# PREDICTIVE MODELING OF THE ACADEMIC PERFORMANCE OF USAF ACADEMY PREPARATORY SCHOOL GRADUATES AT THE USAF ACADEMY

Kenneth R. Anderson

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

December 1974

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>GSA/SM/74D-1 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER<br>AD/A-003 598 |
| 4. TITLE (and Subtitle)<br>PREDICTIVE MODELING OF THE ACADEMIC PERFORMANCE OF USAF ACADEMY PREPARATORY SCHOOL GRADUATES AT THE USAF ACADEMY | | 5. TYPE OF REPORT & PERIOD COVERED<br>MS Thesis<br>6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Kenneth R. Anderson<br>Captain, USAF | | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>AIR FORCE INSTITUTE OF TECHNOLOGY (AFIT-ENS)<br>WPAFB, OHIO 45433 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>AIR FORCE INSTITUTE OF TECHNOLOGY (AFIT-ENS)<br>WPAFB, OHIO 45433 | | 12. REPORT DATE<br>December 1974<br>13. NUMBER OF PAGES<br>85 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release; IAW AFR 190-17

JERRY C. HIX, Capt, USAF
Director of Information

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

COLLEGE ENTRANCE EXAMINATION BOARD TESTS
GRADE POINT AVERAGE
PREDICTIVE MODELING
PREP SCHOOL
REGRESSION
USAF ACADEMY PREP SCHOOL

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

SEE OVER



DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20

This study used linear models developed by stepwise multiple regression to examine the relationship between performance at the USAF Academy Preparatory School and subsequent grade point average (GPA) at the USAF Academy. Data collected on Prep School graduates in the Academy classes of 71 through 77 who had completed at least their first year at the Academy provided the following as potentially significant quantitative factors for predicting Academy GPA: Prep School grades recorded as overall GPA, academic GPA, separate grade averages by subject, percentage grade averages for math and English using a grading method proposed by the Prep School; College Entrance Examination Board Aptitude and Achievement scores in math and English taken prior to and during the Prep School program; Prior Academic Record (high school); Physical Aptitude Exam Score; and Athletic and Nonathletic Activity Indices (high school). Academy freshman year GPA was used as the dependent variable for most of the regression runs and the highest $R^2$ value was .613. Conclusions included the following: given Prep School grades as predictor factors, the other quantitative factors do not significantly contribute to predicting Academy GPA; Prep School overall GPA predicts Academy GPA as well as academic GPA or separate grade averages; last quarter Prep School grades increase the predictive ability of Prep School grades; the proposed Prep School percentage grading method results in slightly higher ability to predict Academy GPA than does the conventional grading method.

# PREDICTIVE MODELING OF THE ACADEMIC PERFORMANCE OF USAF ACADEMY PREPARATORY SCHOOL GRADUATES AT THE USAF ACADEMY

THESIS

GSA/SM/74D-1

Kenneth R. Anderson
Captain USAF



Approved for public release; distribution unlimited

# PREDICTIVE MODELING OF THE ACADEMIC PERFORMANCE OF USAF ACADEMY PREPARATORY SCHOOL GRADUATES AT THE USAF ACADEMY

THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology

Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

by

Kenneth R. Anderson
Captain USAF

Graduate Systems Analysis
December 1974



ic

## Preface

This thesis is the result of my efforts to examine the relationship between performance at the USAF Academy Preparatory School and subsequent grade point average at the USAF Academy. This study was done for the Prep School and it is hoped that the Prep School will find the results useful in making policy decisions in the specific areas examined by this study.

Many people contributed to this thesis effort. I would like to express my sincere gratitude to Dr. Joe Cain, my advisor, for his guidance and assistance. I am indebted to Major Jerome A. Atkins, Mathematics Department Head, USAFA Preparatory School, for requesting this study and assisting in collecting the data. I am also indebted to Mr. Risdon J. Westen, Director of Evaluations, USAF Academy, for his counsel and willingness to provide data necessary for this study. I also wish to acknowledge the help of Mrs. Betty Roth, Statistical Assistant in Evaluations, USAF Academy, for taking time from a pressing schedule to manually extract and record a great deal of required data. Special thanks goes to Mrs. Dorothy H. Campbell, Mathematics Department secretary, USAFA Preparatory School, for her extensive assistance in extracting data from Prep School historical

records. Without her help the data could not have been collected in the time available.

Finally, I wish to convey my sincere thanks and appreciation to my family--Linda, Gwen and Cindy. They have supported me with their love and prayers while enduring the hardships and neglect of this trying period.

I accept sole responsibility for any errors contained in this thesis.

Kenneth R. Anderson

## Contents

## List of Tables

## Abstract

This study used linear models developed by stepwise multiple regression to examine the relationship between performance at the USAF Academy Preparatory School and subsequent grade point average (GPA) at the USAF Academy. Data collected on Prep School graduates in the Academy classes of 71 through 77 who had completed at least their first year at the Academy provided the following as potentially significant quantitative factors for predicting Academy GPA: Prep School grades recorded as overall GPA, academic GPA, separate grade averages by subject, percentage grade averages for math and English using a grading method proposed by the Prep School; College Entrance Examination Board Aptitude and Achievement scores in math and English taken prior to and during the Prep School program; Prior Academic Record (high school); Physical Aptitude Exam Score; and Althetic and Nonathletic Activity Indices (high school). Academy freshman year GPA was used as the dependent variable for most of the regression runs and the highest $R^2$ value was .613. Conclusions included the following: given Prep School grades as predictor factors, the other quantitative factors do not significantly contribute to predicting Academy GPA; Prep School overall GPA predicts Academy GPA as well as academic GPA or separate grade averages; last quarter

Prep School grades increase the predictive ability of Prep School grades; the proposed Prep School percentage grading method results in slightly higher ability to predict Academy GPA than does the conventional grading method.

# CHAPTER I

# INTRODUCTION

## Background

The 1972-1973 *History of the United States Air Force Academy Preparatory School* states,

> The mission of the USAF Academy Prep School is to prepare and motivate selected Regular and Reserve enlisted personnel for the USAF Academy and to eliminate candidates obviously lacking the academic potential or military aptitude for the Academy [Ref 10:1].

The Prep School admits about 230 students each year with the USAF Academy Director of Admissions selecting Regular airmen from active duty enlisted applicants and selecting Reservists from Academy candidates who have not received appointments to the Academy. The young men accepting selection as Reservists join the Air Force Reserve, are called to active duty, and receive basic training prior to reporting to the Prep School. During the nine-month Prep School program each student receives military training, physical conditioning, and instruction in English and mathematics. All courses are designed to teach knowledge and skills necessary to do successful work at the Academy as well as to qualify students for entrance to the Academy (Ref 10:1).

Qualification for entrance to the Air Force Academy is based on scores obtained on the English and mathematics sections (Aptitude and Achievement) of the College Entrance

Examination Board tests (CEEB) or on scores obtained on the American College Testing Program examinations (ACT); Prior Academic Record (PAR) which is a high school academic performance measure based on rank in class and size of the class;* score on the Academy's Physical Aptitude Exam; and on two high school activity indexes, one for participation in athletic activities and the other for nonathletic activities. Most of the scores used for qualification to the Academy fit on the 200 to 800 scale common to the CEEB tests. In addition to these numerical scores, qualification to the Academy is subject to some fairly stringent medical standards and to subjective evaluation such as ratings and comments of school authorities and liaison officers. For a thorough description of Academy cadet selection criteria and procedures the interested reader is referred to the current Air Force Academy catalog.

The same qualification standards apply to Prep School graduates and to non-Prep School graduates alike with one major exception. When a Prep School graduate's final GPA after being multiplied by 200 is higher than his Prior Academic Record computed by the Academy before he came to the Prep School, the scaled Prep School GAP (GPA times 200) is used in place of his PAR. While other Academy

*Using the size of the class is an effort to quality adjust the measure of prior academic performance. Support for the contention that the size of the school is one of the most important determinants of the quality of schooling is given in a dissertation study by Welch (Ref 11).

applicants can use either CEEB scores or ACT scores, Prep School students use only CEEB scores and in an effort to maximize these scores they usually take the CEEB tests three times during the Prep School year.

The Prep School nine month program is divided into three academic quarters. Every quarter each student concurrently takes two math courses, two English courses, a physical training course, and a military training course. A conventional four point letter grading system is used with a 4.00 representing an A. The relative weights of the courses in determining GPA are as follows: twenty percent for each of the four academic courses, twelve percent for the physical training course and eight percent for the military training course.

Some Prep School administrators have proposed a percentage grading system for their academic courses in an effort to make a Prep School student's grades more indicative of his actual performance at the Prep School. The present grading method involves assigning conventional letter grades for each course based on raw percentage averages earned in that course. Using certain criteria, grade cut lines are established for each course that divide the raw percentage measurement scale into intervals and all the raw percentage scores in a given interval are assigned the same letter grade. The criteria used to determine the

grade cut lines result in different cut lines for different courses.

The proposed percentage grading system would, in addition to assigning a conventional letter grade, assign a percentage grade for each course based on the raw percentage score earned in that course. Percentage grades in the 90s would correspond to the letter grade A, in the 80s to a B, in the 70s to a C, in the 60s to a D, and below 60 to an F. A percentage grade would also indicate the relative position that the raw percentage score fell between two letter grade cut line scores. For example, a percentage grade of 85 percent indicates that the raw percentage score fell half way between the lowest B cut line and the lowest A cut line. Thus, the percentage grades for a given course are determined from raw percentage order of merits by using a linear transformation which normalizes the cut lines used to assign letter grades to the desired 90, 80, 70, and 60.

Since Prep School testing and grading is very objective and very standardized within each course, and because the course raw percentage scores are typically based on a large number of possible points earned on several quizzes and tests, a percentage grade can be a better indicator of a student's performance in a course than a letter grade which, for example, fails to differentiate between a low B and a high B.

## Objective

Various unpublished studies have been done by the Air Force Academy on the performance of Prep School graduates at the Academy, relative to the performance of the student body as a whole. It is not the intent of this study to duplicate nor extend those studies. This study is being done for the USAFA Prep School to investigate the extent to which Prep School performance predicts subsequent Academy grade point average. Specifically the following "objective questions" were investigated:

1. Among the available and potentially available measures of high school and Prep School performance, what are the significant ones in terms of predictive modeling of Air Force Academy grade point average for Air Force Academy Prep School graduates?

2. What are the relative importances of the significant predictor factors?

3. Do the significant predictor factors remain the same when predicting Academy cumulative GPA at the end of a given class's first year at the Academy, second year, third year, and fourth year?

4. Is Prep School academic grade point average a significantly better predictor factor than Prep School cumulative GPA which includes military training and physical training grades?

5. Are separate English and math grade averages significantly better than academic cumulative GPA as factors affecting Academic GPA?

6. Does the proposed Prep School percentage grading system for academic courses result in significantly better predictor factors than the standard letter grading system?

7. Do the Prep School grades earned during the last third of the Prep School year contribute significantly to the predictive capability of Prep School grades?

8. Are the College Entrance Examination Board scores earned during the Prep School year significantly better as predictor factors than the CEEB scores earned prior to beginning the Prep School program?

# CHAPTER II

## METHODOLOGY

In this chapter the method used to investigate how various explanatory (independent) variables affect Academy GPA is discussed. The variables used as potentially significant explanatory factors and the data collected are described in the next chapter.

### Simple Correlation

The simplest measure of the relationship between a single dependent variable and a single independent variable is the zero-order (simple) correlation coefficient. However, simple correlation coefficients do not allow for control of the affects of other variables. A simple correlation coefficient between a dependent variable and an independent variable can be partly due to the affect of one or more other variables that are correlated with both the dependent variable and the independent variable. Only when the explanatory variables are independent of each other will individual simple correlations with the dependent variable be free of hidden effects from other explanatory variables. The primary explanatory variables used in this study were significantly correlated with each other, ruling out independence.

## Least-Squares Multiple Regression

In order to allow for the effect of several explanatory variables on the one dependent variable, Academy GPA, the statistical technique of least-squares multiple regression analysis was used. This technique fits a linear combination of independent variables to the data and results in a prediction equation. The model hypothesized to represent the relationship between the dependent variable and the independent variables is of the form

$$Y = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad (1)$$

where

Y is the dependent variable,

$x_1, x_2, \ldots, x_k$ is a set of independent variables,

$\beta_1, \beta_2, \ldots, \beta_k$ are unknown parameters,

and $\varepsilon$ is a random error.

The standard assumptions enabling inferences to be drawn from multiple regression analysis have been made in this study. That is, it is assumed that the random error, $\varepsilon$, is a normally and independently distributed random variable with constant variance and expected value zero.

The least-squares multiple regression technique estimates the $\beta$ parameters using the criterion of minimizing the sum of the squared differences between the values predicted by the regression equation and the actual values of the dependent variable. Under the assumptions about the error term these estimates are unbiased and can be tested

for statistical significance. For a given regression equation the squared multiple correlation coefficient, $R^2$, measures the proportion of total variation about the mean, $\bar{Y}$, explained by the regression equation and is thus an indication of how well the data fits the regression line.

## Stepwise Multiple Regression

The writer used the stepwise regression procedure as the basic tool to obtain a "best" regression equation for a given battery of potentially significant independent variables. The actual computer program used was the Biomedical Computer Program for stepwise regression, BMD02R. The stepwise routine calculates a series of regression equations arriving at one that by certain criteria is the "best" regression equation. The basic procedure is described in the next paragraph. For a comprehensive description of stepwise multiple regression see Reference 2.

The first independent variable to be entered by the stepwise regression routine is the one with the largest zero-order correlation with the dependent variable. A least-squares regression equation is calculated using this one independent variable. Then partial correlation coefficients of the remaining independent variables with the dependent variable (after allowance for the independent variable that entered in the first step) are calculated. The variable with the highest partial correlation becomes the next variable to enter and a new regression equation is

calculated. This procedure is repeated with the variable entered at each step being the one with the highest partial correlation with the independent variable (after allowance for the independent variables already in the regression equation). The routine is terminated when the partial F value (equivalent to the t-ratio) for the coefficient of the variable to be entered next is less than a preselected percentage point of the appropriate F distribution. There is one additional and important operation in the stepwise routine. At each stage the coefficients of the variables incorporated into the model in previous stages are re-examined for statistical significance. A variable which may have entered significantly at an early stage may become superflous at a later stage because of the relationships between it and other variables now in the regression equation. When at any stage the partial F value for the coefficient of a variable in the regression is less than a preselected value, that variable is removed and a new regression equation is calculated.

In this study the number of observations in each regression run was large enough that the critical F point was 3.84 (at the ninety-five percent confidence level) for all the regression runs. However, the writer set the minimum F level for inclusion into the regression model at 2.0 in order to decrease the chance of stopping too soon. Although the usual stepwise pattern had decreasing partial

F values for succeeding variable entries, occasionally the first variable to enter at an insignificant level was followed by a variable which entered at a significant level and the new partial F values for coefficients of variables already in the regression all remained or became significant. By setting the F for inclusion at a value less than the critical F of 3.84, variables were occasionally determined to be significant predictor factors that otherwise would have been overlooked.

## Multiple Regression Analysis

While the actual analysis on the data will follow in later chapters, the statistical measures used will be described here.

A least-squares multiple regression equation has several measures associated with it that were used in this study. The following were available directly from the computer printout: the least-squares estimated coefficients; the standard error of each estimated coefficient; partial F values; $R^2$, the square of the multiple correlation coefficient; and s, the standard error of the estimate.

In order to answer the objective questions it was also necessary to use some measure of the relative importance of individual explanatory variables in a given regression equation. The absolute magnitude of the estimated coefficients is inadequate even when the independent variables are all scaled to a common scale because of the

lack of independence between the independent variables. Only when the independent variables are themselves independent of each other can an individual coefficient be correctly interpreted as the marginal contribution to the dependent variable of a unit change in the independent variable (with the other independent variables held constant). In this study, as is usually the case with real world variables, the explanatory variables were not independent of each other.

While there is no completely satisfactory method to measure the relative importance of individual explanatory variables, this writer used the method of comparing incremental contributions to $R^2$. If a given regression equation contains K explanatory variables and $R^2$ is the squared multiple correlation for the regression, then the incremental contribution of the $h^{th}$ explanatory variable is given by $R^2 - R_h^2$ where $R_h^2$ is the squared multiple correlation for a regression on the K-1 variables obtained by deleting the $h^{th}$ variable. Thus, for each of the K variables, the incremental contribution to $R^2$ is the increase in $R^2$ that would be obtained if that variable had entered the regression last. Because the independent variables are not independent of each other, the sum of the incremental contributions will not be equal to $R^2$. The difference is called the multicollinearity effect and may be either positive or negative (Ref 9:163-180).

The incremental contributions were calculated using the formula

$$R^2 - R_h^2 = \frac{(1-R^2)t_h^2}{D} \qquad (2)$$

where $R^2$, $t_h^2$, and D all pertain to the single regression on all K variables. Specifically, $t_h^2$ is the square of the t-ratio for the $h^{th}$ variable (equivalent to the partial F value for the $h^{th}$ variable), and D is the degrees of freedom available for estimating the variance of the error term (D equals the number of observations minus the number of parameters in the regression equation (Ref 9:175).

# CHAPTER III

# DATA

## Data Collection and Grouping

Data was gathered on seven Prep School classes, beginning with the class that graduated from the Prep School in 1967. The class that graduated most recently, 1974, was not included in the study because Academy data was not yet available on it. From this point on the seven classes will be identified by the year they graduated or will graduate from the Academy. So this study involves Prep School graduates in the Academy class of 71 through the class of 77. The first four of these classes had graduated from the Academy at the time of this study, while the most recent three were still at the Academy.

The data collected was grouped three ways and separate regression analyses were done on each group. First, data was available on a larger number of potentially significant predictor factors for the two most recent Prep School classes. Thus, the classes of 76 and 77 comprised a grouping for which end of freshman year GPA was the dependent variable and on which the most extensive set of regression analyses were conducted. There were 164 observations in this grouping which will be referred to as the 76-77 grouping.

The entire seven classes formed the second grouping and contained 644 observations. End of freshman year GPA was used as the dependent variable and where possible the results of the analyses on the 71-77 grouping were compared with those on the 76-77 grouping.

The third grouping consisted of only the four classes that had graduated from the Academy, the classes of 71 through 74. A regression analyses using end of freshman year GPA was conducted on 387 observations representing the students in the 71-74 grouping who had completed their freshman year. In addition separate regression runs were made on 285 observations representing students in the 71-74 grouping who graduated from the Academy using end of freshman year, end of sophomore year, end of junior year, and end of senior year GPA successively as the dependent variable.

Only Prep School graduates who had completed their freshman year at the Academy were considered in this study and some of them were eliminated from the data base. For each of the classes of 74 through 77, the Prep School enrolled from five to twenty mid-year Regular entrants and these students were not considered because of their small number and because of the variety of ways they were integrated into the Prep School program from year to year. Of the remaining Prep School graduates who completed their first year at the Academy, about three percent were not

included in the data base because of missing data. It is assumed that these omissions are randomly distributed throughout the observations.

## Quantitative Independent Variables

Data was collected on the following measures of Prep School and prior to Prep School performance:

### Prior to Prep School

1. Prior Academic Record (PAR)
2. Athletic Activity Index
3. Nonathletic Activity Index
4. August CEEB scores (the CEEB tests taken just prior to beginning the Prep School program)

### Prep School

1. High CEEB scores (the highest scores on each of the four separate CEEB sections in all the times the student took the CEEB battery; Note that the four High CEEB scores may each have been taken on a different battery of CEEB tests.)
2. Physical Aptitude Exam (PAE)
3. Prep School Grades (for the years possible each of the following were recorded for data two-thirds of the way through the year and at the end of the school year)
   a. Cumulative GPA
   b. Academic cumulative GPA (omits physical training, PT, and military training, MT, grades)

c. Separate cumulative grade averages for math, English, and combined PT and MT

d. Academic percentage grade averages for math and English

e. Separate percentage grade averages for math and English

Qualitative Independent Variables

The following qualitative variables, which had a value of one when the student exhibited the quality and a value of zero otherwise, were identified for each student in the data base:

1. Accelerated math student at Prep School
2. Regular
3. Recognized athlete
4. Minority race
5. Father career military
6. Medically pilot or navigator qualified

All two-way interactions between these six qualitative variables were also used as variables. For example, if a student was both a Regular and a member of a minority race the qualitative variable for interaction between these two qualities would have a value of one, otherwise a value of zero. In addition, in order to control for differences between Prep School classes, a qualitative variable was used to indicate the student's class. This variable took on the integer values of zero or one, one through seven,

and one through four for the groupings 76-77, 71-77, and 71-74 respectively.

## Additional Variables Incorporated

Several of the independent variables were highly correlated with each other and so couldn't be included together in regression runs. The interested reader is referred to Theil (Ref 9), Draper and Smith (Ref 2), or to most any text on regression analysis for a description of what breaks down in the regression procedure when independent variables exhibit collinearity. In order to include some of these variables in the same regression runs, "difference" variables were incorporated.

August CEEB scores were highly correlated with high CEEB scores and so variables were formed that measured the difference between the High CEEB scores and the August CEEB scores. Then these "difference" variables were used together with the August CEEB scores in the same regression runs.

The various ways of measuring Prep School grades also exhibited collinearity. To compare Prep School academic cumulative GPA with Prep School overall cumulative GPA which includes physical and military training grades, it was desirable to have both of these measures represented in a single regression run. These two measures were naturally highly correlated so the academic cumulative GPA was used together with a "difference" variable that

measured the difference between academic cumulative GPA and overall cumulative GPA. This "difference" variable represented the contribution of physical and military training grades to overall cumulative GPA.

Variables were also incorporated to measure the difference between Prep School conventional grade point averages and Prep School percentage grade averages. After scaling the percentage grades to be comparable to the four point grading scale, "difference" variables were formed for math, English, and academic cumulative grades. To get the percentage grade averages and the conventional grade averages on a common scale a linear transformation was used that mapped the percentage middle A, 95 percent, to a 4.00; the percentage middle B, 85 percent, to a 3.00, and so on. This amounted to multiplying each percentage grade average by one tenth and subtracting 5.50. These "difference" variables were included in regression runs with the conventional grade averages.

Similarly, for the 76-77 data grouping "difference" variables were incorporated so that grades through the second academic quarter and the difference between end of year grades and end of second quarter grades could be in the same regression models. This provided a means to study the question of whether or not third (last) quarter grades contributed significantly to the predictive ability of Prep School grades.

## Problems in Obtaining Data

Direct data was available on most of the variables from Prep School and Academy permanent records. However, separate math, English, and combined PT and MT grades had to be computed for each student from records showing his thirty-six or more separate letter grades.

One of the main things the Prep School wanted studied was a comparison of the proposed percentage grading system for Prep School academic courses and the currently used letter grading system in terms of their impact on Academy GPA. In order to study this question, individual course percentage grades were calculated from historical data at the Prep School for each student represented in the data base. Then, for each student, math percentage grade averages were calculated as the average of all that student's individual math course percentage grades. Similarly, English percentage grade averages and academic percentage grade averages were calculated for each student.

Data was available to compute math percentage grade averages through the entire Prep School year for only the most recent two Prep School classes, the Academy classes of 76 and 77. Prior to that the grading used during the last third of the year was not standardized and overall raw percentage order of merits were not available. However, normalized math percentage grade averages were calculated

for the first two quarters on all seven of the Prep School classes.

Data was only available to compute English percentage grade averages through the first two quarters of the year and then only for the Academy classes of 76 and 77. Prior to the class of 76 the English raw percentage orders of merit were not on record. Again, last quarter grading was not standardized.

The normalized percentage grade averages that were computed are summarized in Table I.

Table I

Computed Normalized Percentage Grade Averages

| | | CLASS | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 71 | 72 | 73 | 74 | 75 | 76 | 77 |
| MATH | End of 2nd Qtr. | X | X | X | X | X | X | X |
| MATH | End of Year | | | | | | X | X |
| ENGLISH | End of 2nd Qtr. | | | | | | X | X |
| ENGLISH | End of Year | | | | | | | |

X indicates normalized percentage grades were computed.

## Comparison of Groupings

Appendix A defines symbols used for variables in subsequent tables in the text. For several of the quantitative variables, Appendix B provides a comparison of means, standard deviations, and simple correlations with

end of Academy freshman year GPA for the different data groupings. On the 71-74 Graduates grouping, simple correlations with end of sophomore, end of junior, and end of senior year GPA were also listed.

To compare one data grouping with another, statistical tests were performed to determine whether or not the means for each variable were significantly different. The test statistic used to compare two means was

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (3)$$

where $\bar{X}$ denotes the mean, s denotes the square of the standard deviation, and n denotes the number of observations. The subscript 1 denotes one data grouping and 2 denotes the other. For large sample sizes ($n \geq 30$) this statistic is approximately normal in distribution under the hypothesis that there is no difference in means. At the ninety-five percent confidence level, which was used by this writer, two means are significantly different when the test statistic Z takes on a value greater than 1.96 or less than -1.96.

In comparing the 76-77 grouping with the 71-77 grouping the following variables were found to have significantly different means:

NATHACT IDX: Nonathletic Activity Index

AENGACH: August English Achievement CEEB score

| | |
|---|---|
| AMTHAPT: | August Math Aptitude CEEB score |
| AMTHACH: | August Math Achievement CEEB score |
| ENGAPT: | High English Aptitude CEEB score |
| MTHAPT: | High Math Aptitude CEEB score |
| MTHACH: | High Math Achievement CEEB score |
| DMTHACH: | MTHACH minus AMTHACH |
| PAE: | Physical Aptitude Exam score at Prep School |
| MTHGPA: | Prep School math cumulative GPA through two-thirds of the year |
| %MTHGPA: | Prep School math percentage grade average through two-thirds of the year |

The significant differences between the means of the above variables indicates that the 76-77 grouping is not totally representative of the combined seven classes. It is interesting to note that while the means of math GPA differed significantly, the means of English GPA and cumulative GPA did not.

In comparing the 71-74 Graduates with the 71-74 Freshmen, none of the means were significantly different.

When the 71-74 Graduates grouping was compared with the 71-77 grouping the following variables had significantly different means:

| | |
|---|---|
| AENGAPT: | August English Aptitude CEEB score |
| AENGACH: | August English Achievement CEEB score |
| AMTHAPT: | August Math Aptitude CEEB score |
| AMTHACH: | August Math Achievement CEEB score |

| | |
|---|---|
| ENGAPT: | High English Aptitude CEEB score |
| MTHAPT: | High Math Aptitude CEEB score |
| MTHACH: | High Math Achievement CEEB score |
| MTHGPA: | Prep School math cumulative GPA through two-thirds of the year |
| %MTHGPA: | Prep School math percentage grade average through two-thirds of the year |

Thus, while no difference between the 71-74 Graduates grouping and the 71-74 Freshmen grouping could be detected by comparing means of individual variables, it can be concluded that the 71-74 Graduates grouping is not representative of the entire seven class grouping.

A comparison of the means and standard deviations for end of freshman year GPA for the different data groupings is provided in Table II. The table also includes data on GPA at the end of each year at the Academy for the 71-74 Graduates grouping.

Table II

Mean Academy Cumulative GPA by Data Grouping

| | DATA GROUPING | | | |
|---|---|---|---|---|
| | 76-77 | 71-77 | 71-74 Freshmen | 71-74 Graduates |
| Mean GPA 1st Year | 248 | 249 | 250 | 260 |
| Std Deviation | 46 | 48 | 49 | 45 |
| Mean GPA 2nd Year | | | | 255 |
| Std Deviation | | | | 41 |
| Mean GPA 3rd Year | | | | 258 |
| Std Deviation | | | | 39 |
| Mean GPA 4th Year | | | | 263 |
| Std Deviation | | | | 37 |

# CHAPTER IV

## 76-77 ANALYSIS

This chapter describes the regression analyses on the 76-77 data grouping. The analyses on the 71-77 and 71-74 data groupings follow in the next two chapters. However, since the classes of 76 and 77 were the most recent Prep School classes used in this study, it is reasonable to expect the results of the 76-77 analyses to be more indicative of results that could be expected for subsequent classes.

As mentioned in Chapter III, this was the only grouping for which Prep School percentage English grades were available. It is also the only grouping for which Prep School grades were used at the end of the academic year as well as after the second quarter. For the other groupings Prep School grades were looked at only after the second quarter because third quarter percentage grades were not available for either math or English.

### Variables Included in all 76-77 Runs

There were three ways Prep School grades were considered: overall GPA, academic GPA, and separate grade averages by subject. End of Academy freshman year GPA was the dependent variable on all 76-77 runs and unless otherwise stated all runs in both categories included the

following as potentially significant explanatory variables, which will be referred to as the "basic set:"

All the qualitative variables

All the interaction variables

Prior Academic Record before entering Prep School

Physical Aptitude Exam score

Athletic Activity Index

Nonathletic Activity Index

August CEEB scores

Differences between High CEEB and August CEEB scores

Prep School Cumulative GPA

First, the category of Prep School overall GPA was considered. As explained in Chapter II, "difference" variables were incorporated for certain pairs of variables that were highly correlated. In addition to the variables listed in the "basic set," the variable for Prep School cumulative GPA after the second quarter (CUMGPA) and the "difference" variable measuring the contribution of third quarter grades to the final cumulative GPA, (FCUMGPA-CUMGPA), were both included as potentially significant predictor variables in the first regression run.

The "best" regression equation obtained by the stepwise procedure was

$$Y = -.62 + .916x_1 + 1.57x_2 + 24.83x_3 - 10.49x_4 - 38.54x_5 \quad (4)$$

where

Y is predicted end of freshman year GPA

$x_1$ is CUMGPA

$x_2$ is (FCUMGPA-CUMGPA)

$x_3$ is ACCEL MATH: Accelerated math student at Prep School

$x_4$ is CLASS of 77

$x_5$ is INT-ACT (DADMIL-REGULAR): The interaction term between the variable for father career military (DADMIL) and the variable for prior active enlisted time in military service (REGULAR). See Appendix B.

Note, all the measures of GPA were scaled by a factor of 100. For example, 320 was used for a GPA of 3.20.

The value of $R^2$ was .613 and s, the standard error of the estimate, was 29.3. The above five explanatory variables had an affect on Academy GPA, together explaining 61.3 percent of the total variation. However, the value for $R^2$ is not high enough to reliably predict the response of an individual observation. That is, if the values of the five significant variables for a specific student were substituted into the equation, very little confidence could be placed in the predicted freshman year GPA. Fortunately, however, examination of the objective questions is not dependent on the regression equations having a high value for $R^2$. What variables are significant predictor factors and a measure of their relative importance can still be examined.

Recall that the qualitative variables have a value of 1 for observations having the quality and a value of 0 otherwise. Thus, the fact that the variable CLASS entered the model means that the constant term remains -.62 for observations from the class of 76, but becomes (-.62) plus (-10.49), or -11.11, for the class of 77. Similarly, the other significant qualitative variables cause the constant term to change for observations possessing the respective qualities. Thus, the three significant qualitative variables imply eight separate regression equations, one for each of the eight subsets of a three element set. The constant term would be the only thing that differed in the eight equations.

Instead of writing out the "best" regression equation as was done in equation (4), from this point on the format of Table III will be used to show the results of a "best" regression equation.

The variables are listed in the order they entered the stepwise regression model. The standard error of the estimated coefficient is a measure of the variability of the estimated coefficient and can be used in determining confidence intervals for the estimated coefficients. The partial F value is the square of the individual t-ratio and is simply the square of the estimated coefficient divided by the standard error of that estimated coefficient. All the regression runs in this study had large sample sizes

and, as discussed in Chapter II, 3.84 was the critical F value for significance at the ninety-five percent confidence level. Chapter II also included a discussion of the incremental contributions to $R^2$ which provide a measure of the part of $R^2$ that is accounted for by each significant variable.

Table III

Regression on 76-77 Data--Cumulative GPA

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| CUMGPA(271)* | .916 | .064 | 203.30 | .498 |
| FCUMGPA-CUMGPA(2)* | 1.57 | .24 | 42.47 | .104 |
| ACCEL-MATH | 24.83 | 7.16 | 12.02 | .030 |
| CLASS | -10.49 | 4.66 | 5.08 | .012 |
| INT-ACT (DADMIL-REGULAR) | -38.54 | 17.17 | 5.04 | .012 |

Constant Term: -.62
$R^2$ = .613
s = 29.3
Mean value of dependent variable: 248
*Mean value of variable in parentheses

Because of the large sample sizes, confidence intervals for the estimated coefficients in all the regression runs for all the data groupings can be calculated using

(ESTIMATED COEFFICIENT) $\pm$ 1.96 x (STD ERROR OF COEFF) (5)

For example, the true value of the CUMGPA coefficient lies between .790 and 1.04 with ninety-five percent confidence. Confidence intervals are discussed here to acquaint the unfamiliar reader with the fact that there is often sizeable variability associated with the estimated coefficients. Note that a good approximation of a ninety-five percent confidence interval is simply the estimated coefficient plus or minus twice the standard error of the estimated coefficient.

For the battery of potentially significant explanatory variables used in the first regression run, it can be concluded that Prep School cumulative GPA was the only significant quantitative predictor factor and that third quarter grades did contribute significantly to the predictive ability of Prep School cumulative GPA. Note that Prior Academic Record before entering the Prep School did not enter the model. Also, none of the CEEB scores entered.

## Prep School Academic GPA

The next regression run examined academic cumulative GPA versus overall cumulative GPA. It also examined final grades versus end of second quarter grades. In order to examine these, a rather involved set of "difference" variables was incorporated. Table IV shows the Prep School grades it was desirable to have in the same regression run.

Table IV

Variables for Academic GPA and Overall GPA

| | END OF 2ND QUARTER | END OF YEAR |
|---|---|---|
| ACADEMIC GPA | ACADCUM | FACADCUM |
| OVERALL GPA | CUMGPA | FCUMGPA |

All of these variables were highly correlated with each other. The variable ACADCUM was used directly and difference variables were used to include the affects of the other variables. Table V shows the difference variables that were used and what they measure.

Table V

Difference Variables for Contribution of Third Quarter and Nonacademic Grades

| | END OF 2ND QUARTER | 3RD QTR CONTRIBUTION |
|---|---|---|
| ACADEMIC GPA | ACADCUM | FACADCUM-ACADCUM |
| CONTRIBUTION OF NONACADEMIC GRADES TO OVERALL GPA | CUMGPA-ACADCUM | (FCUMGPA-FACADCUM) minus (CUMGPA-ACADCUM) |

Table VI gives the results of the "best" regression equation using the variables in Table V and the "basic set" as potentially significant explanatory variables.

Again Prep School grades, thi. time academic GPA, was the only significant quantitative factor and again third quarter grades contributed significantly to the

Table VI

Regression on 76-77 Data--Academic GPA

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ACADCUM(264)* | .781 | .054 | 208.39 | .513 |
| FACADCUM-ACADCUM(8)* | 1.31 | .20 | 43.09 | .106 |
| ACCEL MATH | 20.55 | 7.13 | 8.31 | .020 |
| INT-ACT (DADMIL-REGULAR) | -38.78 | 17.23 | 5.07 | .012 |

Constant term: 38.81
$R^2$ = .608
s = 29.37
Mean value of dependent variable: 248
*Mean value of variable in parentheses

predictive ability of the Prep School grades. However, since the difference between cumulative GPA and academic cumulative GPA was not a significant factor, physical and military training grades did not significantly contribute to the predictive ability of cumulative GPA. This still does not answer the question as to whether or not academic GPA is significantly better than overall GPA as a predictor factor. The $R^2$ values from Table III and Table VI cannot be compared because the model in Table III contains the variable CLASS while Table VI does not. In order to compare two models differing only by one having academic GPA and the other overall GPA, the variable CLASS was deleted from

the model in Table III and another regression run. This model, containing overall GPA, resulted in an $R^2$ of .600 which is .008 less than the value of $R^2$ for the model containing academic GPA (Table VI). Thus, less than an additional one percent of the total variation was explained by using academic cumulative GPA instead of overall cumulative GPA.

The predictive abilities of academic GPA and overall GPA were compared another way. Since the only quantitative variables in the models in both Table III and Table VI were the appropriate measures of Prep School grades, their relative importance was compared by examining the square of their simple correlations with Academy freshman GPA, which were obtained from Appendix B. Table VII provides the comparisons.

Table VII

Comparison of $R^2$ for Overall GPA and Academic GPA

| | OVERALL GPA | ACADEMIC GPA |
|---|---|---|
| END OF SECOND QUARTER GRADES | .489 | .491 |
| END OF YEAR GRADES | .558 | .567 |

The difference of .003 for end of second quarter grades and .009 for end of year grades indicates that using academic GPA explained less than an additional one percent

of the total variance over that explained by overall GPA. It is this writer's opinion that this does not indicate that academic cumulative GPA is a significantly better predictor factor than overall GPA. Thus, while physical and military training grades did not contribute to the predictive ability of overall GPA, neither did they detract.

In the same manner, academic GPA was compared with percentage academic grade average. From Appendix B the simple correlation between ACADCUM and freshman year GPA was .701 and r between %ACADCUM and freshman year GPA was .721. The corresponding values of $R^2$ are .491 and .520. The difference of .029 indicates that an additional 2.9 percent of the total variance was explained by percentage academic cumulative GPA over conventional academic cumulative GPA.

Separate Prep School Grades

The other basic category of Prep School grades was separate grades by subject. The variables listed as the "basic" set were also included as potentially significant predictor variables in all runs in this category.

The separate grades considered were math, English, combined physical and military training, percentage math, and percentage English. In order to include the affect of both conventional and percentage grades, "difference" variables were used. "Difference" variables were also included in order to examine end of second quarter grades

versus final grades. The variables included as potentially significant predictor factors are listed below. Their values after the second academic quarter were used and the difference between their end of year values and end of second quarter values were also used to provide a measure of the affect of third quarter grades.

English GPA

Math GPA

Combined PT and MT GPA

Difference between math percentage grade average and math GPA. (The percentage grade averages were first scaled as described in Chapter III.)

In addition, for end of second quarter grades, the difference between English percentage grades were not available for the third quarter.

Table VIII provides the results of the "best" regression equation obtained by the stepwise multiple regression procedure. The results were similar to the results obtained for overall GPA. Prep School separate grades were the only significant quantitative factors and again third quarter grades added to the predictive ability of the Prep School grades. Math GPA explained a larger portion of the variance than did English GPA. The variables measuring the difference between conventional grades and percentage grades did not enter the model and thus the model did not detect any significant contribution by percentage grades. However, because the simple correlation

coefficients of all the math and English percentage grade averages were somewhat higher than the correlation coefficients of the corresponding conventional grade point averages (see Appendix B), and because the percentage academic GPA explained an additional 2.9 percent of the total variance over conventional academic GPA, the relative importance of the two grading methods was examined further.

Table VIII

Regression on 76-77 Data--Separate Grade Averages

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| MTHGPA(268)* | .412 | .039 | 109.94 | .271 |
| ENGGPA(260)* | .384 | .045 | 73.95 | .182 |
| FENGGPA-ENGGPA(4)* | .777 | .156 | 24.91 | .061 |
| FMTHGPA-MTHGPA(-2)* | .588 | .131 | 20.23 | .050 |
| ACCEL MATH | 18.86 | 7.26 | 6.75 | .017 |
| INT-ACT (DADMIL-REGULAR) | -41.27 | 17.38 | 5.64 | .014 |

Constant term: 34.39
$R^2$ = .613
s = 29.37
Mean value of dependent variable: 248
*Mean value of variable in parentheses

A regression run was made on just math GPA and English GPA after the second quarter and another run was made using just percentage math and percentage English

grade averages after the second quarter. These runs did not include any other variables as potentially significant explanatory variables. Math GPA with English GPA resulted in an $R^2$ of .492 and math percentage grade average with English percentage grade average resulted in an $R^2$ of .520. The difference of .028 means that an additional 2.8 percent of the total variance was explained by percentage grades over conventional grades.

These values of $R^2$ were compared with the values of $R^2$ for academic GPA and percentage academic GPA to examine separate grades versus combined grades. Table IX provides the comparison for the end of second quarter grades.

Table IX

Comparison of $R^2$ for End of Second Quarter Conventional, Percentage, Separate, and Academic Grades

| | SEPARATE MATH, ENGLISH GRADES | ACADEMIC GPA |
|---|---|---|
| CONVENTIONAL GRADES | .492 | .491 |
| PERCENTAGE GRADES | .520 | .520 |

It is obvious that using math and English grades separately gained nothing, in terms of variance explained, over using the academic cumulative GPA.

Two more runs were made in order to compare final math GPA with final percentage math average. Both runs

included final English GPA and again no other variables were allowed to enter the model. Final math GPA with final English GPA resulted in an $R^2$ of .567. Final math percentage grade average with final English GPA resulted in an $R^2$ of .574. Thus, an additional .7 percent of the total variance was explained by using final math percentage grade average instead of conventional math GPA.

Thus, percentage grades were slightly better predictor factors than conventional grades, where the criterion for "better" was explaining a larger portion of the total variance in Academy freshman year GPA.

Table X summarizes the values of $R^2$ for the various ways that percentage and conventional grades have been considered for both end of year and end of second quarter.

Table X

Comparison of $R^2$ for Different Grade Measures
At End of Second Quarter and End of Year

| | $R^2$ | |
|---|---|---|
| | END OF SECOND QUARTER GRADES | END OF YEAR (FINAL) GRADES |
| CUMGPA | .489 | .558 |
| ACADCUM | .491 | .567 |
| %ACADCUM | .520 | - |
| SEPARATE MATH, ENGLISH | .492 | .567 |
| SEPARATE %MATH, %ENGLISH | .520 | - |
| SEPARATE %MATH, ENGLISH | - | .574 |

## Conclusions About the 76-77 Grouping

The results of the regression analyses on the 76-77 data grouping are summarized below in terms of answers to the "Objective Questions." The third objective question, which will be examined using the results of the 71-74 grouping, is the only objective question that cannot be addressed with the results of the 76-77 grouping.

### Question One

Among the available and potentially available measures of Prep School performance, only Prep School grades were found to be significant quantitative factors in terms of predictive modeling of Air Force Academy freshman grade point average for Air Force Academy Prep School graduates.

### Question Two

The relative importances of the significant predictor factors were determined by the magnitudes of their incremental contributions to $R^2$. Under this criterion for relative importance, math GPA was more important than English GPA for both conventional and percentage methods of measurement.

### Question Four

Prep School academic GPA was not a significantly better predictor factor than overall cumulative GPA which includes military and physical training grades.

Question Five

Using both math and English grade point averages, but separately, was not better than using academic cumulative GPA.

Question Six

The proposed Prep School percentage grading method resulted in slightly higher predictive ability than the conventional grading method.

Question Seven

Prep School grades earned during the last third of the Prep School academic year did significantly increase the predictive ability of Prep School grades.

Question Eight

Neither August CEEB scores nor High CEEB scores were found to be significant predictor factors.

# CHAPTER V

## 71-77 ANALYSIS

The analysis on the 71-77 data grouping is described in this chapter. As with the 76-77 analyses, end of freshman year GPA was the dependent variable on all the regression runs in this grouping. Later in the chapter the results of the 71-77 analyses are compared with the results on the 76-77 grouping.

### Variables Included in All 71-77 Runs

For the 71-77 data grouping all measures of Prep School grades were at the end of the second academic quarter. All 71-77 runs included the following as potentially significant explanatory variables, which will be referred to as the "basic set."

All the qualitative variables

All the interaction variables

Prior Academic Record before entering Prep School

Physical Aptitude Exam score

Athletic Activity Index

Nonathletic Activity Index

August CEEB scores

Differences between High CEEB and August CEEB scores

## Prep School Cumulative GPA

In addition to the basic set the first regression run included Prep School cumulative GPA as a potentially significant explanatory variable. The results of this stepwise regression run are tabulated in Table XI. The format for the regression results is the same as in the previous chapter, with the variables listed in the order they entered the stepwise procedure.

Table XI

Regression on 71-77 Data--Cumulative GPA

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| CUMGPA(275)* | .604 | .038 | 246.90 | .209 |
| AENGAPT(512)* | .182 | .027 | 44.14 | .037 |
| DENGAPT(62)* | .172 | .036 | 22.51 | .019 |
| AENGACH(478)* | -.104 | .025 | 17.56 | .015 |
| PRIOR PAR(491)* | .046 | .018 | 6.62 | .006 |
| DADMIL | -11.16 | 3.53 | 9.96 | .008 |
| INT-ACT (ACCEL-DADMIL) | 13.12 | 6.05 | 4.70 | .004 |

Constant term: 9.36
$R^2$ = .462
s = 35.43
Mean value of dependent variable: 249
*Mean value of variable in parentheses

Five quantitative variables entered the model at significant levels, with Prep School cumulative GPA

explaining the bulk of the total variance. The variable DENGAPT entering significantly indicates that the increase, while at the Prep School, in the English Aptitude CEEB scores was significant in explaining part of the total variance in freshman GPA.

The reader needs to be aware of the fact that while the English CEEB scores entered the model and the math CEEB scores did not, this does not imply that English CEEB scores are better predictors of Academy GPA than math CEEB scores when either are used alone. Appendix B gives the simple correlations with end of freshman GPA and both math CEEB scores have higher correlation coefficients than their corresponding English CEEB scores for the 71-77 data grouping. Thus, if used alone, math CEEB scores were better predictor factors of Academy freshman GPA than were English CEEB scores. However, all the variables in the regression model can be thought of as working together to explain as much of the total variation in freshman GPA as possible. The fact that English CEEB scores entered the model simply means that the English CEEB scores explained some of the total variation in freshman GPA that was not explained by the other significant explanatory variables in the model, in this case Prep School cumulative GPA and Prior Academic Record before entering the Prep School. Similarly, the fact that math CEEB scores did not enter the model means that the math CEEB scores did not explain

any of the total variation that was not already explained by CUMGPA, PRIOR PAR, and the English CEEB scores.

The reader also needs to be aware that the negative coefficient for the August English Achievement (AENGACH) variable does not imply that the higher the AENGACH score the lower the expected Academy freshman GPA will be. The entries in Appendix B show that there is a significant positive correlation between AENGACH and freshman GPA. As explained in Chapter II, the coefficient on the AENGACH variable in the regression model cannot be interpreted as a marginal contribution to predicting freshman GPA because the AENGACH variable is not independent of the other explanatory variables. All the variables in the regression model work together and interact with each other and the individual coefficients are not necessarily meaningful in terms of the marginal contributions of single variables.

## Prep School Academic GPA

The next regression run examined Prep School academic cumulative GPA. In addition to the basic set, this run included Prep School academic cumulative GPA and the difference between overall cumulative GPA and academic cumulative GPA as potentially significant predictor variables. Table XII shows the results.

The fact that the difference variable (CUMGPA minus ACADCUM) entered the model at a significant level

means that physical and military training grades significantly contributed to the predictive ability of cumulative GPA.

Table XII

Regression on 71-77 Data--Academic GPA

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ACADCUM(270)* | .593 | .042 | 198.11 | .168 |
| CUMGPA-ACADCUM(5)* | .523 | .139 | 14.26 | .012 |
| AENGAPT(512)* | .178 | .028 | 40.98 | .035 |
| DENGAPT(62)* | .170 | .036 | 21.62 | .018 |
| AENGACH(487)* | -.105 | .025 | 17.83 | .015 |
| PRIOR PAR(491)* | .046 | .018 | 6.64 | .006 |
| DADMIL | -10.96 | 3.55 | 9.51 | .008 |
| INT-ACT (ACCEL-DADMIL) | 12.81 | 6.08 | 4.45 | .004 |

Constant term: 14.86
$R^2$ = .462
s = 35.45
Mean value of dependent variable: 249
*Mean value of dependent variable in parentheses

Separate Prep School Grade Averages

Next, separate Prep School grade averages were used. The potentially significant explanatory variables were the basic set plus math GPA, English GPA, combined physical and military training GPA, and the variable for

the difference between math GPA and percentage math grade average. Table XIII gives the outcome of the regression run.

Table XIII

Regression on 71-77 Data--Separate Grade Averages

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| MTHGPA(284)* | .250 | .023 | 120.53 | .102 |
| ENGGPA(255)* | .236 | .028 | 69.18 | .059 |
| PTMTGPA(294)* | .106 | .028 | 14.36 | .012 |
| AENGAPT(512)* | .180 | .028 | 40.81 | .035 |
| DENGAPT(62)* | .172 | .037 | 21.62 | .018 |
| AENGACH(487)* | -.103 | .025 | 16.53 | .014 |
| PRIOR PAR(491)* | .046 | .018 | 6.72 | .006 |
| DADMIL | -10.91 | 3.56 | 9.40 | .008 |
| INT-ACT (ACCEL-DADMIL) | 13.12 | 6.05 | 4.70 | .004 |

Constant term: 13.21
$R^2$ = .462
s = 35.48
Mean value of dependent variable: 249
*Mean value of variable in parentheses

## Conclusions About 71-77 Grouping

Within three significant digits the value of $R^2$ was identical for all three models on the 71-77 data grouping. On all three models the same variables (other than the different measures of Prep School grades) entered as

significant predictor factors. The conclusions about the results of the 71-77 data grouping are given below in terms of answers to the applicable "Objective Questions."

Question One

In terms of predicting freshman year GPA, the significant predictor factors were Prep School grades, English CEEB scores, and Prior Academic Record before entering the Prep School.

Question Two

Prep School grades were by far the most important predictor factors and when looked at separately, math GPA was more important than English GPA. Physical and military training grades were significant, but contributed very little in comparison to math and English grades. English CEEB scores were more important than Prior Academic Record before entering the Prep School.

Question Four

Prep School academic GPA was not a better predictor factor than overall cumulative GPA.

Question Five

Using separate Prep School math and English GPA was not better than using academic GPA.

Question Six

Percentage math grading method was not found to be better than conventional math GPA.

Question Eight

English CEEB scores were found to be significant predictor factors. Given the August English CEEB scores, the High English Achievement score contributed nothing in terms of increasing predictive ability, while the High English Aptitude score did.

## Comparison of 71-77 and 76-77 Results

The results on the 71-77 grouping cannot be compared directly with the results on the 76-77 grouping because the affect of third quarter Prep School grades were considered in the 76-77 grouping, but not in the 71-77 grouping. In order to make comparison meaningful, three additional regression runs were made on the 76-77 grouping using Prep School grade measures just through the second academic quarter. For each measure of Prep School grades the potentially significant predictor factors were the same as the ones used in the corresponding 71-77 run. The results are presented in Tables XIV, XV, and XVI.

Table XIV

Regression on 76-77 Data--Cumulative GPA
at End of Second Quarter

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| CUMGPA(271)* | .668 | .073 | 83.8 | .255 |
| PRIOR PAR(484)* | .104 | .038 | 7.32 | .022 |

Constant term: 18.74
$R^2$ = .510
s = 32.37
Mean value of dependent variable: 248
*Mean value of variable in parentheses

Table XV

Regression on 76-77 Data--Academic GPA
at End of Second Quarter

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ACADCUM(264)* | .571 | .061 | 86.40 | .261 |
| PRIOR PAR(484)* | .103 | .039 | 7.16 | .022 |

Constant term: 47.73
$R^2$ = .513
s = 35.54
Mean value of dependent variable: 248
*Mean value of variable in parentheses

Table XVI

Regression on 76-77 Data--Separate Grade Averages at End of Second Quarter

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| MTHGPA(268)* | .315 | .043 | 53.84 | .160 |
| ENGGPA(260)* | .281 | .044 | 41.70 | .124 |
| PRIOR PAR(484)* | .096 | .039 | 6.17 | .018 |
| %MTHGPA-MTHGPA(-3)* | .380 | .185 | 4.20 | .012 |

Constant term: 45.72
$R^2$ = .562
s = 32.32
Mean value of dependent variable: 248
*Mean value of variable in parentheses

The results of the above three runs on the 76-77 grouping differ somewhat from the results of the previous chapter. No qualitative variables entered in the above three runs and PRIOR PAR did enter in all the above, but not in any of the runs in the previous chapter. Thus, given Prep School grades through the end of the school year, PRIOR PAR contributed nothing towards explaining any of the remaining variation, but when Prep School grades were used only through the second quarter, PRIOR PAR explained a small but significant portion of the remaining variation.

The similarities and differences between the 71-77 runs and the above three 76-77 runs are described below.

Similarities

Both groupings had Prep School grades as the main contributor to the prediction model. Neither grouping resulted in academic GPA being a better predictor factor than overall cumulative GPA.

Differences

English CEEB scores were significant predictor factors for the 71-77 grouping, but not for the 76-77 grouping. Math percentage grade average was a slightly better predictor factor than conventional math GPA for the 76-77 grouping, but not for the 71-77 grouping.

# CHAPTER VI

## 71-74 ANALYSIS

The 71-74 data grouping was formed for the single purpose of examining the third objective question which deals with determining the significant factors for separately predicting Academy GPA at the end of each year at the Academy. The other objective questions have already been examined using the analyses on the most current data available (the 76-77 data grouping) and will not be addressed for this grouping.

### Regression Runs on 71-74 Graduates

The 285 observations representing the students in the 71-74 grouping who completed all four years at the Academy are referred to as the "71-74 Graduates" in Appendix B which, for each of the potentially significant quantitative variables, provides the mean, standard deviation, and simple correlation with cumulative GPA at the end of each year at the Academy. This 71-74 Graduates grouping was used in four stepwise multiple regression runs with cumulative Academy GPA at the end of the freshman year as the dependent variable for the first run, end of sophomore year cumulative GPA as the dependent variable for the second run, end of junior year cumulative GPA for the third run, and end of senior year cumulative GPA for the fourth

run. All four runs included the following as potentially significant predictor factors:

Math GPA

Difference between math GPA and percentage math grade average

English GPA

Combined physical and military training GPA

Differences between High CEEB and August CEEB scores

Prior Academic Record before entering Prep School

Physical Aptitude Exam score

Athletic Activity Index

Nonathletic Activity Index

All the qualitative variables

All the interaction variables

As with the 71-77 analyses in the previous chapter, all measures of Prep School grades were at the end of the second academic quarter. The results of the four regression runs are presented in the tables that follow.

There are several things worth noting in comparing the factors which were significant for predicting GPA at the end of each year at the Academy. Prep School English GPA was the most important contributor towards explaining variance for all four years. August English CEEB scores were significant all four years as was the improvement in English Aptitude while at the Prep School. Prior Academic

Table XVII

Regression on 71-74 Graduates--Separate Grade Averages
GPA at End of Freshman Year as Dependent Variable

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ENGGPA(259)* | .239 | .043 | 30.85 | .069 |
| MTHGPA(300)* | .167 | .035 | 22.62 | .051 |
| INT-ACT (ACCEL-MINORITY) | -68.09 | 25.80 | 6.97 | .016 |
| PRIOR PAR(498)* | .075 | .026 | 8.21 | .018 |
| AENGACH(496)* | -.139 | .040 | 11.90 | .027 |
| AENGAPT(523)* | .200 | .042 | 23.01 | .056 |
| DENGAPT(62)* | .199 | .058 | 11.94 | .027 |
| DADMIL | -10.64 | 5.04 | 4.44 | .010 |

Constant term: 66.01
$R^2$ = .378
s = 35.93
Mean value of the dependent variable: 260
*Mean value of variable in parentheses

Table XVIII

Regression on 71-74 Graduates--Separate Grade Averages
GPA at end of Sophomore Year as Dependent Variable

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ENGGPA(259)* | .234 | .038 | 37.36 | .081 |
| MTHGPA(300)* | .121 | .032 | 14.72 | .032 |
| AENGAPT(523)* | .224 | .037 | 36.37 | .079 |
| AENGACH(496)* | -.130 | .036 | 13.12 | .028 |
| DENGAPT(62)* | .190 | .051 | 14.03 | .030 |
| PRIOR PAR(498)* | .059 | .023 | 6.43 | .014 |
| DADMIL | -10.09 | 4.49 | 5.04 | .011 |
| INT-ACT (REGULAR-MINORITY) | -35.78 | 16.28 | 4.83 | .010 |

Constant term: 67.69
$R^2$ = .401
s = 31.99
Mean value of the dependent variable: 255
*Mean value of variable in parentheses

## Table XIX

Regression on 71-74 Graduates--Separate Grade Averages
GPA at End of Junior Year as Dependent Variable

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ENGGPA(259)* | .283 | .036 | 60.32 | .134 |
| AMTHAPT(617)* | .134 | .036 | 13.52 | .030 |
| DMTHAPT(64)* | .122 | .052 | 5.40 | .012 |
| AENGACH(496)* | -.198 | .050 | 15.60 | .035 |
| AENGAPT(523)* | .204 | .045 | 20.67 | .046 |
| DENGAPT(62)* | .184 | .054 | 11.84 | .026 |
| PRIOR PAR(498)* | .058 | .023 | 6.52 | .014 |
| DENGACH(116)* | -.102 | .049 | 4.36 | .010 |
| INT--ACT (REGULAR-MINORITY) | -31.76 | 15.84 | 4.02 | .009 |

Constant term: 58.25
$R^2$ = .384
s = 31.24
Mean value of dependent variable: 258
*Mean value of variable in parentheses

Table XX

Regression on 71-74 Graduates--Separate Grade Averages
GPA at End of Senior Year as Dependent Variable

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ENGGPA(259)* | .254 | .034 | 57.74 | .123 |
| AMTHAPT(617)* | .116 | .034 | 11.63 | .026 |
| DMTHAPT(64)* | .126 | .049 | 6.48 | .014 |
| INT-ACT (REGULAR-NONRATED) | -21.59 | 10.19 | 4.49 | .010 |
| INT-ACT (REGULAR-MINORITY) | -35.10 | 14.86 | 5.58 | .012 |
| PRIOR PAR(498)* | .056 | .022 | 6.84 | .015 |
| AENGACH(496)* | -.115 | .033 | 12.07 | .027 |
| AENGAPT(523)* | .143 | .036 | 15.45 | .035 |
| DENGAPT(62)* | .134 | .047 | 8.02 | .018 |

Constant term: 65.33
$R^2$ = .379
s = 29.45
Mean value of dependent variable: 260
*Mean value of variable in parentheses

Record was the smallest quantitative contributor towards explaining variance in Academy GPA for all four models.

The only striking difference in the predictor factors from year to year was the fact that math GPA entered significantly for only the first two years while August Math Aptitude and improvement in Math Aptitude were significant predictor factors for only the last two years. The reader needs to be aware that while the Math Aptitude CEEB variable replaced Math GPA for the junior and senior year models, this does not mean Math GPA was not significantly correlated with Academy GPA after the junior and senior years. Math GPA was significantly correlated with Academy GPA at the end of each year at the Academy (see Appendix B). The implication is simply that the Math Aptitude CEEB variable explained basically the same variation that Math GPA would have explained and more.

The overall fit of each model to the data, as measured by $R^2$, was about the same for all the models with the model for GPA at the end of the sophomore year having the highest $R^2$.

Regression Run on 71-74 Freshmen

The "71-74 Freshmen" grouping refers to 387 observations in the 71-74 grouping representing students who had completed at least their freshman year at the Academy. Chapter III describes that in comparing the 71-74 Graduates grouping with the 71-74 Freshmen grouping, none of the means

of the potentially significant explanatory variables were found to be significantly different for the two groupings. That is, for each explanatory variable, the mean for the 71-74 Graduates was not significantly different from the mean for the 71-74 Freshmen at the ninety-five percent confidence level. However, it cannot be concluded that the two groups are necessarily representative of each other and so a step-wise regression run was made on the 71-74 Freshman grouping using end of freshmen year GPA as the dependent variable. The results are given in Table XXI.

Table XXI

Regression on 71-74 Freshmen--Separate Grade Averages
GPA at End of Freshman Year as Dependent Variable

| Variable | Estimated Coefficient | Std Error of Coeff | Partial F Value | Incremental Contribution to $R^2$ |
|---|---|---|---|---|
| ENGGPA(251)* | .221 | .039 | 32.66 | .049 |
| MTHGPA(290)* | .237 | .031 | 57.50 | .087 |
| AENGAPT(518)* | .222 | .037 | 35.21 | .053 |
| PTMTGPA(294)* | .106 | .034 | 9.50 | .014 |
| DENGAPT(62)* | .202 | .049 | 16.69 | .025 |
| AENGACH(494)* | -.116 | .035 | 10.98 | .017 |
| PRIOR PAR(488)* | .053 | .024 | 4.94 | .007 |

Constant term: -1.15
$R^2$ = .427
s = 37.53
Mean value of dependent variable: 250
*Mean value of variable in parentheses

The mean Academy GPA at the end of the freshman year was 250 for the 387 observations in the 71-74 Freshmen grouping and was 260 for the 285 observations in the 71-74 Graduates grouping. The mean end of freshman year GPA for the 102 students who completed their freshman year, but for some reason did not graduate was calculated to be 222.

In comparing the models for predicting Academy GPA at the end of the freshman year for the 71-74 Graduates (Table XVII) and the 71-74 Freshman (Table XXI), the relative importances of math GPA and English GPA are reversed in the two models. For all the Prep School graduates who finished their freshman year (Table XXI), math GPA had almost twice the incremental contribution to $R^2$ that English GPA had. However, using just the Prep School students who graduated from the Academy (Table XVII), the incremental contribution to $R^2$ by English GPA was significantly higher than the contribution by math GPA. This reversal in relative importance of two significant variables for two subgroupings of the 71-74 grouping is an indication that slight differences in results are likely to occur for different subsets of a population.

It is also interesting to note that combined physical and military training GPA was a significant predictor factor for the grouping of all the students who

finished their freshman year, but was not for the grouping of just the students who graduated.

## Conclusions About 71-74 Grouping

The third Objective Question is the only Objective Question addressed with the 71-74 analyses.

### Question Three

The factors that were significant for predicting Academy end of freshman GPA remained the same and in the same position of relative importance for predicting Academy GPA at the end of the sophomore, junior and senior year, with one exception. The one exception was that math GPA, which was significant in the models for predicting GPA at the end of the freshman and sophomore years, was replaced by Math Aptitude CEEB score for the models predicting GPA at the end of the junior and senior years.

# CHAPTER VII

## ASSUMPTIONS AND OVERALL CONCLUSIONS

Conclusions based on the results of the multiple regression analyses for each data grouping have been presented in the three analysis chapters. However, the value to the Prep School of the results of this study are mainly in their applicability to future Prep School classes. The only explicit assumptions made for the analyses were relative to the regression model and were described in Chapter II. If some additional assumptions are made, overall conclusions about future Prep School classes can be stated.

### Assumptions

1. It is assumed that the 76-77 data grouping is representative of subsequent (future) classes. This assumption allows the conclusions based on the 76-77 analyses to be extrapolated to subsequent classes. It is this writ r's opinion that this is a reasonable assumption as long as the extrapolation is not extended for more than a few subsequent classes, say four or five, and only if the Prep School and Academy academic programs do not change significantly. The fact that significant predictor factors have changed over time in the past was evidenced by comparing the results of the 76-77 grouping with the results of the 71-77 grouping. While most of the conclusions

on the two groupings were identical, English CEEB scores were significant predictor factors for the 71-77 grouping, but not for the 76-77 grouping.

2. It is assumed that the 76-77 data grouping is representative of all Prep School graduates who entered the Academy classes of 76 and 77 and not just those who finished at least their first year at the Academy. In other words it is assumed that if all the Prep School graduates entering the Academy classes of 76 and 77 would have completed their first year at the Academy the relative impact of the various explanatory variables on their Academy GPA would have been the same as it was for the 76-77 data grouping. This is a big assumption because approximately thirty percent of the Prep School graduates entering the Academy classes of 76 and 77 failed to complete their first year at the Academy, and because the slight differences between the results of the 71-74 Freshmen and the 71-74 Graduates groupings warns that different subsets of a population are likely to yield slightly different results. It is this writer's opinion that in terms of applicability to future Prep School classes some of the conclusions about the 76-77 analyses are fairly insensitive to the validity of this assumption while others are fairly sensitive. The insensitive conclusions will be included in a "strong" conclusions category and the sensitive ones in a "secondary" conclusions category.

3. It is assumed that the result of the 71-74 analyses that significant predictor factors remain the same and in the same relative position of importance for predicting Academy GPA at the end of each year at the Academy is applicable to future classes. It is assumed that the one exception to the above result for the 71-74 grouping (the variable math GPA being replaced by the variable Math Aptitude CEEB score for predicting junior and senior year cumulative GPA) is not applicable to future classes. Based on simple correlations described below it is this writer's opinion that the deletion of the above exception is valid.

In comparing the simple correlations with freshman GPA (see Appendix B) of math GPA and Math Aptitude CEEB scores it is observed that in going from the 71-74 grouping to the 76-77 grouping the correlation coefficients for math GPA significantly increase, while the correlation coefficients for Math Aptitude (August and High) significantly decrease. Thus, if the data were available to run regressions for predicting GPA at the end of each year at the Academy for the 76-77 grouping, it is reasonable to expect that the variable math GPA would be significant in all the runs and would not be replaced by a variable measuring Math Aptitude CEEB score. Since it has been assumed that the 76-77 grouping is indicative of future Prep School classes, it is reasonable to extrapolate these anticipated results to future classes.

## Overall Conclusions Relative to Future Prep School Classes

### Strong Conclusions

These conclusions are strongly supported by the results of this study.

1. Prep School grades are significant factors in predictive modeling of Academy freshman GPA, and given Prep School grades as predictor factors, the quantitative factors Prior Academic Record before entering the Prep School, August CEEB scores, High CEEB scores, Physical Aptitude Exam Score, Athletic Activity Index, and Non-athletic Activity Index do not significantly contribute to predicting Academy freshman GPA.

2. Prep School cumulative GPA, which includes physical training and military training grades, predicts Academy freshman GPA as well as Prep School Academic GPA or separate math and English grade point averages.

3. Third quarter Prep School grades significantly increase the ability of Prep School grades to predict Academy freshman GPA.

4. The models developed in this study are not adequate to reliably predict Academy freshman GPA for individual Prep School students.

### Secondary Conclusions

These conclusions are supported by the results of this study, but the supportive evidence is not as

overwhelming as for the "strong conclusions," and in terms of applicability to future Prep School classes the "secondary conclusions" are somewhat more sensitive to the validity of the second and third stated assumptions.

1. The proposed Prep School percentage grading method results in slightly higher ability to predict Academy freshman GPA than does the conventional grading method.

2. The predictor factors determined to be significant in the 76-77 analyses are expected to remain the same and in the same relative positions of importance in predictive modeling of Academy GPA at the end of each Academy year for future Prep School classes.

## Bibliography


1. Dixon, W. J., editor. *BMD Biomedical Computer Programs* (Second Edition). University of California Publications in Automatic Computation No. 2. Berkeley: University of California Press, 1970.

2. Draper, N. R. and H. Smith. *Applied Regression Analysis*. New York: John Wiley and Sons, Inc., 1966.

3. Goldberger, Arthur S. *Econometric Theory*. New York: John Wiley and Sons, Inc., 1964.

4. Goldberger, Arthur S. *Topics in Regression Analysis*. New York: The Macmillan Company, 1968.

5. Johnston, J. *Econometric Methods* (Second Edition). New York: McGraw-Hill Book Company, 1972.

6. Kerlinger, Fred N. *Foundations of Behavioral Research* (Second Edition). New York: Holt, Rinehart and Winston, Inc., 1973.

7. Mendenhall, William. *Introduction of Linear Models and The Design and Analysis of Experiments*. Belmont, California: Wadsworth Publishing Company, Inc., 1968.

8. Mendenhall, William and Richard L. Scheaffer. *Mathematical Statistics with Applications*. Belmont, California: Wadsworth Publishing Company, Inc., 1973.

9. Theil, Henri. *Principles of Econometrics*. New York: John Wiley and Sons, Inc., 1971.

10. United States Air Force Academy. *History of the United States Air Force Academy Preparatory School 1972-1973*. Unpublished history. United States Air Force Academy Colorado: United States Air Force Academy, 1973.

11. Welch, Finnis. *The Determinants of the Return to Schooling in Rural Farm Areas*. Unpublished dissertation. Chicago: The University of Chicago, 1966.

# APPENDIX A

## SYMBOLS DEFINED

### Independent Quantitative Variables

| | |
|---|---|
| PRIOR PAR | Prior Academic Record before coming to Prep School |
| ATHACT IDX | Athletic Activity Index (high school) |
| NATHACT IDX | Nonathletic Activity Index (high school) |
| AENGAPT | August English Aptitude CEEB score |
| AENGACH | August English Achievement CEEB score |
| AMTHAPT | August Math Aptitude CEEB score |
| AMTHACH | August Math Achievement CEEB score |
| ENGAPT | High English Aptitude CEEB score |
| ENGACH | High English Achievement CEEB score |
| MTHAPT | High Math Aptitude CEEB score |
| MTHACH | High Math Achievement CEEB score |
| DENGAPT | ENGAPT minus AENGAPT |
| DENGACH | ENGACH minus AENGACH |
| DMTHAPT | MTHAPT minus AMTHAPT |
| DMTHACH | MTHACH minus AMTHACH |
| PAE | Physical Aptitude Exam score at Prep School |
| CUMGPA | Prep School cumulative GPA through two-thirds of the year |
| FCUMGPA | Prep School final cumulative GPA |
| MTHGPA | Prep School math cumulative GPA through two-thirds of the year |

| | |
|---|---|
| FMTHGPA | Prep School final math cumulative GPA |
| ENGGPA | Prep School English cumulative GPA through two-thirds of the year |
| FENGGPA | Prep School final English cumulative GPA |
| FPMTGPA | Combined Prep School Physical and Military Training cumulative GPA through two-thirds of the year |
| FPTMTGPA | Final Prep School Physical and Military Training combined cumulative GPA |
| ACADCUM | Prep School academic cumulative GPA through two-thirds of the year |
| FACADCUM | Prep School final academic cumulative GPA |
| %MTHGPA | Prep School math percentage grade average through two-thirds of the year |
| F%MTHGPA | Prep School final math percentage grade average |
| %ENGGPA | Prep School English percentage grade average through two-thirds of the year |
| %ACADCUM | Prep School academic percentage grade average through two-thirds of the year |

Independent Qualitative Variables

| | |
|---|---|
| ACCEL MATH | Accelerated math student at Prep School |
| REGULAR | Prior active enlisted time in Military Service |
| ATHLETE | Recognized athlete |
| MINORITY | Minority race |
| DADMIL | Father career military |
| NONRATED | Not medically qualified for pilot or navigator training |
| CLASS | Prep School class |

Interaction Terms

All two-way interaction terms between the qualitative variables are identified by the symbol INT-ACT followed by the symbols for the two variables in parenthesis. For example, INT-ACT (DADMIL-ATHLETE) represents the variable for interaction between DADMIL and ATHLETE.

# APPENDIX B

## MEANS, STANDARD DEVIATIONS, AND SIMPLE CORRELATIONS WITH ACADEMY GPA

| | CLASS | | | | | | |
|---|---|---|---|---|---|---|---|
| Variable | 76-77 | 71-77 | 71-74 Fresh. | 71-74 Grad. | | | |
| | | | | | $r_2$ | $r_3$ | $r_4$ |
| PRIOR PAR | | | | | | | |
| $\bar{x}$ | 484 | 491 | 488 | 498 | | | |
| s | 78 | 91 | 91 | 91 | | | |
| r | .502 | .380 | .340 | .317 | .291 | .285 | .304 |
| ATHACT IDX | | | | | | | |
| $\bar{x}$ | 534 | 535 | 537 | 534 | | | |
| s | 104 | 122 | 126 | 126 | | | |
| r | -.069 | -.101 | -.059 | -.067 | -.132 | -.142 | -.116 |
| NATHACT IDX | | | | | | | |
| $\bar{x}$ | 572 | 522 | 506 | 511 | | | |
| s | 118 | 114 | 108 | 105 | | | |
| r | .081 | .066 | .082 | .091 | .087 | .078 | .095 |
| AENGAPT | | | | | | | |
| $\bar{x}$ | 499 | 512 | 518 | 523 | | | |
| s | 80 | 79 | 77 | 77 | | | |
| r | .194 | .246 | .256 | .216 | .284 | .276 | .264 |
| AENGACH | | | | | | | |
| $\bar{x}$ | 471 | 487 | 494 | 496 | | | |
| s | 80 | 77 | 74 | 73 | | | |
| r | .154 | .172 | .152 | .136 | .177 | .182 | .167 |
| AMTHAPT | | | | | | | |
| $\bar{x}$ | 582 | 603 | 611 | 617 | | | |
| s | 67 | 71 | 70 | 68 | | | |
| r | .193 | .297 | .318 | .252 | .248 | .227 | .210 |

$\bar{x}$: mean, s: standard deviation.
r: simple correlation with GPA after first year at Academy.
$r_i$: simple correlation with GPA after $i^{th}$ year at Academy, i=2,3,4.

| Variable | CLASS 76-77 | 71-77 | 71-74 Fresh. | 71-74 Grad. | $r_2$ | $r_3$ | $r_4$ |
|---|---|---|---|---|---|---|---|
| AMTHACH | | | | | | | |
| $\bar{x}$ | 588 | 579 | 590 | 597 | | | |
| s | 65 | 75 | 77 | 76 | | | |
| r | .281 | .340 | .321 | .264 | .257 | .253 | .239 |
| ENGAPT | | | | | | | |
| $\bar{x}$ | 558 | 574 | 580 | 585 | | | |
| s | 70 | 71 | 70 | 69 | | | |
| r | .306 | .340 | .342 | .296 | .367 | .362 | .347 |
| ENGACH | | | | | | | |
| $\bar{x}$ | 596 | 607 | 609 | 613 | | | |
| s | 65 | 66 | 64 | 65 | | | |
| r | .259 | .225 | .192 | .182 | .249 | .234 | .228 |
| MTHAPT | | | | | | | |
| $\bar{x}$ | 655 | 669 | 676 | 681 | | | |
| s | 55 | 60 | 59 | 58 | | | |
| r | .198 | .313 | .326 | .277 | .312 | .280 | .275 |
| MTHACH | | | | | | | |
| $\bar{x}$ | 676 | 705 | 715 | 720 | | | |
| s | 59 | 61 | 58 | 56 | | | |
| r | .336 | .344 | .341 | .288 | .310 | .290 | .294 |
| DENGAPT | | | | | | | |
| $\bar{x}$ | 59 | 62 | 62 | 62 | | | |
| s | 47 | 46 | 46 | 45 | | | |
| r | .126 | .107 | .089 | .087 | .080 | .086 | .084 |
| DENGACH | | | | | | | |
| $\bar{x}$ | 125 | 120 | 115 | 116 | | | |
| s | 56 | 56 | 55 | 56 | | | |
| r | .078 | .025 | .020 | .034 | .057 | .034 | .046 |
| DMTHAPT | | | | | | | |
| $\bar{x}$ | 73 | 66 | 65 | 64 | | | |
| s | 42 | 43 | 44 | 44 | | | |
| r | -.044 | -.053 | -.071 | -.025 | .029 | .018 | .039 |

| Variable | CLASS 76-77 | 71-77 | 71-74 Fresh. | 71-74 Grad. | $r_2$ | $r_3$ | $r_4$ |
|---|---|---|---|---|---|---|---|
| DMTHACH | | | | | | | |
| $\bar{x}$ | 117 | 126 | 125 | 123 | | | |
| s | 44 | 48 | 48 | 47 | | | |
| r | .033 | -.096 | -.101 | -.081 | -.043 | -.043 | -.034 |
| PAE | | | | | | | |
| $\bar{x}$ | 666 | 644 | 634 | 636 | | | |
| s | 94 | 88 | 86 | 90 | | | |
| r | -.037 | -.008 | .026 | .044 | .018 | -.011 | .007 |
| CUMGPA* | | | | | | | |
| $\bar{x}$ | 271 | 275 | 276 | 282 | | | |
| s | 41 | 44 | 45 | 44 | | | |
| r | .699 | .636 | .603 | .541 | .533 | .516 | .510 |
| FCUMGPA* | | | | | | | |
| $\bar{x}$ | 273 | | | | | | |
| s | 37 | | | | | | |
| r | .747 | | | | | | |
| ACADCUM | | | | | | | |
| $\bar{x}$ | 264 | 270 | 271 | 279 | | | |
| s | 49 | 53 | 54 | 53 | | | |
| r | .701 | .633 | .596 | .537 | .536 | .517 | .502 |
| FACADCUM | | | | | | | |
| $\bar{x}$ | 265 | | | | | | |
| s | 44 | | | | | | |
| r | .753 | | | | | | |
| %ACADCUM* | | | | | | | |
| $\bar{x}$ | 810 | | | | | | |
| s | 49 | | | | | | |
| r | .721 | | | | | | |
| MTHGPA | | | | | | | |
| $\bar{x}$ | 268 | 284 | 290 | 300 | | | |
| s | 68 | 69 | 68 | 66 | | | |
| r | .550 | .519 | .490 | .406 | .379 | .351 | .337 |

*All measures of conventional GPA are scaled by a factor of 100.

| Variable | CLASS 76-77 | 71-77 | 71-74 Fresh. | 71-74 Grad. | $r_2$ | $r_3$ | $r_4$ |
|---|---|---|---|---|---|---|---|
| FMTHGPA | | | | | | | |
| $\bar{x}$ | 266 | | | | | | |
| s | 61 | | | | | | |
| r | .602 | | | | | | |
| %MTHGPA* | | | | | | | |
| $\bar{x}$ | 816 | 829 | 834 | 842 | | | |
| s | 65 | 66 | 65 | 62 | | | |
| r | .565 | .522 | .489 | .396 | .369 | .338 | .322 |
| F%MTHGPA** | | | | | | | |
| $\bar{x}$ | 813 | | | | | | |
| s | 60 | | | | | | |
| r | .604 | | | | | | |
| ENGGPA* | | | | | | | |
| $\bar{x}$ | 260 | 255 | 251 | 259 | | | |
| s | 63 | 62 | 62 | 62 | | | |
| r | .505 | .507 | .499 | .475 | .503 | .501 | .491 |
| FENGGPA* | | | | | | | |
| $\bar{x}$ | 264 | | | | | | |
| s | 56 | | | | | | |
| r | .539 | | | | | | |
| %ENGGPA** | | | | | | | |
| $\bar{x}$ | 804 | | | | | | |
| s | 64 | | | | | | |
| r | .535 | | | | | | |
| PTMTGPA* | | | | | | | |
| $\bar{x}$ | 296 | 294 | 294 | 295 | | | |
| s | 43 | 53 | 58 | 60 | | | |
| r | .153 | .085 | .093 | .088 | .062 | .066 | .096 |
| FPTMTGPA* | | | | | | | |
| $\bar{x}$ | 306 | | | | | | |
| s | 40 | | | | | | |
| r | .135 | | | | | | |

*All measures of conventional GPA are scaled by a factor of 100.

**All measures of percentage GPA are scaled by a factor of 10.

# VITA

Kenneth Ray Anderson was born on 19 December 1945 in Gridley, California. He graduated from Newberg High School, Newberg, Oregon, in 1964, attended Oregon State University and then Oregon College of Education where he received the degree of Bachelor of Science in 1969 and a commission in the United States Air Force through the Air Force Reserve Officer Training Corps program. After serving as a mathematics instructor at the United States Air Force Academy Preparatory School for four years he entered the Graduate Systems Analysis program in the Air Force Institute of Technology's Resident School of Engineering.

Permanent Address:

Route 3, Box 668
Dundee, Oregon 97115

This thesis was typed by Mrs. Phyllis Reynolds.