the mean and variance of standard deviation.

In general, only the first two moments of the distribution are discussed as they appear to do a reasonable job. They are simple to compute, and the third and fourth moments require advanced mathematical techniques which operate against their practical applicability. Prior to the advent of computers, it was almost impossible, but now their computation is becoming more practical, and such a procedure should be explored.

The arithmetic average or the mean, the first moment of the distribution, serves as the most commonly used index of the central tendency of a distribution because it is simple to compute. Although the other moments of the distribution add to the description of the distribution, they do not do so significantly, and they are very difficult to calculate. It is a unique feature of the mean that the resulting mean deviation of the scores in the distribution will always be zero. It will be that score in the distribution in which there are equal deviations around it. If it is desirable for the average of the signed errors to be zero, the mean then becomes the best guess for any randomly drawn score from the distribution. This will be true in distributions other than the normal distribution, but it is extremely important in inferential statistics. An alternate description of the mean is that it is the "expectation" or "expected value" of a random variable. Since the mean is an arithmetic average, which is one of several types of averages, over an indefinite number of cases the mean is the average value for any random case of the distribution.

As important as the mean is to inferential statistics, it has a serious disadvantage: its sensitivity to changes at the extremes of the distribution. This may result in a skewed distribution. Changing only one extreme score in the distribution may greatly affect the mean. "The occurrence of even a few very high or very low cases can seriously distort the impression of the distribution given by the mean, provided that one mistakenly interprets the mean as the typical value" (Hays, 1963,

p. 175). An average does not tell anything about the spread or dispersion of scores of the distribution. It does not indicate the shape of the distribution. The average is computed "to see through the variability, so to speak, to the general location" (Wallis & Roberts, 1956, p. 244).

Dispersion, or variability, of the scores in a distribution indicates how the scores deviate from the central tendency. It tells how well or how poorly the measure of central tendency, in this case, the mean, describes the random scores, indicating how the observations tend to be, or "not" to be, like the average. The variance of a distribution is defined as "the average of the square of the deviations of the measurements about their mean" (Mendenhall, 1964, p. 30). The data will be more variable and thus exhibit a higher variance when the average of the squared deviations is large than when it is small.

A sample of independent observations of a variable will produce a sample mean that estimates the population mean and a sample variance that estimates the population variance. These values of a sample can be described as a joint event. Hays provided the following principle: "Given random and independent observations, the sample mean M and the sample variance . . . are independent if and only if the population distribution is normal" (Hays, 1963, p. 233). Sampling a normal population, the sample mean does not control the sample variance, and vice versa. If the population is not normal, these sample statistics, the mean and the variance, are not independent across samples. In making inferences from the variance, violating the normality assumption is not serious for a large N, thirty or greater, but inferences should not be made when a small or moderate N, less than thirty, is obtained. The size of the sample distribution will determine how small the variance will be. As the size of the distribution decreases, the variance increases.

Academic success is measured by the grades obtained. After the grades which the student receives in his courses are converted to a quantitative scale, they are averaged to obtain a grade point average. This is an arithmetic mean; it is one of several measures of central tendency. It is used frequently because it is simple to compute. The grade point average then typifies all the grades earned by the student and is used to describe the student's academic work.

However, the grade point average does not provide a complete description of the student's academic performance. There may be some extremely low grades obtained to offset an otherwise good academic record or a few extremely high grades intended to improve an otherwise mediocre academic record. It becomes important, therefore, to consider the variance of the grades, the second moment of the distribution, obtained in order to get a more complete picture of the student's performance.

A high grade point average would be expected to exhibit a low variance. As the grade point average approaches the maximum level, 4.00 on a four-point scale, the amount of variance exhibited among earned grades would become smaller until a variance of zero is attained. On the other hand, a low grade point average would be expected to exhibit a high variance. The student makes a series of attempts to maintain a minimum grade point average required by the college. In most cases, those individuals who possess a low grade point average and a low grade point variance have fallen below the minimum point and no longer attend college.

Although the above theoretical model is commonly applied, it is not reasonable to do so because of the restraints placed on the student as he progresses through his academic career. As the student selects the coursework he will follow, there are required courses and electives. Even when there is a choice involved in the selection, the student often elects a course with a specific purpose in mind--either interest or a desire to improve or maintain a grade point average. Therefore, the distribution scores of his academic performance are not random.

Hays stated that the use of sample statistics as "estimators" or population values, as in the case of the grade point average and variance, is a large problem in inferential statistics. A sample statistic value must contain "evidence about the value of the corresponding population values" (Hays, 1963, p. 196). Inferring from a sample value to a population value is called "point estimation" because the estimate is made up of a single value. The sample only represents a small subset of the larger population. Since chance effects will be operating, it is unwise to say that any estimate is exactly like the population value. Hays (1963) proposed four properties that a statistic should have in order to be a good estimator of a population parameter. They included:

1. The sample statistic should be an unbiased estimate.
2. The sample statistic should be a consistent estimate.
3. The sample statistic should be an efficient estimate.
4. The sample statistic should be a sufficient estimate.

<u>The sample statistic should be unbiased</u>. That is, the sample

statistic is unbiased as an estimate of the population parameter if the expectation of the sample statistic is the same as the population parameter. On the average, with all possible random samples, the sample statistic equals the population parameter. The sample mean is an unbiased estimate of the population mean; however, the sample variance is a biased estimate of the population variance. On the average, the sample variance underestimates the population variance when the sample size is small. However, they are almost equal for large samples. Also, large samples produce better estimators of the population mean than do small ones.

The sample statistic should be a consistent estimate. The larger the sample size the sample statistic will have a higher probability of being close to the population value. The sample mean and the sample variance are considered as consistent estimators since they tend to get closer to the population value as the sample size is increased.

A third property of the sample statistic should be the relative efficiency of the estimate. The standard deviation or standard error "represents the extent of the difference that chance factors tend to create between a sample estimate and a true parameter value" (Hays, 1963, p. 199). A good estimator will have small standard errors. Since the standard error indicates the tendency for the sample statistic to deviate by chance from the population value, the statistic with the smaller standard error will be more efficient. It would be expected that the grade point variance would be smaller within a specific discipline, and there would be better prediction for that discipline than to all disciplines combined, for both the high grade point average group and the low grade point average group. The grade point variance will be smaller for the grade point average within the major area of study than for the total grade point average. It will also be less than for the area of greatest number of hours. Therefore, the grade point average for the major area of concentration will have a smaller standard error and be a more efficient predictor. However, as the size of the sample decreases, as it does when moving from the area of total grade point average (TGPA) to the area of the grade point average in the major field (MGPA), the variance increases. This may obviate any decrease in variance in the major area due to specialized interest.

The fourth property is that of sufficiency. A sufficient estimator is one that contains "'all' the information available in the data about the value of θ" (Hays, 1963, p. 200). It is "best" in that referring to other aspects of the sample data does not improve the estimate. The sample mean is a sufficient estimator of the population mean, since it is not necessary to refer to any other statistic in estimating or interpreting the population mean.

Statement of the Problem

A different approach to the index of academic achievement at the college level other than total grade point average is needed because of the accumulation of error variance. One approach to the control of the sources of variance in grades, in an effort to increase the strength of the relationship between predictor variables and criterion measures, could be through the separation of courses into broad areas of concentration of study.

This study attempts to establish the grade point averages and variances of several broad areas of general education, major area, and area of greatest number of semester hours as substitute criteria for the total grade point average as a criterion of academic performance. It also seeks to explore the patterns of academic achievement of students, females and males combined and separately, in these various broad areas of academic work. A further attempt at increasing the strength of the relationship between the predictor variables and criterion measures will be made by adding the second moment of the distribution, through a multiple correlation technique, the variance of obtained grades, to the first moment of the distribution, the grade point average, the presently used criterion measure.

Hypotheses

In this study it is hypothesized that:

Research Hypothesis 1: The correlation coefficients which describe the relationship between the measures of potential performance (American College Tests Battery and Otis Quick-Scoring Mental Ability Test: Gamma Test) and the criterion measures, the grade point average and variance in the major area and the grade point average and variance in the area of greatest number of semester hours, will be significantly greater than the correlation coefficients between the measures of potential performance (American College Tests Battery and Otis Quick-Scoring Mental Ability Test: Gamma Test) and the criterion measures, total grade point average and total grade point variance.

$H_0$1A: Only chance differences will be found among the correlation coefficients which describe the relationship between the measures of potential performance (American College Tests Battery and Otis Quick-Scoring Mental Ability Test: Gamma Test) and the achievement (total grade point average, grade point average in major area, and grade point average in the area of greatest number of semester hours).

H_o1B: Only chance differences will be found among the correlation coefficients which describe the relationship between the measures of potential performance (American College Tests Battery and Otis Quick-Scoring Mental Ability Test: Gamma Test) and the measures of variance (total grade point variance, grade point variance in the major area, and grade point variance in the area of greatest number of semester hours).

H_o1C: Only chance differences will be found between the correlation coefficients for the measures of potential performance and criterion measures for females and the correlation coefficients for the measures of potential performance and criterion measures for males.

Research Hypothesis 2: The multiple correlation coefficients, utilizing the American College Test Composite Scores and the Otis Quick-Scoring Mental Ability Test with the criterion measures of grade point average and variance will be significantly greater than the correlation coefficients of either the American College Test Composite Score or the Otis Quick-Scoring Mental Ability Test with the criterion measures of grade point average and variance.

H_o2: Only chance differences will be found between the multiple correlation coefficients, utilizing American College Test Composite Score and the Otis Quick-Scoring Mental Ability Test with the criterion measures of grade point average and variance, for the combined group and the correlation coefficients of either the American College Test

Composite Score or the Otis with the criterion measures of grade point average and variance.

Research Hypothesis 3: The multiple correlation coefficients, utilizing grade point average and grade point variance with the American College Test Composite Score or the Otis Quick-Scoring Mental Ability Test, will be significantly greater than the correlation coefficients of either grade point average or grade point variance with the American College Test Composite Score or the Otis Quick-Scoring Mental Ability Test.

H_o3A: Only chance differences will be found between the multiple correlation coefficients, utilizing grade point average and grade point variance with the American College Test Composite Score or the Otis Quick-Scoring Mental Ability Test, for the combined group, and the correlation coefficients of either the grade point average or the grade point variance with the measures of American College Test Composite Score and the Otis.

H_o3B: Only chance differences will be found between the multiple correlation coefficients, utilizing grade point average and grade point variance with the American College Test Composite Score or the Otis Quick-Scoring Mental Ability Test, for females, and the correlation coefficients of either the grade point average or the grade point variance with the measures of American College Test Composite Score and the Otis.

H_o3C: Only chance differences will be found between the multiple correlation coefficients, utilizing grade point average and grade point variance with the American College Test Composite Score or the Otis

Quick-Scoring Mental Ability Test, for males, and the correlation coefficients of either the grade point average or the grade point variance with the measures of American College Test Composite Score and the Otis.

Significance of the Study

Because of its widespread use in the academic community, the grade point average as a criterion measure of academic success needs to be investigated. The grade point average and its efficacy as a criterion measure of academic performance comes to the forefront when it is used in counseling the student concerning his future course of academic work in college. Further, as a discriminator in the admission and retention of students, more efficient criterion measures of academic performance can lead to an improvement in the management and quality of the various college programs.

The over-all grade point average is not a highly efficient measure of academic performance. Investigations in which the over-all grade point average was used as the criterion measure of academic performance have only accounted for approximately twenty-five percent of the variance among grades, the currently used evaluation measure. Various batteries of scholastic and aptitude tests have been developed in an effort to increase predictive efficiency of academic performance, utilizing the over-all grade point average as the criterion measure for academic success. However, "more efficient prediction of college or university performance will depend as much or more on improvement of college evaluation of performance as on improvement of such prediction instruments as the ACT" (Engelhart, 1965, p. 7). Attention now needs to be focused upon the problem of eliminating the error components involved in student evaluation and upon identifying the various components in student academic performance as measured by teachers' grades.

CHAPTER II

REVIEW OF RELATED LITERATURE

The purpose of this chapter is to present a review of the studies that have been concerned with the prediction of college academic performance in relation to the type of predictor variables that have been employed and the statistical approach utilized. There are two types of predictor variables used: intellective and nonintellective. The factor of sex-linked differences in prediction is also presented. The grade point average as a criterion measure of college success is investigated.

Several approaches have been pursued in an effort to investigate and explain the academic achievement of college level students. These studies have generally focused on the problem of predicting the future academic performance of college-bound individuals. Intellective factors have received most of the attention in the studies that have been conducted, and, in most cases, the strength of the relationship between these intellective predictor variables and the criterion measure of success, usually total grade point average, was not strong--approximately +0.50. Such a correlation means that approximately seventy-five percent of the variance among grades is unexplained, which results in poor prediction of academic work. This is reportedly due to the fact that the proper independent or predictor variables have not been identified and measurement errors in the predictors and the criterion measures are large and unexplained.

There are many components involved in the estimation of college achievement. Goodstein, Crites, and Heilbrun (1963) estimated that college achievement, as measured by over-all grade point average, is composed of several well-established variance components: intellective, (35%); nonintellective, (15%); unknown, (40%); and error, (10%). Their estimates were based upon the relationships between measures of aptitude and achievement, the predictor variables, and the success criterion, grade point average. Both variables are described as being "continuous variables, unrestricted in range, and linearly related, as specified in the zero-order and multiple regression models" (Goodstein, Crites, & Heilbrun, 1963, p. 175). Fishman and Pasanella (1960) reviewed the literature concerned with the components of college achievement. The major component in predicting academic success is an intellective function.

GPA as a Criterion Measure

The student's grades, as represented by the grade point average, have served as the traditional criterion of academic achievement. Other criterion measures, such as graduation from college, intellectual curiosity, career success, verbal expression, have been used in some studies, but for the most part the grade point average has been the sole index of academic performance.

Few studies have utilized GPA in specific courses or GPA beyond the freshman year in college. The first term or first year GPA was used in most cases. There have been many arguments against the use of GPA as a criterion of success, but it is still the most widely used measure of college achievement.

Munger (1957) argued that the first-semester GPA was too limited in its approach as a measure of success. By using this criterion, it is not possible to predict if the student will graduate from college. It is preferable to gauge how long the student will stay in college or if he will graduate rather than the size of the first-semester GPA. Overall success is more important than success in the first semester.

The first-semester GPA is the best predictor of the subsequent semester GPA.

> While it must be admitted that first-semester GPA is more closely related to second semester GPA (r of about .80) than the standard college grade-predictors are related to first-semester GPA (multiple correlation of about .64), the amazing thing is that there is such a high relationship between a combination of data available before the first semester is begun and performance for the first semester; it should not be expected that the first-semester GPA be predicted as accurately as the second-semester GPA can be predicted from the first-semester GPA. Measurement specialists refer to this as the unreliability of college GPA (i.e., if college GPA cannot predict itself very well then it is unlikely that it can be predicted very accurately by anything else (Fricke, 1956-1957, 40-41).

Based on this statement, high school achievement and ability tests should not be used to predict beyond the first semester of college. He argued that low correlations have resulted because the GPA for the first, second, third, and fourth year in college has been used in conjunction with predictor variables acquired before the student entered college. Those individuals who have not remained in college beyond the first or second year are not considered in these studies so that "the relationship between college achievement and predictors of college achievement appears to be much lower than it actually is" (Fricke, 1956-57, p. 40).

Chansky (1964) discussed the statistical operations that the GPA permits. He asked the question of whether parametric statistics can be applied to grading. The normal distribution is an underlying assumption of the interval scale used in product moment correlations. After reviewing several studies, Chansky reached the conclusion that teacher's grades did not form a normal distribution, but a skewed one. Correcting for individual differences in grading among teachers accentuated this discrepancy. "Even if the students had at the beginning of a course been assigned to a class by some random procedure, instruction, a biasing treatment, destroys the assumption of randomness" (Chansky, 1964, p. 96). Therefore, such computations are not permissible. He recommended the use of ordinal scales for graded achievement in research; the GPA would be the median grade.

Holland and Richards (1965) felt that the college grading practices needed to be examined. College grades do not predict well for success after college. The primary purpose of a college education supposedly is preparation for life, but under the present system college mainly prepares the individual for further graduate education.

Intellective Factors in Prediction

The range for the correlation coefficients between intellective predictors and future academic work runs from +0.30 to +0.70 (Lavin, 1965). Some of these correlations are the results of investigators using different types of ability tests. While some use a standard intelligence test, others use tests that specifically predict academic performance, such as the American College Tests Battery (ACT), the

American Council on Education Psychological Examinations (ACE), or the Scholastic Aptitude Test (SAT). Another predictor of future academic performance is the measure of past academic performance. Quite often, high school performance is used to predict college performance. Lavin (1965) reminds the reader that factors other than ability enter into the high school academic record.

In reviewing the studies concerned with academic prediction since 1950, Schroeder and Sledge stated that "intellective factors found, in decreasing order of importance, were high school achievement (grade point average slightly superior to rank in class), subject matter test scores, and measures of mental ability" (Schroeder & Sledge, 1966, p. 97). Understandably, specific course grades in high school correlated more highly with similar college course grades than with total college grades. The differences between the correlations for all intellective factors was small (±0.10). Fricke (1956-57) reported consistently averaged correlations of 0.55 between high school achievement and college achievement and correlations of 0.45 between standardized tests of academic ability and college achievement. By combining the two factors, high school achievement and ability test scores, multiple correlations of 0.64 are obtained with college achievement.

Brice (1957), basing his results on the high school and college records of 244 students, achieved better prediction using teacher predictive rating of future success or failure than with the objective predictors of the Otis and ACE tests. The subjective predictor showed a significant correlation coefficient of 0.48 with freshman year GPA while the Otis only established a correlation of 0.12 with freshman GPA.

The high school senior year GPA had a significant correlation of 0.41 with freshman GPA. These results were more inaccurate for the predicted failures than for the predicted successful students. In another study (Lanigan, 1947), Pearsonian correlations of low magnitude between the Otis and six areas of academic work--English (0.29), social science (0.53), language (0.23), mathematics (0.24), science (0.53), and fine arts (0.38)--led the investigator to conclude that the use of the Otis as a single predictive measure for college success, except in the areas of science and social studies, would be unprofitable.

Holland and Astin (1962) concluded that the best predictors of college academic performance were good high school grades and high self-ratings of scholarship. The least effective predictor was scholastic aptitude tests. His subjects included, however, only high aptitude students. Their assessment occurred over intervals from one to four years.

Nonintellective Factors in Prediction

The nonintellective factors involved in college achievement consists of personality and study skill variables. Goodstein, Crites, and Heilbrun (1963) gave three bases for inferring the existence of these variables. Only part of the variance in college achievement is explained by ability variables and measurement errors. Secondly, by utilizing optimally-weighted measures of intellective factors and past academic performance, higher correlations are achieved with college GPA than with either predictor alone. "Finally, there is some empirical evidence which shows a slight correlation (Mdn $r = .22$) between personality variables, as assessed by the Rorschach, Minnesota Multiphasic

Personality Inventory (MMPI), and similar measures, and college achievement" (Goodstein, Crites, & Heilbrun, 1963, p. 176). Moderately high correlations (Mdn r = .47) were reported for study habits inventories and GPA. They concluded that about five to fifteen percent of the variance in college achievement is contributed by nonintellective factors.

In studies concerned with nonintellective factors, Schroeder and Sledge (1966) found that the major influence came from affective factors. Such affective factors include interests, motivations, attitudes, beliefs, values, and adjustments. Public high school graduates were equal or superior to private high school graduates in their college achievement, but no significant relationship was found between the size of high school and achievement. As a general rule, women were superior to men in achievement.

In their study of background factors relating to achievement (Schroeder & Sledge, 1966), these factors were most predictive of over-all grade point average and least predictive of Mathematics GPA. These factors accounted for 40.4 percent of over-all college grade point variance and 32.8 percent of pure science. "The overall college grade-point regression also exhibited the smallest standard error of estimate (.6325), followed by language (.6931), social science (.7372), pure science (.9053), mathematics (.9804), and technical regression (2.0440)" (Schroeder & Sledge, 1966, p. 101). They concluded that "non-academic predictors are quite unimportant except with reference to prediction of technical and social science achievement" (Schroeder & Sledge, 1966, p. 102).

Goodstein, Crites, and Heilbrun (1963) attempted to identify some of the nonintellective variance in college achievement. The data were analyzed for within-levels of ability and configural analyses of personality variables as they related to TGPA. No consistent patterns of personality characteristics relating to GPA were found. Following along the same line, Holland and Richards studied the correlations between academic achievement tests, average school grades, and extra-curricular achievement. Very low correlations were found (Mdn $r = .04$). These results suggested that "academic and nonacademic accomplishment are relatively independent dimensions of talent" (Holland & Richards, 1965, p. 165).

Sex-Linked Factors in Prediction

Another factor involved in obtaining such different correlations between predictors and academic performance criterion is the sex composition of the different studies that have been conducted. In some of the studies the statistics were computed for males and females together while in others the sample was comprised of all males or all females. In most studies that consider sex-linked differences in academic performance, there appears to be such a factor operating.

Jackson (1955) indicated that women manifested higher academic performance than men when obtained grades served as the criterion. Also, women perform more in accordance with their measured ability than the men. Jex (1966) stated that women were consistently more predictable than men. He also found that women received better grades in college as well as in high school than did men. In Abelson's study (1952) there was a sex-linked difference for predictability of academic performance;

females had a higher predictability. He stated that this was probably due to the fact that girls' college grades possessed greater homogeneity and had a smaller standard deviation than those of the boys. Using a differential prediction approach, Berdie (1955) reported higher correlations for females than for males.

Multiple Correlation Prediction

Studies in which a battery of predictors were used to predict college performance produced a somewhat higher correlation than for zero-order correlations. A ceiling is reached quickly in multiple correlation so it is not generally valuable to add more than three or four variables to the prediction. The multiple correlations reported by Cronbach (1949) ran around +0.60 to +0.70. Recent research indicated an average multiple correlation of about +0.65.

Fishman and Pasanella (1960) reviewed 216 studies in which multiple correlations were calculated using intellective predictors with the college freshman average. They reported correlation coefficients from 0.37 to 0.83, with a median coefficient of 0.62. For grades beyond the freshman year, eleven studies had multiple correlations ranging from 0.50 to 0.72 and a median correlation of 0.65. The usual predictor combination was an aptitude test score and the high school record. In twenty-one studies, using aptitude test score and high school record, the multiple correlation increased from 0.00 to 0.23 over the zero-order correlation based on high school average alone. A median increase of 0.07 was achieved. A similar rise of 0.00 to 0.38 was gained in 181 studies when any one of the intellective predictors was added to the high school record. However, when high ability students were used (Holland, 1958), the multiple correlation dropped to 0.22. He concluded that this decrease was the result of widely different grading practices and the restricted range of talent among the subjects.

Of the intellective variables used in the multiple correlation prediction studies, the high school average or high school rank appears to be the best single predictor. "Few studies came to the point of combining intellective and nonintellective predictors by means of multiple-correlation techniques. Where this was done, the gain in multiple correlation attributable to the nonintellective predictor was discouragingly small" (Fishman & Pasanella, 1960, p. 303). Holland (1962) reported the best multiple predictors of college achievement as good high school grades and high self-ratings of scholarship. However, he was working only with the high aptitude students. Scholastic aptitude tests were the least efficient predictors of college achievement.

Bloom and Peters (1959) devised a method for increasing the correlation between high school grades and college grades. By using a scaling method to obviate high school grading standard differences, they obtained an average multiple correlation of +0.75. Fishman (1962) concluded from the study by Bloom and Peters that the variance accounted for by the nonintellective factors could thereby be decreased. He further concluded that these factors, as approached by current research, appear to be similar to the usual predictors of academic performance. "Bloom's work also reinforces the earlier arguments of those who have pointed to needed refinements in the college criterion as the next major target for further predictive improvement" (Fishman & Pasanella, 1960, p. 302).

However, Lindquist (1963) attained results, using internal scaling of high school grades to predict college grades, that were unsatisfactory. He stated that this method was not a promising approach for improving the prediction of college grades.

Crawford and Burnham (1945) used standardized Navy achievement tests as criteria in the Navy V-12 Program and obtained multiple correlations of 0.74 for the usual academic predictors. It is this approach that Fishman recommended, and he stated that the multiple correlations "will be more representative of the true predictive power of our cognitive predictors" when this is done (Fishman, 1958, p. 349).

Differential Prediction

Another approach is that followed by Horst (1957) in which the differential prediction technique is used. Instead of predicting an over-all academic performance level, prediction is made for specific course areas. The predictor batteries employed depend upon the particular course area for which grades are being predicted. Thirteen measures are used in this technique. There are seven entrance tests taken the freshman year in college. High school grade point averages are calculated for six high school subject areas. The age and the sex of the student are also considered.

Using these variables, prediction can be made with "considerable accuracy in each of 32 different course areas at the University as well as the over-all University grade point average" (Horst, 1955-1956, p. 458). Following the idea that global predictors and criteria concentrate on over-all academic performance, multidimensional predictors and criteria should be used for assessing performance since the academic work of many students will differ from one subject to another. Differential prediction is not used for admission at the University of Washington but for guidance in future academic work. The student is able to find out his weakest and strongest areas through this technique.

Horst found median correlations of about +0.50, with a range of +0.13 to +0.89. "In general, it has turned out that of the students who were predicted to do very well, less than 10 per cent did poorly and of the students who were predicted to do poorly, less than 10 per cent did well" (Horst, 1955-1956, p. 462). The middle group could not be predicted as well, as is the case in many studies.

Differential prediction differs from global prediction primarily because the criterion is not an undifferentiated average as in the global prediction. The institutions which use this process are interested in "guided admission" rather than selection. "This model is also geared toward the educational philosophy which claims to be interested in the 'particular' areas of excellence or insufficiency of each candidate rather than in his 'average' excellence or insufficiency" (Fishman & Pasanella, 1960, p. 307).

Stone (1954) applied multiple regression equations in differential prediction of performance in four college curricula: commerce, elementary education, physical sciences, and social sciences. He used the high school GPA (HSGPA), the American Council in Education Psychological Examination (ACE), and the Cooperative General Culture Test (CGCT) as predictors to curriculum GPA. The most efficient single predictor was the high school GPA. For commerce and elementary education, using a

HSGPA and ACE total, multiple correlation coefficients of 0.63 and 0.73, respectively, were achieved. A battery of HSGPA, ACE total, CGCT Literature, and General Science test achieved a multiple correlation of 0.73 for the physical sciences.

Summary

From a review of the literature the use of intellective factors in the prediction of future academic performance appears to be the most promising approach. Correlations between intellective factors and over-all grade point average ranged from +0.30 to +0.70, the average correlation appearing around +0.50. Nonintellective factors, such as personality traits, interests, and study-habits consistently have produced low correlations with the over-all grade point average, leading many investigators to doubt their value in the prediction of future academic work. When considering these studies it is very important to determine whether the sex-linked factor had been considered. Throughout these studies, the females have been significantly more predictable than the males.

Through the use of multiple correlation techniques the correlation between predictor variables and criterion measures were increased somewhat. This correlation generally increased to about +0.65. The predictor variables for these studies included an aptitude test score and the high school record of the student. When nonintellective variables were combined with the intellective variables in a multiple correlation technique, the results were generally neglible.

It has been repeatedly brought out in the literature that the problems involved in prediction studies quite often rested with the criterion measures of academic performance rather than with the predictor variables. A systematic approach is needed which will investigate these criterion problems more fully. Supposedly, when this is done, predictive efficiency will be increased. Much of the error variance results from the differences in grading standards among instructors as well as departments; there is no single accepted concept concerning the meaning of grades.

CHAPTER III

METHODOLOGY

The main interest of this study was in the strength of the relationship between the predictor variables, Otis and ACT, and the criterion measures, grade point average and variance of college academic performance. A Pearsonian coefficient of correlation was computed between the predictor variables and each of the criterion measures of academic performance to determine the strength of their relationship. A series of multiple coefficients of correlation were calculated to further explore the relationships existing between the predictor variables and criterion measures. A factor analysis was then run in an effort to identify the factors upon which the twelve variables were loading. The data, in all cases, were analyzed with regard to three aspects: the combined group, the males, and the females.

Subjects

The subjects included in the study were comprised of the students enrolled in Education 120, Psychology of Education, at the University of Oklahoma for the spring semester, 1968-1969. Education 120 is a required course for all students seeking a teaching certificate and admitted to the teacher education program in the College of Education. It is generally the second course that these students take within the College of Education. In addition to the prerequisites for the course, Psychology 1 and Sociology 1, they will have been enrolled in Education 52; enrollment in Education 52 does not require admission to the College of Education.

Before the students are allowed to enroll in Education 120, they must be admitted to the College of Education which involves the attainment and maintenance of the following cumulative grade point average: 2.00 for 24-44 semester hours; 2.05 for 45-60 semester hours; 2.10 for 61-74 semester hours; 2.15 for 75-89 semester hours; 2.20 for 90-104 semester hours; and 2.25 for 105-124 semester hours. The extremely low ability student, or the student who is not performing at the anticipated level, and who has not been able to maintain the minimum grade point requirement very often is no longer attending the university.

Originally, there were 410 students included in the study. A certain number of students had to be eliminated from the sample because not all the necessary information on them was available. Only those who had taken the American College Tests Battery for admittance to the University of Oklahoma and had taken the Otis Quick-Scoring Mental Ability Test: Gamma Test were included. All graduate students were excluded from the sample since they were not required to take the ACT and since the focus of the study was on the prediction of undergraduate academic performance. Most of the transfer students were not required to take the ACT to be admitted to the University of Oklahoma and thus were eliminated from the sample. Those transfer students who had taken the ACT and had taken at least one semester of previous work at the

University of Oklahoma were kept in the sample. Because of grading differences and course differences between colleges, those students with previous work at other colleges, but not at the University of Oklahoma, were excluded. The Otis Quick-Scoring Mental Ability Test was administered to all students enrolled in Education 120 at the beginning of the semester by the instructors of the course. However, due to some absences not all students took the test. The final sample consisted of 222 students who had been admitted to the teacher certification program and on whom all scores used in the study had been secured.

Inferences can be made to similar groups, but not beyond such groups without much caution. Because many factors are operating in the admission of these students to the College of Education and Education 120 is a required course for a teaching certificate, this sample is not representative of the university as a whole. The student at the lower end of the continuum of academic performance is not represented in the sample. It can only be said to be representative of students who have been admitted to the College of Education.

Instruments

The American College Tests Battery is a widely used college entrance examination used for the purpose of selecting those individuals who will be able to succeed in the college curriculum. This Battery measures the student's educational development but also indirectly measures his intellectual ability. It has been described as a combination scholastic aptitude examination and achievement test battery (Downie, 1967). Anastasi (1968, p. 231) characterized the ACT as overlapping "traditional aptitude and achievement test, focusing on the basic intellectual skills required for satisfactory performance in college." By measuring the student's educational development in applying his knowledge to problem solving, it is possible to predict how successful he will be in college.

The test battery is made up of four tests, each with its own reliability coefficients, based on 900 high school seniors---English Usage, (0.90); Mathematics Usage, (0.89); Social Sciences Reading, (0.86); and Natural Sciences Reading, (0.83). The composite score of the ACT has a reliability coefficient of 0.95. Intercorrelations were computed with the data for the four tests. They were as follows: English and Mathematics, (0.53); English and Social Science, (0.63); English and Natural Science, (0.58); Mathematics and Social Science, (0.55); Natural Science and Mathematics, (0.64); and Social Science and Natural Science, (0.68). Because of these high intercorrelations, the ACT should not be used for differential predictions. Multiple correlation coefficients based on the 1962 data of 132 participating colleges were reported to indicate the predictive validity of the ACT. Using the ACT scores and high school grades to predict college performance, the following multiple correlation coefficients were obtained: 0.74, 0.67, and 0.58, for the 90th. percentile, 50th. percentile, and 10th. percentile, respectively (Engelhart, 1965). The five scores on the American College Tests Battery were obtained for each student from his file in the registrar's office.

The Otis Quick-Scoring Mental Ability Test is comprised of three tests. The Gamma Test, which was administered to the students enrolled in Education 120 by the instructors of the course, is designed for high school and college students and is a revision and extension of the Higher Examinations of the Otis Self-Administering Tests of Mental Ability. "The purpose of the three tests in the series is to measure mental ability---thinking power or the degree of maturity of the mind" (Otis, 1954, p. 1). While some of the answers to the questions depend upon schooling to some extent, the goal was to use that type of question as little as possible so as to depend as much on thinking as possible. Horrocks and Schoonover (1963) made the following statement: "Validities and reliabilities of the Otis Quick-scoring Mental Ability Tests are relatively satisfactory when compared with those cited for other measures of group intelligence, but unfortunately the test manual is quite vague as to the nature of the normative population and of the normative sample" (Horrocks & Schoonover, 1963, p. 347).

Area Groupings

One of the uncontrolled sources of variance in the measurement of academic performance, as reported in the literature, is the result of comparing grades obtained in various academic disciplines that are not, in fact, comparable. In an effort to control some of this variance, and thus increase the efficiency of prediction, it became necessary to classify the different aspects of the student's academic course work into comparable groupings. The broad fields of academic study were classified into eighteen academic areas for this study. These areas are presented in Table I with the curricular areas grouped under each area grouping. Grade point averages and variances for the 222 subjects were computed for these classifications, depending upon the focus of the student's work.

TABLE I

EIGHTEEN AREAS OF UNDERGRADUATE CURRICULAR AREAS AT THE UNIVERSITY OF OKLAHOMA

Area Groupings	Curricular Areas
Aviation	Aviation
Behavioral Sciences	Psychology; Social Work; Sociology; Anthropology
Business	Accounting; Business Administration; Business Communication; Business Law; Economics; Finance; Management; Marketing; Office Administration
Communication	Speech; Journalism; Library Science
Education	Education
English	English
Fine Arts	Fine Arts; Art; Drama; Dance; Music; Music Technique; Music Theory; Music Education
Physical Education	Physical Education
Home Economics	Home Economics
Languages	Classical Culture; Modern Languages
Law	Law
Medical Sciences	Pharmacy; Nursing; Physical Therapy
Military Sciences	Military Sciences
Natural Sciences	Astronomy; Mathematics; Physics; Geology and Geophysics; Meterology
Philosophy	Philosophy
Physical Sciences	Bacteriology; Botany; Microbiology; Physiology; Zoology; Biology; Chemistry
Social Sciences	History; History of Science; Political Science; Geography

Only those courses actually offered by each of the areas are included in the grouping. Prerequisites that are organized under another area's domain were not included for that area. For each student, a grade point average and grade point variance were calculated for the total number of hours, hours in the major field, and the area with the greatest number of semester hours.

Procedure

A copy of the official transcript of each student enrolled in Education 120 was obtained from the registrar's office. Using the curricular areas designed for the study as a guide, the total grade point average and variance (TGPA and TGPV, respectively), the grade point average and variance for the major concentration (MGPA and MGPV, respectively), and the grade point average and variance for the area with the greatest number of cumulative hours (HGPA and HGPV, respectively) were calculated for each student. The five ACT scores for each student were obtained from the registrar's office and the test score on the Otis Quick-Scoring Mental Ability Test: Gamma Test was secured.

Each student's major area was determined by consulting a data information sheet which had been completed at the beginning of the semester by each of the students enrolled in Education 120. A difficulty arose with respect to elementary education majors, who up until enrollment in Education 120 had not completed any coursework in their major area. Building upon the idea that they should perform better, because of aptitude and interest, in prerequisite courses designed for elementary education majors only, a group of courses was identified to represent their major area of concentration. These are the subject areas with which an elementary school teacher must be familiar in order to teach elementary school children, and they would be comparable to courses required in any major field in the secondary school program. The courses included in the elementary education major are Geography 61, Botany 4, Physics 4, Fine Arts 3 and 4, Math 70, Library Science 308, and Health, Physical Education, and Recreation 90. Those individuals majoring in special education presented a similar problem. Their major area was represented by the required courses of Physics 4 and Botany 4.

The area chosen for greatest concentration of work was simply that area in which was recorded the greatest number of semester hours accumulated. Often this area corresponded with the major area. When a tie occurred with the area with the greatest number of hours, a table of random numbers was consulted to determine the choice.

The computer program used provided calculations giving the means and standard deviations for the twelve variables, Pearsonian correlations and multiple correlations between the predictor variables and the criterion variables. A Varimax orthogonal rotated factor analysis was then run for the twelve variables to identify the factors operating. The computations were accomplished using the IBM 360/40 at the Merrick Computing Center of the University of Oklahoma.

These series of calculations were computed for the entire sample, both males and females combined, and then for males and females separately. The literature concerned with the prediction of academic performance indicated that a sex-linked factor was operating. Females were consistently more predictable with regard to their future academic performance than the males. Comparisons were made between the males and the females.

A comparison of the means and standard deviations of the twelve variables was made. It would be expected that these statistical values would differ significantly between variables. A further comparison of the ACT scores of the students in this sample with that of ACT scores of the University of Oklahoma 1962 freshman class and the ACT scores of the University College advisees of the College of Education, 1967-1968, was made.

As proposed in the theoretical model supporting the use of GPA as an index of academic performance, there must be independence between GPA and GPV. Intercorrelations, using a Pearsonian coefficient of linear correlation, between TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV were calculated to determine if significant relationships existed among these variables and to ascertain if, in fact, they were independent.

Intercorrelations, using the Pearsonian coefficient of linear correlation, were computed between the predictor variables (Otis and ACT---English Usage, Mathematics Usage, Social Sciences Reading, Natural Sciences Reading, and Composite ACT) and the criterion measures of academic performance (TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV) in an attempt to discover which variables exhibited the strongest relationship.

While a correlation coefficient describes the strength of the relationship existing between two variables, a regression equation is used to establish the degree of linear regression of variable Y (the dependent variable) on variable X (the independent variable). In one instance, that of correlation, there is no clear-cut delineation between the two variables as to which is the independent or predictor variables. Hays (1963, p. 503) pointed out that the correlation coefficient is a "'symmetric' measure of linear relationship," and as such the designation of the independent and dependent variable is not necessary since the measure of linear prediction is the same. However, this symmetry does not exist for the regression equation. For a regression, one variable is definitely assigned as the independent variable; it is not free to vary. In the correlation coefficient the variables can take on any observed value for an individual. The regression procedures apply to those studies in which the results of "experimental" treatments are predicted using a linear function rule, whereas, correlations are used for linear prediction of natural traits, as in this study. Natural traits are those that the individual brings to the study. He has acquired them previously. Furthermore, the regression applies to the population and the correlation to the sample.

When these techniques are used simply for a descriptive purpose, it is not necessary to meet any assumptions concerning the form of the distribution. The only requirement to be met is that a number of distinct cases has been obtained, each with two numerical scores, X and Y (Hays, 1963). When this is met, it becomes possible to calculate correlations and regressions. As in the case of this investigation, when the interest is merely in exploring the relationships between variables, a correlation approach suffices.

Multiple correlations, with Otis and the composite score of the ACT as the predictor variables and the proposed criteria for academic performance (TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV) as the criterion measures, were calculated to test the strength of the relationship between these combined variables with the criterion measures. An increased efficiency of prediction is expected through the use of multiple correlation techniques.

Intercorrelations were also computed among the Otis Quick-Scoring Mental Ability Test and the five scores of the American College Tests Battery to estimate the strength or significance of relationships among these predictor variables. Guilford stated that "in building a battery of tests to predict a criterion, test makers should try to maximize the validity of each test and to minimize the correlations between tests" (Guilford, 1965, p. 403). As the correlations between the independent variables increases, the multiple correlation between the independent and dependent variables will decrease. A group of highly correlated predictor variables are predicting the same factors. The less duplication involved in prediction the more efficient the predictor battery. Adding the second highly correlated variable would not add significantly to the prediction.

Coefficients of multiple correlation were calculated between GPA combined with GPV in each of the three major groupings and the Otis and the ACT (composite score), respectively, to estimate the strength of the relationship when the second moment of the distribution is considered along with the first moment of the distribution.

A final analysis was made by performing a Varimax orthogonal factor analysis to determine common factors which might emerge from the twelve-variable matrix. A factor analysis serves two purposes: (1) summarizes the information of large number of variables; and

(2) presents a psychological interpretation of the common factors. The factor analysis was utilized to identify the underlying structure of the student's academic performance in college. For an adequate description of the data, the number of factors involved must be taken into consideration. One way of looking at and interpreting this underlying structure is to execute a Varimax factor analysis. Thurstone devised the rules for simple structure:

> (1) Each row of the factor matrix should have at least one zero.
> (2) If there are "m" common factors, each column of the factor matrix should have at least "m" zeros.
> (3) For every pair of columns of the factor matrix there should be several variables whose entries vanish in one column but not in the other.
> (4) For every pair of columns of the factor matrix, a large proportion of the variables should have vanishing entries in both columns when there are four or more factors.
> (5) For every pair of columns of the factor matrix there should be only a small number of variables with non-vanishing entries in both columns (Cooley & Lohnes, 1962, p. 161).

CHAPTER IV

RESULTS AND DISCUSSION

Predictor Variables and Criterion Measures

Initially, the means and standard deviations were calculated for all variables used in the study for the combined group, the males, and the females. In Table II the means and standard deviations for the criterion measures, TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV for the combined group, are presented. It will be noticed that the magnitude of the mean grade point average in each category (over-all coursework, major, and area of greatest concentration) followed the expected pattern with respect to the fact that the largest mean was that of MGPA (2.87). This higher grade point average would be expected because of the student's supposed interest and aptitude in his chosen field. The next highest grade point average is that of HGPA (2.75). A test of significance, using a t test for the difference between correlated means (Guilford, 1965), between TGPA and MGPA was calculated since the largest difference existed between these two means. The results indicated no significant difference between the two means ($t = 0.33$, $p > .05$). No further tests were conducted since the difference between any of the other means would be smaller and therefore insignificant. Therefore, the students did not perform better in their major area as suggested. As expected, the variance

TABLE II

MEANS AND STANDARD DEVIATIONS OF THE CRITERION MEASURES OF TOTAL GRADE POINT AVERAGE AND VARIANCE, GRADE POINT AVERAGE AND VARIANCE IN MAJOR, AND GRADE POINT AVERAGE AND VARIANCE IN AREA WITH THE GREATEST NUMBER OF HOURS FOR THE COMBINED GROUP

	GPA	GPV	MGPA	MGPV	HGPA	HGPV
Mean	2.71	0.61	2.87	0.37	2.75	0.43
Standard Deviation	0.55	0.35	0.77	0.47	0.72	0.43

for the major area was smaller, although, again, not significantly ($t = 0.79$, $p > .05$), than the total grade point variance.

Criterion measures for males and females. In Table III the means and standard deviations of the criterion measures for males and females are presented. When the means and variances of the criterion measures for males were considered, the differences occurred in the same direction as that for the combined group, with the MGPA being higher (2.50) than the TGPA (2.42). A test of significance, using a test for the difference between correlated means, revealed no statistical difference between means ($t = .72$, $p > .05$). The same pattern appeared for the females (TGPA = 2.42; MGPA = 2.99) although no significant differences between means were found ($t = .39$, $p > .05$).

In comparing the GPA's of the males and the females, the females consistently earned a higher GPA--2.80 as opposed to 2.42 for TGPA and 2.99 as opposed to 2.50 for MGPA. To test the significance of the difference between these means, a t test for difference between independent means, with unequal N's and unequal variances (Hays, 1963) was computed. There was a significant difference between the males and females for TGPA and MGPA ($t = 4.81$, $p < .01$; $t = 4.22$, $p < .01$, respectively). This was the general pattern reported in the literature concerning the greater achievement of females over males. In addition, the females showed a significantly smaller GPV than the males in all categories. The following results between males and females were significant beyond the .01 level for TGPV ($t = 4.03$) and MGPV ($t = 2.87$). HGPV (2.05) was significant beyond the .05 level. It must be remembered that as N decreases, as it does in this study from 170 for females to 52 for males, the variance will increase.

TABLE III

MEANS AND STANDARD DEVIATIONS OF THE CRITERION MEASURES OF TOTAL GRADE POINT AVERAGE AND VARIANCE, GRADE POINT AVERAGE AND VARIANCE IN MAJOR, AND GRADE POINT AVERAGE AND VARIANCE IN AREA WITH THE GREATEST NUMBER OF HOURS FOR MALES AND FEMALES

	GPA	GPV	MGPA	MGPV	HGPA	HGPV
			Males			
Mean	2.42	0.80	2.50	0.56	2.53	0.56
Standard Deviation	0.48	0.42	0.72	0.59	0.68	0.56
			Females			
Mean	2.80	0.55	2.99	0.31	2.81	0.39
Standard Deviation	0.54	0.30	0.75	0.41	0.72	0.37

Predictor variables for the combined group. The means and standard deviations for the predictor variables, the Otis Quick-Scoring Mental Ability Test and the four tests of the American College Tests Battery, plus the composite score for the ACT, for the combined group, are presented in Table IV. Norms for the ACT for the University of Oklahoma 1962 freshman class are presented in Table V. Because the data was incomplete and the sampling procedure was unknown, statistical comparisons were not possible. The standard deviations for the norm group were computed on only a random half of each group. It should be noted in Table V, that in all categories, the ACT scores obtained by education majors produced means that were lower than those obtained by the university students combined. As noted in Table IV, the students in the present investigation obtained ACT scores whose means were all above those presented by the university students. The Counseling Center of the University of Oklahoma reported that by 1967 the average ACT composite score had increased by two points (Foster, 1967). Also, the standard deviations in the present study were smaller than those computed for the 1962 freshman class ACT scores.

In Table VI are presented the means and standard deviations of the ACT test scores obtained by University College advisees of the College of Education, 1967-1968. In comparing the results of the 1967-1968 University College advisees and the students in the present study, it was found that the students in the present study scored higher, on the average, than the 1967-1968 University College advisees.

The means and standard deviations of the Otis and the ACT scores for males and for females are presented in Table VII. When comparing

TABLE IV

MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR VARIABLES OF THE OTIS MENTAL ABILITIES TEST AND THE AMERICAN COLLEGE TESTS BATTERY FOR THE COMBINED GROUP

	OTIS	ACT				
	Gamma Test	English Usage	Mathematics Usage	Social Science Reading	Natural Science Reading	Composite Score
Mean	111.86	21.46	20.65	22.75	21.50	21.71
Standard Deviation	8.27	3.89	5.40	4.77	5.48	3.94

TABLE V

MEANS AND STANDARD DEVIATIONS FOR ACT TESTS FOR 1962 FRESHMEN: EDUCATION MAJORS AND UNIVERSITY STUDENTS

		ACT TESTS				
		English Usage	Mathematics Usage	Social Science Reading	Natural Science Reading	Composite Score
Education Majors	Mean	19.96	17.37	18.62	18.18	18.66
	Standard Deviation	4.5	5.9	5.5	5.7	4.5
Total University	Mean	20.26	20.55	20.97	20.97	20.83
	Standard Deviation	4.6	6.5	5.6	5.9	4.9

TABLE VI

MEANS AND STANDARD DEVIATIONS FOR ACT TESTS FOR 1967-1968 UNIVERSITY COLLEGE ADVISEES OF THE COLLEGE OF EDUCATION

	ACT TESTS				
	English Usage	Mathematics Usage	Social Science Reading	Natural Science Reading	Composite Score
Mean	20.16	18.72	20.37	19.20	19.92
Standard Deviation	3.73	5.23	5.31	4.78	4.21

TABLE VII

MEANS AND STANDARD DEVIATIONS OF THE PREDICTOR VARIABLES OF THE OTIS MENTAL ABILITIES TEST AND THE AMERICAN COLLEGE TESTS BATTERY FOR MALES AND FEMALES

	OTIS	ACT				
	Gamma Test	English Usage	Mathematics Usage	Social Science Reading	Natural Science Reading	Composite Score
			Males			
Mean	113.35	20.10	21.40	23.12	22.87	22.02
Standard Deviation	8.03	4.41	5.60	4.28	6.32	4.20
			Females			
Mean	111.40	21.88	20.42	22.64	21.09	21.62
Standard Deviation	8.31	3.63	5.34	4.91	5.15	3.86

the scores obtained on the Otis and the ACT composite, the males scored higher. Only on the ACT English Usage subtest did the females score higher than the males.

<u>GPA-GPV intercorrelations</u>. Intercorrelations among the variables concerned with measures of academic performance, grade point average and variance in the three categories, total course work, major field, and area of greatest concentration, for the combined group are presented in Table VIII. A consistent pattern developed between the grade point averages and grade point variances. A negative correlation was indicated, in each case, between the GPA and GPV for each of the three categories. In addition to being negative correlations, the Pearson coefficients were low in the case of MGPA-MGPV (-0.36) and HGPA-HGPV (-0.32). The correlation coefficient for TGPA-TGPV was -0.61; this was the strongest relationship for the three GPA-GPV pairings.

The same pattern of negative correlations between TGPA-TGPV, MGPA-MGPV, and HGPA-HGPV held for females and for males. These intercorrelations for the criterion measures for females and males are presented in Tables IX and X. As postulated in the theoretical model, the grade point average and the grade point variance must be independent. These correlations of low to moderate magnitude (-0.65, -0.33, and -0.30 for females in the three categories and -0.43, -0.32, and -0.31 for males in the three categories) indicated that such independence between the mean and variance did not exist. All of these correlations were significant beyond the .01 level of confidence (.01 = .181 for the combined group; .01 = .208 for the females; and .01 = .354 for the males) except the correlations for HGPA-HGPV and MGPA-MGPV for the males, which were significant beyond the .05 level (.05 = .273).

TABLE VIII

PEARSON CORRELATION COEFFICIENTS BETWEEN CRITERION VARIABLES FOR THE COMBINED GROUP

	TGPV	MGPA	MGPV	HGPA	HGPV
TGPA	-0.61	0.77	-0.33	0.80	-0.32
TGPV		-0.42	0.63	-0.46	0.67
MGPA			-0.36	0.68	-0.28
MGPV				-0.32	0.62
HGPA					-0.32

TABLE IX

PEARSON CORRELATION COEFFICIENTS BETWEEN CRITERION VARIABLES FOR FEMALES

	TGPV	MGPA	MGPV	HGPA	HGPV
TGPA	-0.65	0.76	-0.33	0.81	-0.35
TGPV		-0.42	0.54	-0.50	0.59
MGPA			-0.33	0.62	-0.28
MGPV				-0.30	0.45
HGPA					-0.30

TABLE X

PEARSON CORRELATION COEFFICIENTS BETWEEN
CRITERION VARIABLES FOR MALES

	TGPV	MGPA	MGPV	HGPA	HGPV
TGPA	-0.43	0.70	-0.17	0.73	-0.14
TGPV		-0.25	0.71	-0.32	0.77
MGPA			-0.32	0.83	-0.18
MGPV				-0.31	0.84
HGPA					-0.31

Although the distributions were skewed positively, the sample size (N = 222) for the combined group was sufficiently large to obviate this skewness as a factor affecting the independence of the two values, the mean and the variance of the scores. An N of 170 for females and an N of 52 for males would also be sufficiently large enough to ameliorate this skewness. Hays (1963) stated that non-normality is not a serious consideration when N is sufficiently large. As can be seen in Table IX and Table X, significant negative correlation coefficients were found consistently across the three categories between the GPA and the GPV, indicating that some factor was operating so that the means and variances were not independent as suggested by the theoretical model.

These persistent negative correlations between GPA and GPV can be explained because of the high correlations between TGPA and MGPA and HGPA. A negative correlation was exhibited between TGPA and TGPV. Because of the high positive correlation of MGPA (0.77 for the combined group; 0.76 for the females; and 0.70 for the males) and HGPA (0.80 for the combined group; 0.81 for the females; and 0.73 for the males) with TGPA, the same negative relationship would exist between those GPAs and their respective GPVs.

A higher correlation existed for the females for TGPA-TGPV (-0.65) than for males (-0.43). A test of significance, using Fisher's z transformation technique for uncorrelated data, was calculated for these two correlations. The results indicated no statistical difference ($z = 1.06$; $p > .05$). Since the greater difference between the larger coefficients was not significant, the smaller differences would not be significant either. The relationship between TGPA-TGPV was stronger than the relationships existing for MGPA-MGPV and for HGPA-HGPV for both males and females as well as the combined group. This would be expected for two reasons. When changing from the total area to the major field or the area with the greatest number of hours, (1) the variance increases and (2) the distribution becomes truncated.

Predictor Variables and Criterion Measures Intercorrelations

Research hypothesis 1 predicted that the Pearsonian coefficients of correlation between the predictor variables, Otis and the five scores of the ACT, and criterion measures, MGPA, MGPV, HGPA, and HGPV, would be significantly greater than the Pearsonian coefficients of correlation between the predictor variables, Otis and the five scores of the ACT, and criterion measures, TGPA and TGPV. To test this research hypothesis, three null hypotheses were formulated: (1) that of only chance differences between the correlation coefficients between MGPA and HGPA, respectively, and TGPA; (2) that of only chance differences between the correlation coefficients between MGPV and HGPV, respectively, and TGPV; and (3) that of only chance differences between the correlation coefficients for females and for males with respect to the predictor variables and the criterion measures.

Intercorrelations for the combined group. The predictor variables and the criterion measures were analyzed through the use of Pearsonian correlation coefficients to discover if any significant relationships could be found. The findings for the combined group, both males and females, are presented in Table XI. All of the predictor variables and grade point averages reported significant correlation coefficients beyond

TABLE XI

PEARSON CORRELATION COEFFICIENTS BETWEEN THE PREDICTOR VARIABLES (OTIS AND ACT) AND THE CRITERION VARIABLES (GRADE POINT AVERAGE AND VARIANCE) FOR THE COMBINED GROUP

	TGPA	TGPV	MGPA	MGPV	HGPA	HGPV
OTIS	0.44	-0.17	0.31	0.04	0.28	0.00
ACT_E	0.44	-0.22	0.30	-0.05	0.29	-0.01
ACT_M	0.41	-0.14	0.37	0.01	0.34	-0.02
ACT_{SS}	0.42	-0.23	0.24	-0.02	0.27	-0.03
ACT_{NS}	0.31	-0.10	0.21	0.08	0.19	0.01
ACT_{COM}	0.47	-0.21	0.33	0.01	0.33	-0.02

the .01 level of confidence. To be significant at the .01 level, a correlation coefficient of 0.181 must be reached; at the .05 level, a correlation coefficient of 0.138 must be attained. The strongest relationship existed between the composite score of the ACT and TGPA ($r = 0.47$). This was followed next by the Otis and the English Usage Test of the ACT; both showed a correlation coefficient of 0.44 with TGPA.

The correlation coefficients reported for MGPA were somewhat lower than those reported for TGPA. The strength of the relationship between MGPA and Mathematical Usage ($r = 0.37$) was the strongest of the six coefficients obtained for MGPA. The magnitude of the correlation coefficients for HGPA, the largest being 0.34 for ACT Mathematics Usage, was less than those reported for the predictors and TGPA. No tests of significance were computed since the direction of change from TGPA to MGPA and HGPA, respectively, was opposite to that hypothesized. The null hypothesis of only chance differences between the correlation coefficients of the predictor variables and the criterion measures for the combined group between MGPA and HGPA, respectively, and TGPA failed to be rejected.

The lack of increased predictive power from TGPA to MGPA and HGPA does not appear to be a function of N directly since there were an equal number of cases ($N = 222$) for both variables. However, Hays (1963) reported that the sample statistic should be an unbiased estimate of the population parameter, and that better estimators of the population mean are produced by large samples than by small samples. Indirectly, the sample size may have been related to the lack of the expected results. There were fewer semester hours accumulated in each case in the major area and the area of greatest concentration than for the over-all course load. By comparison with the over-all course load, this decrease in semester hours may have caused the mean to become biased as an estimate of MGPA and HGPA, resulting in a lower correlation between the predictor variables and MGPA and HGPA. TGPA was simply a better estimator of the average over-all course work than MGPA or HGPA were of their respective areas.

Negative correlations significant beyond the .01 level were found between TGPV and the predictor variables (ACT English = -0.22; ACT Social Science = -0.23; and ACT composite = -0.21). At the .05 level, significant negative correlations were found between TGPV and Otis (-0.17) and ACT Mathematics (-0.14). There were no significant correlations found between the predictors and MGPV or HGPV. The magnitude of the correlation coefficients decreased when moving from TGPV to MGPV and HGPV. Again, no tests of significance were applied as a result. Therefore, the null hypothesis of only chance differences between the correlation coefficients of the predictor variables and grade point variances in the three categories for the combined group failed to be rejected.

The relationship between GPA and GPV was not strong, as reported in Table VIII, and was moving in the wrong direction; GPV did not add significantly to the description of the distribution, for the combined group. All of the correlation coefficients between TGPV and the predictor variables were negative. Because GPA was positively correlated with the predictor variables, and GPA and GPV were negatively correlated, GPV for the major area and the area of greatest concentration must be negatively correlated with the predictor variables also.

Intercorrelations for males and for females. The Pearsonian correlation coefficients between the predictor variables, Otis and the five scores of the ACT, and the criterion variables, grade point averages and variances for the three categories, for females are presented in Table XII. All of the correlation coefficients between the predictor variables and GPA were significant beyond the .01 level. A correlation coefficient of .159 suffices at the .05 level of confidence. When using the ACT composite score as a predictor of TGPA for females, a correlation coefficient of 0.61 was obtained. This was an increase from the ACT Composite-TGPA correlation coefficient of 0.47 for the combined group. Using the Otis as a predictor of TGPA for females produced a correlation coefficient of 0.56.

For the females, negative correlations existed between the predictor variables and the variances of the three categories without exception. The highest of these was a -0.44 (significant beyond the .01 level) reported for ACT Social Science with TGPV. The next highest coefficient was for the ACT Composite of -0.41 (significant beyond the .01 level). The only significant correlation coefficient (beyond the .05 level) for MGPV was with the Otis (-0.17). The correlation coefficients for HGPV and Otis and ACT Composite, respectively, were reported at -0.20, (significant beyond the .05 level).

The Pearsonian coefficients between the predictor variables and the criterion measures for males are presented in Table XIII. None of the coefficients for GPA in the three categories were significant. To be significant beyond the .01 level a coefficient of .35 must be obtained. The general pattern of obtaining lower correlation coefficients for MGPA

TABLE XII

PEARSON CORRELATION COEFFICIENTS BETWEEN THE PREDICTOR VARIABLES (OTIS AND ACT) AND THE CRITERION VARIABLES (GRADE POINT AVERAGE AND VARIANCE) FOR FEMALES

	TGPA	TGPV	MGPA	MGPV	HGPA	HGPV
OTIS	0.56	-0.36	0.47	-0.17	0.39	-0.20
ACT_E	0.48	-0.31	0.35	-0.13	0.34	-0.13
ACT_M	0.52	-0.30	0.51	-0.12	0.44	-0.18
ACT_{SS}	0.53	-0.44	0.33	-0.14	0.34	-0.18
ACT_{NS}	0.48	-0.27	0.39	-0.03	0.31	-0.15
ACT_{COM}	0.61	-0.41	0.48	-0.12	0.43	-0.20

TABLE XIII

PEARSON CORRELATION COEFFICIENTS BETWEEN THE PREDICTOR VARIABLES (OTIS AND ACT) AND THE CRITERION VARIABLES (GRADE POINT AVERAGE AND VARIANCE) FOR MALES

	TGPA	TGPV	MGPA	MGPV	HGPA	HGPV
OTIS	0.24	0.13	-0.05	0.48	-0.00	0.40
ACT_E	0.19	0.08	0.01	0.22	0.08	0.29
ACT_M	0.20	0.09	0.07	0.25	0.10	0.26
ACT_{SS}	0.10	0.22	0.02	0.26	0.02	0.30
ACT_{NS}	0.07	0.09	-0.12	0.21	-0.03	0.25
ACT_{COM}	0.17	0.14	-0.02	0.28	0.05	0.33

and HGPA than those for TGPA was found for the males with the exception of ACT Natural Science and MGPA, which increased from 0.07 for TGPA to -0.12 for MGPA. However, neither of these correlation coefficients were significantly different from zero.

The males were more variable in the grades they obtained in their course work than the females. The males had a higher correlation between the predictor variables and MGPV and HGPV than they did in MGPA and HGPA. Just the opposite was true for the females where higher correlations existed for GPA than for GPV. A reflection of this greater variability of the males appears from the fact that they had a lower grade point average than the females although their ACT scores and Otis scores were higher than those for the females (see Table III and Table VII).

A 0.48 correlation coefficient was reported between Otis and MGPV for males; a 0.28 correlation coefficient was reported between ACT Composite and MGPV for males. Both of these were significant, the Otis-MGPV beyond the .01 level and the ACT Composite-MGPV beyond the .05 level. The correlation coefficient for Otis-HGPV (0.40) was also significant beyond the .01 level. The ACT Composite-HGPV (0.33) was significant beyond the .05 level.

The null hypothesis of no difference between the males and the females with regard to the correlation coefficients between the predictor variables and the criterion measures was rejected and, therefore, the research hypothesis of statistical differences accepted. Using Fisher's Z transformation for uncorrelated data, a test of significance was computed between the correlation coefficients of the predictor variables and GPAs for the females and males. The results of these tests are

are presented in Table XIV. With the exception of ACT English-HGPA and ACT Natural Science-MGPA, all the results were significant. As expected, the females had significantly high correlation coefficients between predictor variables and GPA for all three categories than the males. The females were more in line with their expected performance than the males.

The Otis and ACT Composite predicted variance for MGPV and HGPV better for the males than for the females. A test of significance, using Fisher's Z transformation for uncorrelated data, was calculated. The results were significant beyond the .01 level of confidence except for ACT Composite-MGPV (Z = 2.50) which was significant beyond the .05 level. Consideration of performance variability of the males is indicated from these results. From the correlations obtained in the study, it appears that the predictors ACT and Otis have a stronger relationship with male variability than they do with male average performance. Because of this stronger relationship with male variability, ACT and the Otis can possibly be profitable measures for admission to and guidance through college if male variability is considered as a criterion measure of academic success.

Multiple Correlation--Otis and ACT

Research hypothesis 2 predicted that the multiple correlation coefficients for the Otis and ACT with the criterion measures would be significantly greater than the zero-order correlations for the Otis and ACT, respectively, with the criterion measures. To test this hypothesis, a null hypothesis of only chance differences between the multiple correlation coefficients for the predictor variables and the criterion measures and the zero-order correlation coefficients for the predictor variables and the criterion measures for the combined group was formulated.

TABLE XIV

SIGNIFICANCE (FISHER'S Z TRANSFORMATION) FOR PEARSON CORRELATION COEFFICIENTS FOR PREDICTOR VARIABLES AND CRITERION MEASURES (GPA) BETWEEN MALES AND FEMALES

	TGPA	Level of Significance	MGPA	Level of Significance	HGPA	Level of Significance
OTIS	2.45	.05	3.41	.01	2.52	.05
ACT_E	2.09	.05	2.17	.05	1.66	n.s.
ACT_M	2.28	.05	3.05	.01	2.27	.05
ACT_{SS}	3.07	.01	2.00	.05	2.04	.05
ACT_{NS}	2.78	.01	1.81	n.s.	2.14	.05
ACT_{COM}	3.29	.01	3.34	.01	2.54	.05

Multiple correlation coefficients were calculated, using the ACT Composite since it was the largest coefficient for the five ACT scores (see Table XI) and the Otis as the predictor variables, to establish the relationship between these two variables combined and the six criteria, as reported in Table XV, for the combined group. The null hypothesis of only chance differences between the multiple correlation coefficients and the zero-order correlation coefficients for the combined group failed to be rejected. The largest multiple correlation coefficient was found for TGPA at 0.55, significant at the .01 level (.01 = .212; .05 = .172). The zero-order correlation between Otis and TGPA was 0.44, and for ACT Composite and TGPA it was 0.47. However, in spite of the strength of the relationship for the multiple correlation coefficient, for practical purposes the large standard error (0.47) negates the significance of the size of the coefficient for prediction purposes. The multiple correlations reported for MGPA (0.45) and HGPA (0.42) were significant at the .01 level, but, again, their standard error, 0.70 and 0.77, respectively, are too large for accurate prediction.

Throughout the literature an increase in prediction was found when a battery of tests was used to predict future academic performance. Although a 0.55 correlation coefficient was achieved for TGPA, this means that only approximately thirty percent of the variance is accounted for through such procedures. Part of this may have been due to the relationship between the predictor variables, Otis and the ACT Composite. The intercorrelation between these two variables was 0.74. Because of this high intercorrelation between the predictor variables little was added to the prediction to the six criteria. Both tests appear to be measuring the same thing.

TABLE XV

MULTIPLE CORRELATIONS OF ACT COMPOSITE SCORE AND OTIS QUICK-SCORING MENTAL ABILITIES TEST AND GRADE POINT AVERAGE AND VARIANCE FOR THE COMBINED GROUP

	PREDICTORS OTIS ACT_{COM}	S.E.
TGPA	0.55	0.47
TGPV	0.29	0.34
MGPA	0.45	0.70
MGPV	0.18	0.47
HGPA	0.42	0.66
HGPV	0.08	0.43

The multiple correlations found for the variances were extremely low. The only one showing significance (0.01 level) was TGPV (0.29). The standard error (0.34) was very high also.

Multiple Correlation--GPA and GPV

Research hypothesis 3 predicted that the multiple correlation coefficients for GPA-GPV with the criterion measures would be significantly greater than the zero-order correlations for the Otis and ACT, respectively, with the criterion measures. To test this hypothesis, three null hypotheses, for the combined group, the males, and the females, of only chance differences between multiple correlation coefficients, utilizing GPA-GPV and Otis and ACT, and the zero-order correlation coefficient between Otis and ACT, respectively, and the criterion measures was formulated.

Multiple correlations (GPA-GPV) for the combined group. In Table XVI are presented the multiple correlation coefficients between GPA combined with GPV within the three major groupings and the Otis and ACT Composite for the combined group. As proposed in the theoretical model an increase in the strength of the relationship between the GPA and predictor variables would occur when the GPV, the second moment of the distribution, was added to the first moment, using a multiple correlation technique. The strongest relationship existed for the ACT Composite and TGPA-TGPV (0.49), but there was not a significant increase in the strength of the relationship when compared with ACT Composite-TGPA (0.47). The null hypothesis of only chance differences between the multiple correlation coefficients of GPA-GPV and the zero-order

TABLE XVI

MULTIPLE CORRELATION COEFFICIENTS OF GRADE POINT AVERAGE AND GRADE POINT VARIANCE AND OTIS MENTAL ABILITIES TEST AND ACT WITHIN GROUPINGS FOR THE COMBINED GROUP

	OTIS	ACT
TGPA - TGPV	.460	.485
MGPA - MGPV	.443	.149
HGPA - HGPV	.298	.286

correlation coefficients of Otis and ACT for the combined group failed to be rejected. This may have been the result of the intercorrelation between TGPA and TGPV for the combined group which was -0.61. As in the previous case with Otis and ACT, a high intercorrelation between variables indicates that they are measuring essentially the same thing; thus, the addition of one to the other in a prediction measure does not add significantly to that prediction.

Multiple correlation (GPA-GPV) for males and for females. In Table XVII the multiple correlation coefficients between GPA combined with GPV for the Otis and ACT Composite for the three major groupings for males and for females are presented. For the females there were no increases from the zero-order correlations, using GPA, to the multiple correlations considering GPV in addition to GPA. Therefore, no tests of significance were computed. The null hypothesis of only chance differences between the multiple correlation coefficients, utilizing GPA-GPV and Otis and ACT, respectively, and the zero-order correlation coefficients for GPA and the predictor variables for females failed to be rejected.

However, for the males there were several increases in the multiple correlation coefficients, using GPA and GPV, from the zero-order correlation coefficients using GPA. To be significant beyond the .05 level and .01 level of confidence, a multiple coefficient of correlation must reach a level of .336 and .410, respectively. In Table XVII, it will be noted that a multiple coefficient of correlation of 0.35 (significant beyond the .05 level) was obtained for the males for TGPA-TGPV with the Otis. This compared with a zero-order correlation coefficient

TABLE XVII

MULTIPLE CORRELATION COEFFICIENTS OF GPA AND GPV AND THE OTIS AND ACT WITHIN GROUPINGS FOR MALES AND FOR FEMALES

	OTIS		ACT	
	Male	Female	Male	Female
TGPA -TGPV	0.35	0.56	0.30	0.61
MGPA -MGPV	0.49	0.47	0.29	0.48
HGPA -HGPV	0.42	0.40	0.39	0.44

of 0.24 for the Otis and TGPA. As a test of significance, first-order partial correlations were calculated and evaluated. Ezekiel (1941, p. 214) defined the coefficient of partial correlation as "a measure of the extent to which that part of the variation in the dependent variable which was 'not' explained by the other independent factors can be explained by the addition of the new factor." The importance of the additional variable is evaluated by ascertaining how much of the variation is attributed by one variable when all the other variables are held constant. A partial correlation is a Pearson product-moment correlation and thus is comparable to zero-order correlations. By testing for significance of the coefficient of correlation of the additional variable, a conclusion could be drawn on the potency of that variable; if it were significantly greater than zero, it was making a real contribution. The criterion for significance for the coefficient of partial correlation for males was found to be .258 (.05 level) and .339 (.01 level) (Guilford, 1965, p. 341). For the males for TGPA-TGPV with the Otis compared with the zero-order correlation coefficient for Otis-TGPA, a partial coefficient of correlation of .266 (significant beyond the .05 level) was obtained. A partial correlation coefficient of .239 was reached for TGPV with ACT Composite-TGPA. This did not reach significance. Therefore, GPV did not add significantly to the relationship of ACT Composite-TGPA for males.

When the relationships between Otis and MGPA and Otis and HGPA were determined, a nonsignificant correlation coefficient was obtained (-0.05 and -0.00, respectively). The addition of a significant coefficient of correlation (Otis-MGPV, 0.48; Otis-HGPV, 0.40) to the relationship between Otis and GPA would increase that relationship to a significant level, resulting in a significant increase in the relationship. Such was the case for Otis and MGPA-MGPV with a coefficient of multiple correlation of 0.49, significant beyond the .01 level. A significant increase also occurred for the Otis and HGPA-HGPV with a multiple correlation coefficient of 0.42, significant beyond the .01 level. A significant increase occurred for the ACT and HGPA-HGPV (0.39) from the ACT Composite-HGPA (0.05). There was no significant increase for the ACT Composite-MGPA-MGPV (0.29) from the ACT Composite-MGPA (-0.02). The addition of a nonsignificant coefficient of correlation to a nonsignificant coefficient did not raise the level of the multiple coefficient of correlation to a significant level. There is no significant statistical difference between two nonsignificant coefficients of correlation.

Adding grade point variance to the mean or grade point average for males resulted in stronger relationships because the Otis and ACT exhibited a stronger relationship for the males with GPV than with GPA. Stronger relationships existed between the Otis and GPA-GPV than the ACT Composite and GPA-GPV because the Otis had a stronger relationship with GPV than ACT Composite.

The null hypothesis of only chance differences between the multiple correlation coefficients, using GPA-GPV, and the zero-order coefficients using Otis and ACT, respectively with GPA, for males, was rejected and the research hypothesis accepted.

Factor Analysis

In order to explore the underlying structure of the twelve variables used in the study, a final analysis of the data was made using varimax rotation of a factor analysis. In Table XVIII are presented the results of the factor analysis of the data for the combined group. From the table it can be seen that three reasonably clear factors emerged from the rotation. Factor A, in the first column, relates to student ability as measured by test scores, Otis and the four subtests and the composite of the ACT. It should be noticed that the mean or grade point average is loading negatively on Factor B, in column 2, and that the variance is loading negatively on Factor C, in column 3. The negative loading is a statistical artifact which results from the mechanics of rotation; it should not be interpreted meaningful in terms of direction or magnitude.

Factor B is the variability of student performance as measured by teacher grades. Factor C is the central tendency of performance of the student as measured by teacher grades. When describing student performance in college work, the variability of the student's performance must be considered along with his average performance. This will yield a more complete description of the student's academic performance.

Thurstone's criteria (Cooley & Lohnes, 1962, p. 161) are useful in interpreting the results of a varimax rotation. The following observations were made. In relation to Rule 1, "each row of the factor matrix should have at least one zero," there were near-zero loadings for at least one of the factors, except for TGPA, the lowest being -0.31 for Factor B, in each row of the matrix. Rule 2 stated that for "m" factors, in this case, three, there should be "m" zeros in each column. Rule 3 "for every pair of columns of the factor matrix there should be several variables whose entries vanish ~~in~~ one column but not in the other," was

TABLE XVIII

ROTATED FACTOR MATRIX FOR THE COMBINED GROUP

Variable	Factor A	Factor B	Factor C	h^2
TGPA	0.335	-0.310	<u>0.817</u>	0.877
TGPV	-0.149	<u>0.822</u>	-0.322	0.802
MGPA	0.162	-0.213	<u>0.849</u>	0.793
MGPV	0.080	<u>0.833</u>	-0.187	0.736
HGPA	0.154	-0.243	<u>0.847</u>	0.800
HGPV	0.025	<u>0.869</u>	-0.112	0.769
OTIS	<u>0.769</u>	0.060	0.275	0.671
ACT_E	<u>0.784</u>	<u>-0.078</u>	0.174	0.651
ACT_M	<u>0.624</u>	0.116	0.408	0.569
ACT_{SS}	<u>0.835</u>	-0.120	0.068	0.716
ACT_{NS}	<u>0.869</u>	0.014	0.026	0.756
ACT_{COM}	<u>0.970</u>	-0.015	0.196	0.979
	4.181	2.364	2.573	Total = 9.118
Percentage	45.85%	25.93%	28.22%	

consistently met for each pair of columns, the entries for one variable vanish in one column but not in the other. For TGPA, there was a .34 and a -.31 in column 1 and 2, respectively, and a .82 in column 3. Rule 4, "for every pair of columns of the factor matrix, a large proportion of the variables should have vanishing entries in both columns when there are four or more factors," did not apply in this case because only three common factors were extracted. Rule 5, which stated that a small number of variables with nonvanishing entries in both columns should exist, was complied with. Again, TGPA does not comply with this rule, with .34, .31, and .82 reported for the three columns. The only other variable, ACT Mathematics that does not meet this requirement, reported .62, .12, and .41 for the three columns. Thus, the factor analysis does not completely meet the criteria for simple structure set forth by Thurstone, but the pattern is a reasonable approximation of simple structure.

The factor analysis for the females is presented in Table XIX and the factor analysis for the males in Table XX. It will be noticed that when the three factor analyses for the combined group, females, and males are compared, they are almost identical. All three factor analyses reasonably approximate the rules set forth by Thurstone.

TABLE XIX

ROTATED FACTOR MATRIX FOR FEMALES

Variable	Factor A	Factor B	Factor C	h^2
TGPA	0.380	-0.339	-0.785	0.876
TGPV	-0.264	0.780	0.314	0.776
MGPA	0.226	-0.201	-0.820	0.764
MGPV	0.020	0.783	0.177	0.645
HGPA	0.165	-0.259	-0.826	0.776
HGPV	-0.092	0.813	0.095	0.679
OTIS	0.753	-0.106	-0.347	0.698
ACT_E	0.823	-0.077	-0.185	0.717
ACT_M	0.533	0.020	-0.558	0.596
ACT_{SS}	0.841	-0.214	-0.105	0.764
ACT_{NS}	0.841	-0.036	-0.168	0.736
ACT_{COM}	0.934	-0.093	-0.308	0.977
	4.115	2.178	2.710	Total = 9.004
Percentage	45.71%	24.19%	30.10%	

TABLE XX

ROTATED FACTOR MATRIX FOR MALES

Variable	Factor A	Factor B	Factor C	h^2
TGPA	0.191	<u>0.869</u>	-0.147	0.813
TGPV	0.023	-0.249	<u>0.857</u>	0.797
MGPA	-0.054	<u>0.919</u>	-0.106	0.859
MGPV	0.216	-0.155	<u>0.892</u>	0.866
HGPA	0.032	<u>0.904</u>	-0.205	0.860
HGPV	0.245	-0.100	<u>0.908</u>	0.894
OTIS	<u>0.696</u>	0.128	0.338	0.615
ACT_E	<u>0.805</u>	0.056	0.061	0.654
ACT_M	<u>0.710</u>	0.164	0.166	0.559
ACT_{SS}	<u>0.761</u>	-0.003	0.136	0.597
ACT_{NS}	<u>0.907</u>	-0.113	-0.005	0.835
ACT_{COM}	<u>0.985</u>	0.019	0.097	0.980
	4.156	2.573	2.603	Total = 9.332
Percentage	44.53%	27.58%	27.89%	

CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Summary

The purpose of this study was to investigate the grade point average, a common criterion of academic success, as an index to academic achievement at the college level. Previous investigators have pointed to the fact that the over-all grade point average, as employed presently by most schools, does not serve as an adequate criterion measure of college academic performance. Too many uncontrolled error variances and unknown components enter the picture when using the over-all grade point average as the representative score, or primary criterion, of the student's academic performance.

It was proposed, following a widely applied theoretical model, that the grade point average, a measure of central tendency, and the grade point variance, a measure of variability, should serve as a more complete index to student academic performance. For this model to be applicable the mean, the grade point average, and the variability, the grade point variance, must be independent. Otherwise, the addition of GPV to GPA would not add significantly to the relationship between predictor variables and criterion measures.

It was further hypothesized that the grade point average in the major field of interest and the area of greatest concentration of semester hours would serve as better indicators of the student's work. By developing a classification of course area or disciplines, the subject areas in which the student enrolls were made more comparable. One of the problems involved in the use of the over-all grade point average is that of comparability across discipline lines. Students will perform better in some areas, supposedly in his major field of interest, than he will in others. When assessing his academic performance, a more accurate description can be obtained by restricting attention to those areas most applicable to his future work and ignoring those ancillary courses which have no real bearing on his future work. The grade point average and grade point variance in the major field of interest and in the area of the greatest number of hours as criterion measures of academic performance were considered.

The usual predictors of college academic success include some measure of ability or aptitude, such as the ACT or SAT, and the student's high school academic record, or some measure of intelligence, such as the Otis or other intelligence test. For the purposes of this study, the ACT and the Otis were chosen as the predictors of academic performance. They were considered separately and then together in a multiple correlation design.

There were 222 students, all enrolled in Education 120 for the spring semester, 1968-1969, included in the study. Scores on the ACT and the Otis Mental Ability Test were secured for each of them and their college transcripts analyzed for estimates of TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV. A series of Pearsonian coefficients of correlation were computed between each of these criterion measures and the predictor variables, Otis and the ACT, in an effort to determine the strength of the relationship that existed between them. To further explore the relationship existing between the predictor variables and the criterion measures, multiple coefficients of correlation were computed, first, using the Otis and ACT, and then using GPA and GPV in the three categories. Finally, three factors were extracted through a factor analysis of the data.

Conclusions

From the results of this study, the following conclusions were reached:

1. The intercorrelations computed between the predictor variables, Otis and the ACT, and the criterion measures, TGPA, TGPV, MGPA, MGPV, HGPA, and HGPV indicated that the grade point averages and variances in the major area and the area of greatest number of semester hours did not exhibit a stronger relationship with the predictor variables than the grade point average and variance in the total course work. No tests of significance were computed since the change in magnitude occurred in the opposite direction of that hypothesized.

2. When comparing the results of the coefficients of correlation for the predictor variables and the criterion measures between the males and females, using Fisher's Z transformation for uncorrelated data, the females exhibited significantly stronger relationships than the males with regard to grade point average for the three categories. Grade point variance and the Otis and ACT Composite (MGPV and HGPV) exhibited a stronger relationship for the males than the females.

3. No significant increase for the combined group occurred from the zero-order correlations of the predictor variables and criterion measures to the multiple correlations, involving the Otis and ACT Composite with the criterion measures. Extremely large standard errors occurred for the multiple correlations so that they could not be used effectively.

4. No significant increases occurred for the combined group from the zero-order correlations between the predictor variables and criterion measures to the multiple correlations, utilizing GPV with GPA. There were no significant increases from the zero-order correlations and the multiple correlations, using GPA and GPV, for females.

5. The addition of GPV to GPA in the multiple correlation for males resulted in a significant increase in the relationship between predictor variables and criterion measures from the zero-order correlations for the predictor variables and criterion measures for the males. Even though this significant increase occurred for males, it did not surpass the zero-order correlations found for females.

Implications

It has been shown, through the factor analysis, that three relatively clear factors--that of ability as measured by test scores, achievement as measured by teachers' grades, and variability of student performance as measured by teachers' grades--are operating within the structure of student academic performance. Student ability as reflected by test scores, such as the ACT or Otis, has been employed quite often in an effort to predict future academic performance as reflected by grades. As has been indicated by the factor analysis, this procedure provides an incomplete description of the student's work. Student performance variability occupies an important place in that description.

Many investigators have become greatly concerned, not with the predictor variables as such, but with the criterion measures of academic performance--teachers' grades, as represented by the grade point average. It is true that if there were one score to be chosen as representative of the student's performance, the grade point average would be best for that score, on the average. However, that "best guess" fluctuates, more or less, depending upon the extreme grades that the student might receive in some of his courses. For this reason, the over-all grade point average has been highly criticized.

It has been suggested that the theoretical model proposed earlier be applied to the description of student academic performance in college. It has further been proposed that not only the first moment of the distribution, the grade point average, be used to describe the student's work, but also the second moment of the distribution, the variability of his work. In order for this to be done, however, the mean, the grade point average, and the variability, the grade point variance, must be independent. Such has been shown not to be the case--at least, when considering the combined group and the females in this sample. A high intercorrelation existed between TGPA-TGPV for the combined group. No increase in prediction occurred for females by adding GPV to GPA. Lower intercorrelations existed for GPA-GPV for males than for females. This may explain why the increase in prediction for males, using GPA-GPV, while no increase occurred for females. By separating the male and female distributions for analysis, the variability in the males' grades could be allowed to give a more complete description of male academic performance. Female academic performance is less variable than that of male performance and therefore variability does not add to the description of their academic ability.

Better prediction from the Otis and the ACT would result, it was hypothesized in this study, by moving from TGPA as a criterion measure of academic performance to MGPA or HGPA. The results of this investigation indicated that this, in fact, does not happen. One of the outstanding patterns throughout all the Pearsonian correlation coefficients computed was the persistent decrease in the relationship between the predictor variables and the criterion measures as the change from TGPA to MGPA and HGPA occurred. This may have been a function of N decreasing in each case. As N decreases, the statistic mean becomes less efficient as an unbiased estimate of the population mean. Fewer hours were accumulated in the major field than in the student's entire course work. It could have been that the ancillary courses were not extraneous to the student's program. When they were removed from the classification of academic areas, decreases in the relationships between variables may have occurred. When the grade point averages and grade point variances were compared in the three categories, no statistical differences could be found, further suggesting that significantly stronger relationships would not exist between the predictor variables and TGPA, MGPA, or HGPA and TGPV, MGPV, or HGPV.

Significantly stronger relationships did exist between the predictor variables and the criterion measures for the females than for the males even when GPV provided an increase for males in the relationship between predictor variables and the criterion measures. Although the males scored higher on the predictor variables than the females, the females possessed higher criterion scores and were more predictable with reference to their future academic work. The use of performance variability can serve as an additional useful index to the academic performance of males.

Recommendations for Further Research

There is a need for further exploration of the criterion measures of academic performance at the college level and their underlying assumptions. A criterion measure other than grade point average needs to be considered because of the lack of efficiency in explaining the variability in academic performance.

The classifications of academic areas should be re-examined in the light of efficiency in terms of the teacher certification program and possibly recategorized into different areas of concentration. Ancillary courses, which were removed before the analyses were made, might have contributed to the relationship between the variables and therefore should be considered.

As the subjects in the study were representative only of those students who had been admitted to the teacher certification program, a similar investigation should be conducted using a sample representative of other colleges and inter-college majors within the university.

Measures can be obtained for different ability level students. The question of whether better prediction for the various levels will occur by adding a measure of variability needs to be investigated.

There is some relationship existing between the student's grade point average and grade point variance. The significant negative correlations indicate that as the grade point average increases, the grade point variance decreases. The patterns of variability among high ability and low achieving students needs to be further analyzed in relation to the student's patterns of academic functioning. The significant negative correlations between the measure of central tendency and measure of variance indicates that they are not independent. This may be the result of the student's manipulation in an effort to gain control of his academic program. It may be caused by administrative policy or by a reaction of the faculty toward what they consider to be an unsatisfactory evaluation process. This lack of independence between GPA and GPV needs to be investigated in view of these possible causes.

The variability of male academic performance in relation to the predictive instruments utilized in the study, provided the groundwork for further investigation into the sex-linked differences involved in academic performance. The patterns of performance for males as compared to those for females as they pursue their academic career should be further explored. The relationship between the variability of performance and the central tendency of performance should be further investigated for males and females.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abelson, R. P. Sex differences in predictability of college grades. Educational and Psychological Measurement, 1952, 12, 638-644.

Aiken, Lewis R., Jr. Interdepartmental variability and student expectations of college grades. Educational and Psychological Measurement, 1964, 24, 823-829.

Aiken, Lewis R., Jr. The grading behavior of a college faculty. Educational and Psychological Measurement, 1963, 23, 319-322.

Anastasi, Anne. Psychological testing. (3d ed.) New York: Macmillan, 1968.

Bass, B. M. Intrauniversity variations in grading practices. Journal of Educational Psychology, 1951, 62, 366-368.

Berdie, R. R. Attitude, achievement, interest, and personality tests: A longitudinal comparison. Journal of Applied Psychology, 1955, 39, 103-114.

Bloom, B. S., and Peters, F. R. The use of academic prediction scales for counselling and selecting college students (Mimeo). (Rev. ed.). Chicago: University of Chicago, Office of the Examiner, 1959.

Brice, Marshall M. A comparison of subjective predictions with objective predictions of college achievement. College and University, 1957, 32, 347-353.

Chansky, Norman M. A note on the grade point average in research. Educational and Psychological Measurement, 24 (no. 1), 1964, 95-99.

Cooley, William W. and Lohnes, Paul R. Multivariate procedures for the behavioral sciences. New York: John Wiley and Sons, Inc., 1962.

Crawford, A. B. and Burnham, P. S. Education testing in the Navy V-12 Program at Yale. Psychological Bulletin, 1945, 42, 301-309.

Cronbach, Lee J. Essentials of psychological testing. New York: Harper and Brothers, 1949.

Downie, N. M. Fundamentals of measurement: Techniques and practices. (2nd ed.) New York: Oxford University Press, 1967.

Engelhart, Max D. Review of forms 4-AC. In Oscar K. Buros (Ed.), The Sixth Mental Measurements Yearbook. Highland Park, N. J.: Gryphon Press, 1965, 2-7.

Ezekial, Mordecai. Methods of correlation analysis. New York: John Wiley and Sons, 1941.

Fishman, Joshua A. The use of quantitative techniques to predict college success. Admissions Information: The Preparation and Use of Admissions Data, 1957, 1, 49-61.

Fishman, Joshua A. Unsolved criterion problems in the selection of college students. Harvard Educational Review, Jan. - Dec., 1958, 28, 340-349.

Fishman, J. A., and Pasanella, A. K. College admission-selection studies. Review of Education Research, 1960, 30, 298-310.

Foster, Dorothy. Memo to Dr. Gerald Kowitz, May 8, 1967.

Fricke, Benno G. Prediction, selection, mortality, and quality control. College and University, 1956-1957, 32, 34-52.

Goodstein, L. D., Crites, J. O., and Heilbrun, A. B. Personality correlates of academic adjustment. Psychological Reports, 1963, 12, 175-196.

Guilford, J. P. Fundamental statistics in psychology and education. (4th ed.) New York: McGraw-Hill Book Company, 1965.

Hays, William L. Statistics for psychologists. New York: Holt, Rinehart, and Winston, 1963.

Holland, John L. Prediction of scholastic success for a high aptitude sample. School and Society, 1958, 86, 290-293.

Holland, John L., and Astin, Alexander W. The prediction of the academic, artistic, scientific, and social achievement of undergraduates of superior scholastic aptitude. Journal of Educational Psychology, 1962, 53, 132-143.

Holland, John L., and Richards, James M., Jr. Academic and nonacademic accomplishment: Correlated or uncorrelated? Journal of Educational Psychology, 1965, 56, 165-174.

Horrocks, John E., and Schoonover, Thelma I. Measurement for teachers. Columbus, Ohio: Charles E. Merrill Publishing Co., 1968.

Horst, P. Differential prediction in college admissions. College Board Review, 1957, 33, 19-23.

Horst, P. The differential prediction of success in various college course areas. College and University, 1955-56, 31, 456-471.

Jackson, R. A. Prediction of the academic success of college freshmen. Journal of Educational Psychology, 1955, 46, 296-301.

Jex, Frank B. Predicting academic success beyond high school. Salt Lake City, Utah: By the Author, 1966.

Lanigan, Mary A. The effectiveness of the Otis, the A. C. E. and the Minnesota speed of reading tests for predicting success in college. Journal of Educational Research, 1947, 41, 289-296.

Lavin, D. E. The prediction of academic performance. New York: Russell Sage Foundation, 1965.

Lindquist, E. L. An evaluation of a technique for scaling high school grades to improve prediction of college success. Educational and Psychological Measurement, 1963, 33, 623-646.

Mendenhall, William. Introduction to statistics. Belmont, California: Wadsworth Publishing Company, 1964.

Morris, V. C. High grades and low grades. College and University, 1953, 28, 317-329.

Munger, Paul L. Can we really predict who will graduate from college? College and University, 1957, 32, 218-221.

Otis, A. S. Otis Quick Scoring Mental Ability Tests: New edition. Manual of directions. New York: Harcourt, Brace and World, 1954.

Schroeder, Wayne L., and Sledge, George W. Factors related to collegiate academic success. Journal of College Student Personnel, 1966, 7, 97-104.

Stern, G. G., Stein, M. I., and Bloom, B. S. Methods in personality assessment: Human behavior in complex social situations. Glencoe, Illinois: Free Press, 1956.

Stone, Joics B. Differential prediction of academic success at Brigham Young University. The Journal of Applied Psychology, 1954, 38, 109-110.

Thompson, W. N. A study of the grading practices of 31 instructors in freshman English. Journal of Educational Research, 1958, 64, 65-68.

Travers, R. M. W., and Gronlund, M. E. The meaning of marks. Journal of Higher Education, 1950, 21, 369-374.

Wallis, W. Allen, and Roberts, Harry V. Statistics, a new approach. Brooklyn, New York: Free Press of Glencoe, 1956.