NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

ADMISSIONS AND PLEBE YEAR DATA AS INDICATORS OF ACADEMIC SUCCESS IN ENGINEERING MAJORS AT THE UNITED STATES NAVAL ACADEMY

by

Nicholas A. Kristof

June 2002

Thesis Co-Advisors:	Alice Crawford Roger Little

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2001	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE: Admissions and Plebe Year Data as Indicators of Academic Success in Engineering Majors at the United States Naval Academy			**5. FUNDING NUMBERS**
6. AUTHOR(S) Nicholas A. Kristof			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)*

This research analyzes the relationship between academic success in high school and at the freshman collegiate level and academic performance in engineering majors at the United States Naval Academy (USNA). The study developed predictive models on success and achievement in engineering by examining nine intellective and ten non-intellective variables. The purpose of the project is to contribute to the improvement of academic advising for students considering engineering majors and thus improve student retention. Regression models are estimated for USNA classes of 1997 through 2000 (N=1,648). Three models are estimated to predict completion of an engineering degree, completion of an engineering degree having achieved superior academics, and cumulative quality point rating. Analysis of various explanatory variables shows that a positive relationship exists between early academic success in math and science at the collegiate level and overall success in an engineering major. First semester academic quality point rating was the single most predictive variable in all models.

14. SUBJECT TERMS United States Naval Academy, Academic Achievement, Academic Advising, Academic Persistence, Engineering Education			**15. NUMBER OF PAGES** 123
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

ADMISSIONS AND PLEBE YEAR DATA AS INDICATORS OF ACADEMIC SUCCESS IN ENGINEERING MAJORS AT THE UNITED STATES NAVAL ACADEMY

Nicholas A. Kristof
Lieutenant, United States Navy
B.S., United States Naval Academy, 1996

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
IN
LEADERSHIP AND HUMAN RESOURCE DEVELOPMENT

from the

NAVAL POSTGRADUATE SCHOOL
June 2002

Author: Nicholas A. Kristof

Approved by: Alice Crawford, Thesis Co-Advisor

Roger Little, Thesis Co-Advisor

Douglas A. Brook, Dean
Graduate School of Business and Public Policy

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This research analyzes the relationship between academic success in high school and at the freshman collegiate level and academic performance in engineering majors at the United States Naval Academy (USNA). The study developed predictive models on success and achievement in engineering by examining nine intellective and ten non-intellective variables. The purpose of the project is to contribute to the improvement of academic advising for students considering engineering majors and thus improve student retention. Regression models are estimated for USNA classes of 1997 through 2000 (N=1,648). Three models are estimated to predict completion of an engineering degree, completion of an engineering degree having achieved superior academics, and cumulative quality point rating. Analysis of various explanatory variables shows that a positive relationship exists between early academic success in math and science at the collegiate level and overall success in an engineering major. First semester academic quality point rating was the single most predictive variable in all models.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my wife, Marcianne, for her patience and support throughout the thesis writing process. She has been, as always, an inspiration to me.

I would like to thank my two research advisors: Professor Alice Crawford of the Naval Postgraduate School and Professor Roger Little of the United States Naval Academy. Their time and effort is much appreciated. I would have been unable to complete this endeavor without their support and guidance.

I would like to thank Alan Harmon and the other individuals who work in the Office of Institutional Research, Planning, and Assessment at the Naval Academy. He, Linda Mallory, and Kenneth Sym were extremely helpful throughout the process. They graciously gave of their time when they surely had little to spare.

Finally, I would like to thank my fellow members of Cohort Five of the Leadership Education and Development Program here at the United States Naval Academy. Their presence and support have allowed me to remain sane.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

The mission of the United States Naval Academy is “To develop midshipmen morally, mentally, and physically…” Within that mission is the understanding that the Navy of the 21st century is challenging and technically demanding, and within this understanding is the necessity for the Naval Academy to produce graduates who possess a background in matters technical. The Navy and the Naval Academy have responded to this in various ways over the past twenty-five years. To understand these responses, one must understand the program of academic majors at the Navy Academy.

The Naval Academy offers nineteen academic majors. Six of these are in various fields of engineering; they are known as Division I majors. Seven are mathematics and the sciences or Division II majors. Four are humanities and social sciences; these are the Division III majors. The final two majors are General Engineering and General Science, which are overviews of engineering and science for those individuals who were unable to complete one of the Division I or II majors.[1]

In the mid 1970s, Admiral Hyman Rickover, head of the U. S. Navy’s Nuclear Propulsion Program, pushed for measures requiring 80 percent of all graduates to have a Division I or II major. Admiral Rickover believed that officers with engineering and science backgrounds, i.e., graduates with technical majors, would be better prepared to serve in the Navy. The Naval Academy operated under these guidelines until the mid 1980s when the Secretary of the Navy, the Honorable John Lehman, removed the 80/20 split. His belief was that any graduate from the Naval Academy could succeed anywhere in the Navy, especially in light of the core academic requirements that ensure graduates possess a broad technical background (Ostendorff, personal communication, May, 2002). All midshipmen are required to take three semesters of calculus followed by a semester of differential equations or probabilities and statistics, two semesters of chemistry, two

[1] This listing does not include the six honors majors available to midshipmen. The Naval Academy offers honors programs in Mathematics, Oceanography (Division II majors), and the four Division I majors (Economics, English, History, and Political Science). The requirements for honors majors typically involve increased coursework and a final thesis or research project (“The Majors Program, Class of 2002,” 1999).

semesters of physics, a semester or more of electrical engineering, and a semester or more of weapon systems engineering. All Academy graduates receive a Bachelors of Science degree in their chosen major. The Naval Academy is accredited by the Middle States Association of Colleges and Secondary Schools. All engineering majors, with the exception of General Engineering, are accredited by the Accreditation Board for Engineering and Technology (ABET).

Since the mid 1980s, when the Secretary of the Navy changed the goal of an 80/20 split, the percentage of midshipmen graduating with a technical degree has decreased. The percentage in recent years has been about 60 percent. Almost two-thirds of that number, or 40 percent of a graduating class, have received degrees from Division I majors.

There is a perception held by midshipmen that the Division I majors are more strenuous. This may or may not be the case, and it is not the aim of this study to examine this issue, but there are differences across the majors Divisions. All majors require approximately the same number of total credit hours for graduation. The averages are 143, 141, and 140 total credit hours for Division I, II, and III majors, respectively. The single greatest difference across the majors Divisions is the number of laboratory hours required. This average for Division I majors is almost 19 hours; the average for Division II majors is about 12[2]; and only one Division III major requires laboratory hours at all—Economics, which requires four[3]. These numbers are in addition to those laboratory hours that the core curriculum described above requires.

This perception of difficulty goes hand in hand with a second belief held by midshipmen, which is that one can achieve a higher Academic Quality Point Rating (AQPR) in a non-technical major. This causes a great deal of concern to midshipmen who early in their Academy career are trying to plan their post-graduation military career. Service Selection, which is the process whereby seniors at the Academy choose their military profession, is based largely on academic standing. For midshipmen who feel

[2] This number drops to eight if the Chemistry major, which requires 30 and possibly up to 40 lab hours, is not included in this calculation. Chemistry, a Division II major, does require the most lab hours of any major offered at the United States Naval Academy ("The Majors Program, Class of 2002," 1999).

[3] This does not take into account extra hours required by Honors Division III majors.

strongly about Service Selection, the choice of academic major, which can have a profound effect on academic standing, is often based on misguided pragmatism. It is common belief that by choosing an easier major, perceived by many to be a Division II or III major, a midshipman will have more freedom to choose a military career.

Major selection occurs near the end of plebe (freshman) year. For many, prior academic interest makes this decision an easy one. For others, including those who are attempting to maximize their positive chances at Service Selection or minimize their academic workload, the choice becomes more difficult. These individuals are often influenced by upper class midshipmen with whom they deal on a regular basis or teammates in the athletic arena. They can be influenced to choose the major that will "make their life easier." Almost always, that major is not a Division I major. What is interesting about this turn of events is that the midshipman who chooses a non-engineering major based on peer pressure very possibly had the ability and interest to succeed as an engineer.

It is true that there are midshipmen who are not academically prepared to succeed in an engineering program. That fact in no way weakens their utility to the Naval Service or means that they will be lesser officers. All midshipmen should choose the major that is right for them. However, this decision should not be made solely based on peer pressure and subjective reasoning.

Throughout the freshman year, midshipmen are exposed to the various academic majors available at the Academy. This process starts during plebe summer, the six-week indoctrination program that occurs the summer prior to their freshman year. Plebes are briefed on all academic departments and shown broad overviews of what each has to offer. During the year, most departments hold one or more 'Majors Fairs' which build upon the information presented during the summer and more fully describe the majors. The frequency of these Fairs typically increases as major selection draws near.

Each plebe is also assigned an academic counselor. This counselor is a member of the faculty who can offer guidance to the midshipman and is charged with helping the freshman adapt to the rigors of Academy academics. This counselor also has the responsibility of helping the midshipman choose his major. Help in this case is merely ensuring that the midshipman is making an informed decision based on counselor experience and student preference.

That this counseling is subjective, based on the opinion of academic counselor and counselee, and not based on objective fact lessens its utility. By examining those factors from admissions data and plebe year performance that lead to academic success in Division I majors, this study aims to develop a model to be used during counseling to better enable the plebe to make an informed majors decision.

B. PURPOSE

The overall purpose of this thesis is to research and develop statistical models suitable for use by midshipmen to aid in their major selection. By identifying independent variables that significantly affect academic success in engineering majors, the Naval Academy could offer its midshipmen a tool to aid in this decision. Improving this decision making process will benefit midshipmen, the academic departments, and the Naval Academy.

C. RESEARCH QUESTIONS

1. What admissions and plebe year variables significantly affect graduation with an engineering degree at the United States Naval Academy?

2. What admissions and plebe year variables significantly affect superior academic performance in engineering majors at the United States Naval Academy?

3. What admissions and plebe year variables significantly affect Cumulative Quality Point Rating (CQPR) for engineers at the United States Naval Academy?

4. Does the Personnel History Questionnaire contribute to the prediction of academic success for engineers at the United States Naval Academy?

5. Can a prediction model be devised to assist midshipmen in choosing an academic major?

D. BENEFITS OF STUDY

The aim of this study is to provide midshipmen with a tool that can be used to help choose an academic major. It is hoped that by improving the tools available to midshipmen making this decision, this study will increase the chances that a midshipman will choose the major for which he or she is best suited.

E. SCOPE, LIMITATIONS AND ASSUMPTIONS

1. Scope

The main focus of this study is to determine if it is possible to predict academic success in Division I (engineering) majors based on demographic, admissions, and plebe year data and then determine those factors that best predict this success. Data were collected for midshipmen from the classes of 1997 through 2000. As this study attempts only to examine academic success within engineering majors, only those midshipmen from these year groups who initially chose a Division I major are included.

The scope of this thesis includes statistical analysis of the data collected using logistic and linear regression techniques followed by a discussion of the results. Finally, this thesis will suggest areas for future research as well as offer recommendations to the Naval Academy to improve the majors selection process.

2. Limitations and Assumptions

First, motivation, desire, and the will to achieve all are vitally important to academic success. An individual who appears on paper to be less able than others may outperform them based on his or her level of motivation. Other social or external influences cannot be taken into account. One such influence may be family pressure, which can have a great effect on academic achievement, even at the college level.

Second, this study suffers from the inability to collect all of the data available at the Naval Academy that may play a role in predicting academic success. Various types of data exist at the Academy but were not available due to media difficulties or time constraints. These issues will be addressed in Chapter VI.

F. ORGANIZATION OF STUDY

Following the introduction, Chapter II examines the literature that relates to predicting academic success for undergraduates pursuing engineering majors. Chapter III describes the data set and data coding to be used for all analyses. Chapter IV describes the statistical methodology used in the study. Chapter V presents and interprets the results of the models. Finally, Chapter VI summarizes conclusions, offers recommendations for future research, and suggests ways for the Naval Academy to improve the majors selection process.

II. LITERATURE REVIEW

A. INTRODUCTION

To provide one of the finest technically based undergraduate educations in the country. -- Strategic Plan, United States Naval Academy

The presence of engineering majors is a significant contributor to the technically based education that the United States Naval Academy (USNA) provides its students. The problems encountered by USNA in attracting and retaining midshipmen in engineering majors are common throughout higher education. The attrition rate for college students pursuing engineering majors is historically about 50 percent (Levin and Wychoff, 1987 and 1990; Benefield, Walker, Halpin, Halpin, and Trentham, 1996; Fletcher, Halpin, and Halpin, 1999; MacGuire and Halpin, 1995). In light of this fact, numerous studies have been conducted with the aim of predicting academic success in engineering courses of study (Levin and Wychoff, 1987 and 1990; Benefield et al., 1996; Fletcher et al., 1999; MacGuire and Halpin, 1995; Durio, Kildow, and Slover, 1980; Muchinsky and Hoyt, 1973; and Shoemaker, 1986). The findings of these previous efforts form the basis for this study.

This chapter is divided into two parts. The first section examines the reasons for attempting to predict academic success for engineers. The second section, which makes up the majority of this chapter, examines several studies concerning academic prediction in undergraduate engineering. These studies form the basis for the present study by describing the various forms of data and analyses used by other researchers in tackling this subject.

B. WHY PREDICT ACADEMIC SUCCESS?

The aim of this study is to determine predictors of academic success for engineering majors. The identification of predictors of persistence and success for engineering students is important to the counseling and advising of such students (LeBold, 1958). Predictors can become significant advising tools that invite students to become actively involved in the advising process (Hayden and Halloway, 1985). By improving academic counseling, administrators should be able to reduce the attrition rate for engineers, currently about fifty percent nationally (Fletcher, Halpin, and Halpin, 1999).

Wankat (1986) discussed the need to improve the academic counseling of engineering students to reduce attrition, much of which may be attributed to inappropriate counseling. Levin and Wychoff (1987) write:

> Current educational practices, especially counseling and advising at the post-secondary level, are both inappropriate and inadequate. They are inappropriate because they do not address many of the characteristics of individual students that relate to persistence and success in their intended educational fields. They are inadequate because information on many of the individual student variables that predict both persistence and success in engineering is not available for academic advising purposes. (p. 3)

The assumption operating in this discussion is that programs of study and students who persist in them are somehow matched. The abilities, preparation, and interests of the successful students appear to "fit" the demands of these disciplines (Yess, 1979). This assumption is at the heart of this study. Stated differently, the assumption, described by Levin and Wychoff (1990), more closely links quantitative studies with academic advising:

> An explicit assumption is being made … concerning the usefulness of predictor information in academic advising, i.e. students are more likely to function well academically and make sound educational decisions when they clearly understand how their personal characteristics relate to the likelihood that they will persist successfully in their chosen field of study. By being well informed, students will be better able to choose early in their educational careers, those curricular paths which fit their interests and abilities. (p. 5)

It must be stressed that academic counseling should not solely rely on the results of quantitative studies. Many individual characteristics enter into a student's academic success: personal motivation, prospects of a high-paying and secure job, support from family and friends, and study skills (O'Connor and McAnulty, 1981). Predictors from quantitative research can only tell one part of the story.

In some ways, the problems faced by the United States Naval Academy are different from those that have been faced by other institutions. Shoemaker's (1986) study was initiated to determine appropriate admissions policies for oversubscribed majors, or those in which the number of eligible freshmen exceeds the number that can be enrolled. O'Connor and McAnulty (1981) report on a situation in which the number of applicants for admission to an engineering school and the number of students that can be accommodated are quite different. They speak of a need to grant admission to students based on the likelihood of their success in obtaining an engineering degree. However, whether the situation concerns an oversubscribed major or an undersubscribed major, appropriate counseling can help to solve the problem. If one accepts the assumptions stated above, studies on prediction can give an engineering administrator several different methods with which to tackle a specific problem (Castaneda and Winer, 1985).

C. STUDIES OF ACADEMIC SUCCESS

1. Academic Success in Community Colleges

This review of the literature begins with research at the community college level; the challenges faced by these institutions are not unlike those faced by colleges, universities, and USNA. Research into academic success at a community college may seem diametrically opposed to the study of academic success at USNA. However, similarities exist and must be considered. In fact, the reason for studying academic success in both types of institution is the same. Yess (1979) states:

> The community college "open door" policy, a non-selective admissions policy permitting open registration for courses of study within budgetary limitations, has been of concern to educators who believe that there is a distinction between mere educational access and educational opportunity. These educators argue that in order to enhance educational opportunity for community college students there is a professional obligation on the part of the community college policy-makers to gather appropriate data concerning which factors contribute to the success of community college students. (p. v)

A similar "open door" policy to that mentioned by Yess exists at USNA for, once admitted, a student is given complete latitude in choosing a major. No concrete method exists of trying to place midshipmen in the major for which they are best suited. Axelrod (as reported in Yess, p. 2) "complains that college faculty believe their student bodies are homogeneous" when, in fact, "students in the various major fields of study differ in intellectual characteristics and personality." If one accepts the idea that a specific major is suited to a specific person, then an institution of higher learning should do everything in its power to place students in the appropriate major.

Roueche and Sims (as reported in Yess, p. 7) attack the “open door” policy because it “affords the student an unrestricted choice in selecting a curriculum for which neither he nor the admissions officer knows he is qualified.” The student is given, in effect, the right to fail at anything he or she wants. They call for the community colleges “to assume the major role in determining, at the outset of the college experience, which students qualify for certain programs and how to channel students who would not suit one program into a more beneficial program of study.” By extension, senior colleges, universities, and USNA should do the same.

In his doctoral dissertation, Yess (1979) examined fifteen variables through stepwise linear regression in an attempt to predict cumulative quality point average (QPA) of students at a Massachusetts community college. He developed models to predict success in seven different programs of study. His general finding was that “predictors which consistently accounted for much of the variance in QPA were the intellective variables: high school English average, SAT Verbal and Math” (p. 104). However, those factors that best predicted success were different for each program of study, and the non-intellective variables (such as gender and marital status) did contribute to the success of each predictive model.

Yess (1979) utilized both intellective and non-intellective variables in his study of academic success; Table 1 summarizes his independent variables. Intellective variables include scores on aptitude tests, high school grades, and the types and number of high school English and math courses taken. Data for these types of variables are easily gathered and have been most widely studied (Yess, 1979, p. 19).

Table 1. Independent Variables from Yess (1979)

Variable Name	Variable Description	Measurement Level
Gender	Self-Explanatory	Discrete Variable (1 = Male, 2 = Female)
Age	Self-Explanatory	Continuous Variable
Stopout	Number of semesters the student discontinued enrollment prior to completing Associate's Degree	Continuous Variable
Number of Transcripts	Sum of community college transcripts sent to senior college	Continuous Variable
Marital Status	Self-Explanatory	Discrete Variable (1 = Single, 2 = Married, 3 = Divorced)
Number of Dependents	Self-Explanatory	Continuous Variable
Related Job Experience	Was job experience related to college program of study?	Discrete Variable (1 = Yes, 2 = No)
Related Career Objective	Was career objective related to college program of study?	Discrete Variable (1 = Yes, 2 = No)
Extracurricular Activities	Sum of extra curricular activities in high school	Continuous Variable
Self-Supporting	Was the college student self-supporting?	Discrete Variable (1 = Yes, 2 = No)
Math Scholastic Aptitude Test Score	Self-Explanatory	Continuous Variable (200 to 800)
Verbal Scholastic Aptitude Test Score	Self-Explanatory	Continuous Variable (200 to 800)
Highest Mathematics	Highest Mathematics Level Achieved in High School	Discrete Variable (1 = arithmetic) (2 = algebra I, geometry) (3 = algebra II, trigonometry) (4 = higher than algebra II or trigonometry)
Number of English Courses	Sum of semesters of high school English courses taken	Continuous Variable
English Grade Average	Sum of grades divided by number of English courses taken	Continuous Variable

Yess (1979) conducted forward stepwise linear regression analysis in an attempt to predict graduation quality point rating (QPR) in seven academic programs: Business Administration, transfer students to Business Administration, Engineering Technology, Executive Secretarial, Law Enforcement, Liberal Arts, and Nursing Education. Of interest to this study are his findings in the Engineering Technology program of study. The results of his analysis of Engineering Technology are summarized in Table 2. His finding was that those five variables that most contributed to the prediction of QPR (in order of their contribution to the total R^2) were: 1) Math Scholastic Aptitude Test score, 2) Gender, 3) whether or not the student was Self-Supporting, 4) Age, and 5) Related Job Experience.

Table 2. Summary of Multiple Regression Analysis for the Engineering Technology Program from Yess (1979)

	R^2 Contribution	
Independent Variable	Cumulative	Additional
Math Scholastic Aptitude Test Score	.062	.062
Gender	.136	.073
Self-Supporting	.167	.031
Age	.259	.092
Related Job Experience	.297	.038
English Grade Average	.321	.025
Verbal Scholastic Aptitude Test Score	.335	.014
Highest Mathematics	.347	.012
Number of English Courses	.369	.021
Marital Status	.381	.013
Related Career Objective	.392	.101
Number of Transcripts	.400	.008
Number of Dependents	.403	.003
Extracurricular Activities	.403	.000

Yess (1979) states that his analysis may have been more complete if he had been able to use other non-intellective variables such as biographical information, socio-economic factors, and personality and interest measures. His review of the literature shows that the wider the range of variables available, the greater the predictive power of any model developed. Those studies that combine both intellective with non-intellective variables have the greatest predictive power.

2. Academic Success at the University Level

This review of academic success at the University level begins with a thesis concerning academic success at the Naval Academy. Watson (2001) examined academic achievement at USNA using the Learning and Study Strategies Inventory (LASSI). The LASSI is a 77-question survey using Likert scales that was developed at the University of Texas at Austin in 1978. It was designed as an assessment tool to identify students' academic strengths and weaknesses and has since been tested and validated by over 30 colleges and universities. The LASSI was administered by the Naval Academy Academic Center during plebe summer to help screen midshipmen for academic intervention; it is no longer given at USNA.

Using its 77 questions, the LASSI provides a percentile score to its taker in ten academic areas: Attitude, Motivation, Time Management, Academic Anxiety, Concentration, Information Processing, Main Ideas, Support Techniques, Self Testing, and Test Preparation (Weinstein, Palmer, and Schulte, 1987). These are more fully described in Table 3. Average performance for each category was determined to be between the 50th and 75th percentile. By comparing a midshipman's results with these averages, the Academic Center was able to identify students who may need assistance in one or more specific areas.

Table 3. Description of the LASSI Variables from Watson (2001)

LASSI Variable	Variable Description
Attitude	Attitude and Interest in academic endeavors.
Motivation	Motivation, diligence, self-discipline, and willingness to work hard.
Time Management	Use of time management principles for academic tasks.
Academic Anxiety	Anxiety and worry about school performance.
Concentration	Concentration and attention to academic tasks.
Information Processing	Information processing, acquiring knowledge, and reasoning.
Main Ideas	Selecting main ideas and recognizing important information.

Table 3. Description of the LASSI Variables from Watson (2001) (Cont.)

LASSI Variable	Variable Description
Support Techniques	Use of support techniques and materials.
Self-Testing	Self testing, reviewing, and preparing for classes.
Test Preparation	Test strategies and preparing for tests.

Watson (2001) conducted linear regression analysis to predict Cumulative Quality Point Rating (CQPR) for midshipmen at the end of their freshman year. His independent variables were High School Class Standing, Scholastic Aptitude Test (SAT) Verbal and Math Scores, and the ten LASSI factors. Three of the LASSI variables, Academic Anxiety, Concentration, and Information Processing, were statistically insignificant and dropped from the CQPR estimation. Table 4 summarizes the regression results.

Table 4. Summary of Multiple Regression Analysis from Watson (2001) Prediction of CQPR at End of Plebe Year (N=3,998)

Independent Variable	B	SE B	β	t	Sig.
Constant	-0.06114	0.107	-	-0.57	0.568
High School Class Standing	-0.00898	0.001	-0.211	-14.89	0.000
SAT Verbal Score	0.00127	0.000	0.157	10.69	0.000
SAT Math Score	0.00286	0.000	0.295	19.84	0.000
Attitude	-0.00099	0.000	-0.043	-2.64	0.008
Motivation	0.00392	0.000	0.166	8.61	0.000
Time Management	0.00113	0.000	0.049	2.79	0.005
Main Ideas	-0.00111	0.000	-0.048	-2.64	0.008
Support Techniques	-0.00223	0.000	-0.105	-6.61	0.000
Self Testing	0.00088	0.000	0.039	2.20	0.028
Test Preparation	0.00123	0.000	0.054	2.86	0.004

Note: $R^2 = 0.327$. $F = 196.73$.

The seven LASSI factors included in the estimation were statistically significant, however, Watson (2001) states that due to the low coefficient of regression (R^2) the model could not be used to accurately predict actual CQPR for an individual. The model could, however, "provide an educator with a relative [academic] performance rating based upon initial entry level variables correlated with academic performance" (Watson, 2001, p. 89). This, too, would be useful to an academic counselor.

The remainder of the studies examined in this literature review deal specifically with the academic success of engineering students, which is the primary focus of this study. Studies concerning engineering success published by researchers at Auburn University in Alabama prove helpful in determining both the scope and success of research conducted in the past as well as providing guidance for research in the future. Three studies, two quantitative and one qualitative, are reviewed here. Benefield et al. (1996) examined student retention in engineering majors to identify at-risk students and design intervention strategies to improve their odds of success. Fletcher, Halpin, and Halpin (1999) studied high school and early college grades in an attempt to predict advancement in and graduation from college. MacGuire and Halpin (1995) conducted a qualitative study into those factors that relate to persistence in undergraduate engineering; qualitative research can be very useful in pointing toward other areas of research that should be investigated. These three research efforts are relevant to a study of the current situation at USNA.

Benefield et al. (1996) conducted an assessment of student retention in pre-engineering curricula; they provide useful information in the design of a predictive model for use at USNA. Their study examined data from pre-engineering students from 1991 through 1995 (N=2,505), which included achievement tests, high school transcripts, the Myers-Briggs Personality Type Indicator, the College Student Inventory, Group Embedded Figures Test, College Freshman Survey, college grade reports, and an exit questionnaire. They found a direct correlation between American College Testing (ACT) test scores and grades in specific courses. "For example, the mean ACT composite score for the 29 students who … received an F in Computer Science was 21.8, while the mean score for the 84 students who received an A was 27.6. A similar relationship holds for the mean ACT math score" (Benefield et al., 1996, p. 3). Correlations were found between ACT scores and successful completion of the pre-engineering program. The correlation coefficient for the ACT composite score and successful completion of pre-engineering was 0.34; the coefficient was 0.38 for completion and the ACT math score. Strong correlations were found between ACT composite scores and semester grade point averages (GPAs) for the first two semesters, as well as between ACT math scores and semester GPAs. Attempts to correlate semester GPAs with student Myers-Briggs Type Indicator preferences proved to be unsuccessful.

Benefield et al. (1996) support the use of multiple regression analyses in predicting academic success. They report:

> Multiple regression analysis showed a strong relationship (regression coefficient of .61) between first quarter GPA and ACT math scores, self-reported high school grades, the study habit scale of the CSI, scores on the Group Embedded Figures Test, the highest educational level of the student's father, and the student's self-rating of his/her academic rating.
>
> A similar analysis for the second quarter GPA (regression coefficient of .57) showed the important independent variables to be ACT math scores, self-reported high school grades, scores on the study habit habits scale of the CSI, and the highest educational level of the student's mother. (p. 7)

In each case, the predictive measures included both intellective and non-intellective factors ranging from ACT scores to the students' self-rated academic preparation to parental education level.

At Auburn University, advancement to a major in the engineering program, i.e., advancing to engineering student status from pre-engineering student status, is an early benchmark of success for engineering students. Fletcher, Halpin, and Halpin (1999) conducted One Way Analysis of Variance (ANOVA) in an attempt to identify pre-engineering students who will be successful. They studied freshman in the pre-engineering program at Auburn in the Summer or Fall Quarter of 1991 (N=868). ANOVA calculations provide the statistical significance, stated as an F statistic, for the difference in the means of variables in different populations. The independent variables were high school math index, high school science index, high school humanities index, high school grade point index, and first quarter college grade point average. In the initial analysis, the dependent variable was engineering status; students were defined as advancers, non-advancers with good grades, or non-advancers with poor grades. The results are summarized in Table 5.

Table 5. ANOVA for Engineering Status and Dependent Variables From Fletcher, Halpin, and Halpin (1999)

Source	df	Sum of Squares	Mean Square	F
Total High School Grade Index				
Status	2	60271.99	30135.99	70.87
Error	649	275961.40	425.21	
Total	651	336233.40		
High School Math Index				
Status	2	4943.33	2471.67	60.99
Error	648	26262.17	40.53	
Total	650	31205.50		
High School Science Index				
Status	2	4776.37	2388.19	45.84
Error	649	33809.01	52.09	
Total	651	38585.39		
High School Humanities Index				
Status	2	11387.77	5693.89	42.60
Error	649	87741.14	133.65	
Total	651	98128.91		
First Quarter College Grade Point Average				
Status	2	313.15	156.58	287.18
Error	833	454.17	0.55	
Total	835	767.32		

Note: F statistic is Significant to < 0.001 for all independent variables.

The F statistics from Table 5 show that there is a strong relationship between each of the high school grade indices and engineering status. Total High School Grade Index shows the strongest relationship with engineering status. However, the strongest relationship to engineering status was with First Quarter College Grade Point Average.

In their second analysis, Fletcher, Halpin, and Halpin (1999) examined graduation status; students were defined as engineering graduate, non-engineering graduate, or non-graduate. The results are summarized in Table 6.

The F statistic shows a relationship between each of the high school grade indices and graduation status, however, the relationships are not as strong as that seen between high school grades and engineering status. Again, the strongest relationship was found between First Quarter College Grade Point Average and the dependent variable.

Table 6. ANOVA for Graduation Status and Dependent Variables From Fletcher, Halpin, and Halpin (1999)

Source	df	Sum of Squares	Mean Square	F
		Total High School Grade Index		
Status	2	15225.66	7612.83	15.62
Error	671	327073.50	487.44	
Total	673	342299.20		
		High School Math Index		
Status	2	1420.44	710.22	15.46
Error	670	30783.13	45.95	
Total	672	32203.57		
		High School Science Index		
Status	2	1648.33	824.16	14.62
Error	671	37808.98	56.35	
Total	673	39457.31		
		High School Humanities Index		
Status	2	2475.06	1237.53	8.47
Error	671	97848.80	145.82	
Total	673	100323.90		
		First Quarter College Grade Point Average		
Status	2	163.09	81.54	111.67
Error	862	629.43	0.73	
Total	864	792.52		

Note: F statistic is Significant to < 0.001 for all independent variables.

From the data shown in Tables 5 and 6, Fletcher, Halpin, and Halpin (1999) conclude that the strongest single predictor of engineering success at Auburn University is First Quarter College Grade Point Average. This agrees with the findings of Pascarella, Duby, Miller, and Rasher (1981) who noted that first quarter grade point average made a significant contribution to models predicting eventual engineering success at the collegiate level.

MacGuire and Halpin (1995) conducted a qualitative study in an attempt to determine causal factors related to engineering success. Their goal was to "understand the students' perspectives on which factors impacted their decision to persist or drop out of the pre-engineering program at a major state university" (p. 2). They interviewed 24 students with equal representation between males and females, between African Americans and Caucasians, and between persisters and non-persisters. For the sake of their study, persisters were those students who completed the pre-engineering curriculum and entered an engineering course of study; non-persisters were those students who left the pre-engineering curriculum to pursue another course of study or leave college. The primary theme to emerge from discussions with non-persisters was a sense of naiveté concerning the nature, amount, and difficulty of the pre-engineering coursework.

Several common themes developed across all sub-groups. These included "issues related to the difficulty of the program, to a lack of preparedness, to coping skills, and to a lack of familiarity with the work of an engineer" (MacGuire and Halpin, 1995, p. 13). Difficulty in this case is not limited to the coursework itself, but also to the feelings that are a result of the lower grades received by these students. Preparedness translates to study skills. The "lack of familiarity" deals specifically with the nature of the advanced coursework. The inclusion of this qualitative study is to show that there are factors related to academic success that cannot be determined in a quantitative manner. In order to fully understand academic success, one must utilize both quantitative and qualitative methods.

Researchers from Pennsylvania State University have also examined engineering success at the undergraduate level. Levin and Wychoff (1987) developed predictive models of both academic success and persistence in engineering curricula using five intellective and nine non-intellective variables. They developed models to determine persistence as well as to predict Cumulative Grade Point Average. In total, four predictive models were developed. Two models, based on linear regression analyses, predict cumulative grade point average (CGPA) and engineering grade point average (EGPA), respectively. Two models, based on logistic regression analyses, predict the log odds of being a "Successful Persister" versus being a "Successful Non-Persister" and the log odds of being a "Successful Persister" versus all other statuses, respectively. The determination of persister versus non-persister is based on student status upon completion of the freshman year; these enrollment statuses are summarized in Table 7. Table 8 lists the independent variables used in their analyses. The results of their analyses are briefly described below.

Table 7. Enrollment Statuses from Levin and Wychoff (1987)

Status	College		Cumulative GPA		CGPA
Successful Persisters	Engineering	and	>= 2.00 and grades >= "C" in 3 of 4 engineering courses	and	>= 2.50
Unsuccessful Persisters	Engineering	and	>= 2.00 or grades >= "C" in less than 3 of 4 engineering courses	or	>= 2.50
Successful Non-Persisters	Out of Engineering			and	>= 2.00
Unsuccessful Non-Persisters	Out of Engineering				
Baccalaureate Non-Persisters	Associate Degree, Dropped, Withdrew				

Table 8. Independent Variables from Levin and Wychoff (1987)

Variable Name	Variable Description	Measurement Level	Data Source
High School Grade Point Average	Converted GPA based on academic courses only	Continuous Variable (0.00 to 4.00)	Admissions
Math Scholastic Aptitude Test (SAT) Score	Self-Explanatory	Continuous Variable (200 to 800)	Admissions
Verbal Scholastic Aptitude Test (SAT) Score	Self-Explanatory	Continuous Variable (200 to 800)	Admissions
Algebra Score	Score on Penn State's Math Placement Test	Continuous Variable (0 to 32)	FTCAP
Chemistry Score	Score on Penn State's Chemistry Placement Test	Continuous Variable (0 to 20)	FTCAP
Gender	Self-Explanatory	Dummy Variable Male or Female	Admissions
Attitude toward High School Mathematics	Student's Reaction to High School Mathematics	Dummy Variable Like or Indifferent/Dislike	FTCAP
Attitude toward High School Physics	Student's Reaction to High School Physics	Dummy Variable Like or Indifferent/Dislike	FTCAP
Attitude toward High School Chemistry	Student's Reaction to High School Chemistry	Dummy Variable Like or Indifferent/Dislike	FTCAP

Table 8. Independent Variables from Levin and Wychoff (1987) (Cont.)

Variable Name	Variable Description	Measurement Level	Data Source
Reason for Engineering Choice	Intrinsic (Genuine) versus Extrinsic (Superficial)	Dummy Variable Genuine or Superficial	FTCAP
College Study Hours	Anticipated study hours per week	Continuous Variable (0 to 60)	FTCAP
Non-Science Points	Consistency of major choice	Continuous Variable (0 to 100)	FTCAP
Certainty	Expressed Certainty regarding intended Major	Discrete Variable Very certain, About 50/50, Slightly Uncertain, Uncertain	FTCAP

Note: FTCAP is the Freshman Testing Counseling and Advising Program, which is provided for all entering freshman at Pennsylvania State University. It includes placement examinations and surveys requesting information regarding high school academic experiences, expectations for college, educational plans, and reasons for attending college. All students are also provided an individualized academic advising interview with a professional advisor.
Note: Non-Science Points is a measure of a student's interest in science programs of study. The higher the value, the more interested in science is the student.

Levin and Wychoff's (1987) finalized model to predict CGPA included eight of the fourteen independent variables. Six of the variables were statistically significant; they were (listed in order of contribution to total R^2): 1) High School Grade Point Average, 2) Math SAT Score, 3) Gender, 4) College Study Hours, 5) Algebra Score, and 6) Chemistry Score. The total R^2 for this model was 0.217.

The finalized model to predict EGPA included eleven of the fourteen independent variables. Eight of the variables were statistically significant; they were (listed in order of contribution to total R^2): 1) Algebra Score, 2) High School Grade Point Average, 3) Math SAT Score, 4) Gender, 5) College Study Hours, 6) Non-Science Points, 7) Chemistry Score, and 8) Reason for Engineering Choice. The total R^2 for this model was 0.280.

The logistic regression model that best predicted the logs odds of successfully persisting versus non-persisting in the School of Engineering included seven of the fourteen independent variables. All were statistically significant; they were (listed in order of contribution to the total chi-square): 1) Algebra Score, 2) High School Grade Point Average, 3) Non-Science Points, 4) Chemistry Score, 5) Reason for Engineering Choice, 6) Verbal SAT Score, and 7) Gender.

The logistic regression model that best predicted the logs odds of successfully persisting in the School of Engineering versus all other statuses included seven of the fourteen independent variables. All were statistically significant; they were (listed in order of contribution to the total chi-square): 1) Non-Science Points, 2) Algebra Score, 3) Gender, 4) Chemistry Score, 5) Attitude toward High School Physics, 6) Verbal SAT Score, and 7) Certainty.

In summary, High School Grade Point Average was the first or second most important variable in three of the four models that Levin and Wychoff (1987) estimated. Algebra Score appeared in all four models and was the first or second most important variable in three. Gender was an important variable in all models as was Chemistry Score. Non-Science Points, a measure of interest in matters related to science, was a predictor in three of the models. Math SAT Score was significant in the prediction of GPA while Verbal SAT Score was significant in models of persistence.

In their second study, Levin and Wychoff (1990) increased the number of independent variables in their analyses to nineteen--ten intellective and nine non-intellective variables. They examined not only general academic performance "but also performance in specific courses considered vital for success in engineering" (p. 2). These courses included college calculus, physics, and chemistry grades as independent variables. Using this expanded data set, persistence was analyzed at the end of the sophomore year as opposed to the end of the freshman year as in their prior study.

Three models were developed to estimate persistence at the end of sophomore year. The first model estimated persistence based solely upon the fourteen variables listed in Table 4; these data are available at the start of the freshman year. The second model estimated persistence using the fourteen independent variables in addition to Calculus I, Physics I, and Chemistry I grades; the grades to these courses are typically available upon completion of the freshman year. The final model used all previous independent variables as well as grades for Calculus II and Physics II, typically completed by the end of a student's third semester.

Levin and Wychoff (1990) hypothesize that the variables that are the best predictors of success change over time. The results of the analyses of the three models demonstrate that this may indeed be the case. The logistic regression model using data available at the start of the freshman year that best predicted the log odds of persisting in engineering included six of the fourteen possible independent variables; they are (in order of contribution to the total chi-square): 1) High School Grade Point Average, 2) Algebra Score, 3) Gender, 4) Non-Science Points, 5) Chemistry Score, and 6) Reason for Engineering Choice. The logistic regression model using data available at the end of the freshman year that best predicted the log odds of persisting in engineering included three of the seventeen possible independent variables; they are (in order of contribution to the total chi-square): 1) Physics I grade, 2) Calculus I grade, and 3) Chemistry I grade. The logistic regression model using data available at the middle of the sophomore year that best predicted the log odds of persisting in engineering included three of the nineteen possible independent variables; they are (in order of contribution to the total chi-square): 1) Calculus II grade, 2) Physics II grade, and 3) Physics I grade. From the results of this second study, the conclusion drawn by Levin and Wychoff is that the most recent math and science grades received by a student are the best predictors of engineering persistence.

The strength of Levin and Wychoff's (1987 and 1990) work is its usefulness as a guide upon which to base an analysis of USNA. They design 'hypothetical students' to demonstrate the practical use of their models. By doing so, they show students and counselors alike the relevance of their work. They cite as a possible outgrowth of their work an interactive computer program to assist academic advisors in helping students choose a major. Finally, they offer that "a standard caution which should be observed whenever statistical data are used in advising the individual student is that any individual case may be an exception to even the most compelling statistics. Therefore, such data should always be placed in the context of more complete personal information" (p. 43). These findings are important and relevant to this thesis.

D. SUMMARY

Studies that attempt to analyze academic success in engineering programs of study have been conducted at numerous colleges and universities. A small number of these, representative of the larger body of literature, have been reviewed here. The findings of this representative sample are indicative of the finding of the larger collection. In analyzing engineering success, two types of predictors, intellective and non-intellective, are examined. Intellective variables, typically easy for researchers to gather, include grades, class standings, and test scores. Non-intellective variables are wide ranging and can include biographical information and survey or interview results. Studies have been conducted using a single independent variable, but the most successful cases of academic prediction involve multiple predictors of both types.

These studies typically make use of two types of analyses: logistic regression and linear regression. Logistic analyses are used to predict an outcome, such as success or failure, while linear analyses are used to predict a continuous variable, such as a grade point average. For both types of analyses, hypothetical cases can be developed to illustrate the practical application of regression coefficients.

The studies reviewed here have concentrated on junior and senior college and university students. From an academic standpoint, this study assumes that there is little difference between a university student majoring in engineering and a midshipman at USNA majoring in engineering. The factors that determine success should be the same.

III. DATA AND DATA SET PREPARATION

A. INTRODUCTION

The purpose of this study is to identify variables that significantly affect academic success in engineering majors. By collecting historical data from individual midshipmen, this study attempts to quantify specific variables, both objective and subjective, that may affect academic success for engineering majors. It is hoped that by quantitatively identifying these variables, a model can be developed that predicts academic success. As a result, the Naval Academy could utilize this model to assist plebes in choosing the academic majors for which they are best suited.

B. DATA SOURCES

This study examined only those individuals whose first choice was an engineering major. The initial data set consisted of 1,666 Naval Academy midshipmen (N=1,666) from the classes of 1997 through 2000. Of the data set, 353 cases were missing data in one or more variables. Data cleaning techniques, which are discussed below, were used to fill in the missing data for 58 of those cases. Two-hundred and seventy-seven cases were missing data from the Personal History Questionnaire (described below); they did not take it as part of their admissions process. The actions taken on these cases will be described in the study Methodology chapter. The remaining 18 cases were deleted from the data set, leaving N=1,648. The reason for dropping these eighteen cases is described below.

Data from three different sources were used to create these cases. All data were collated using midshipman alpha code, which is a six-digit number assigned to each midshipman. The first two digits correspond to year of graduation; the final four correspond to an alphabetical listing of the midshipman's last name.

1. Office of Institutional Research, Planning, and Assessment

The Office of Institutional Research, Planning, and Assessment (IR) of the United States Naval Academy supplied admissions, demographic, and performance data for this study. Data obtained from IR included SAT scores, high school rankings, performance information for plebe year calculus and chemistry, Strong Campbell Interest Inventory scores, and gender and race.

2. Personal History Questionnaire

IR also had access to Personal History Questionnaire (PHQ) data for the classes of 1988 through 2000. The PHQ is a survey with questions pertaining to attitudes, personal history, and family; as such, it is one of the only sources of non-intellective data available to this study. Six variables for analysis were drawn from the eighty-five questions of the PHQ. A complete description of the creation of these variables and a copy of the PHQ is attached as Appendix A. These variables included parent's level of education, semesters of advanced high school mathematics courses taken, and personal attributes such as work ethic, academic preparation, and military aptitude.

3. Mathematics and Chemistry Departments

The Mathematics Department of the United States Naval Academy supplied Pre-Calculus Examination data for those individuals who became part of the study. The Mathematics Department offers this test to all midshipmen during their plebe summer in order to place them in the mathematics course for which they are best suited. The test covers geometry, algebra, and basic trigonometry. Plebes who do well on the test are invited to take a Calculus I placement examination to determine their ability in calculus; validation of Calculus I is possible for those who do well on this second test.

The Chemistry Department supplied Toledo Examination data for those individuals who became part of the study. The Chemistry Department offers this test to all midshipmen during their plebe summer in order to place them in the chemistry course for which they are best suited. The Toledo Examination consists of sixty questions, broken into three sections of twenty questions each. The first two sections cover basic math and algebra; the third section covers basic concepts in chemistry. Plebes who do well on the test are invited to take a Chemistry I placement examination to determine their ability in chemistry; validation of Chemistry I is possible for those who do well on this second test.

C. DATA VARIABLES

This section describes the variables obtained from the three sources of data. For those variables that were derivatives of other variables, a description of each derivation is included.

Data were collected and combined into a master database catalogued into cases by midshipmen alpha code. The variables used in the analysis are shown in Table 9. Further discussion of each variable follows Table 9.

Table 9. Description of Variables

Variable Name	Variable Description	Measurement Level	Data Source
Gender – Female (FEMALE)	Self-Explanatory	Dummy Variable (1 = Female) (0 = Male)	IR
Race – Black (BLACK)	Self-Explanatory	Dummy Variable (1 = Black) (0 = Non-Black)	IR
Race - Asian (ASIAN)	Self-Explanatory	Dummy Variable (1 = Asian) (0 = Non-Asian)	IR
Race – Other (OTHERACE)	Self-Explanatory	Dummy Variable (1 = All other Races) (0 = White, Asian, or Black)	IR
High School Ranking (HS_RANK)	Standing in High School. Determined by Admissions	Continuous Variable (400 to 800)	IR
Average Math Scholastic Aptitude Test Score (SATMAVG)	Self-Explanatory	Continuous Variable (200 to 800)	IR
Average Verbal Scholastic Aptitude Test Score (SATVAVG)	Self-Explanatory	Continuous Variable (200 to 800)	IR
Pre-Calculus Examination Score (PRE_CALC)	Score on Naval Academy's Mathematics Placement Examination	Continuous Variable (0 to 100)	Math Department
Toledo Examination Score (TOT_TOL)	Score on Naval Academy's Chemistry Placement Examination	Continuous Variable (0 to 60)	Chemistry Department

Table 9. Description of Variables (Cont.)

Variable Name	Variable Description	Measurement Level	Data Source
Technical Interest Score from SCII (TISSTD)	Interest Score from Strong Campbell Interest Inventory	Continuous Variable (200 to 800)	IR
First Semester Mathematics Performance (MAT1PERF)	Performance Score for 1st Semester Mathematics Course during Plebe Year	Continuous Variable (0 to 20)	IR
First Semester Chemistry Performance (CHM1PERF)	Performance Score for 1st Semester Chemistry Course during Plebe Year	Continuous Variable (0 to 12)	IR
First Semester Academic QPR (SEM1AQPR)	Overall Quality Point Rating after 1st Semester	Continuous Variable (0.00 to 4.00)	IR
Hardwork Score (HARDWORK)	Candidate's Self-Reported Work Ethic	Continuous Variable (9 to 42)	PHQ
Military Aptitude Score (MIL_APT)	Candidate's Self-Reported Military Aptitude	Continuous Variable (3 to 15)	PHQ
Academic Preparation Score (AC_PREP)	Candidate's Self-Reported Level of Academic Preparation	Continuous Variable (4 to 20)	PHQ
Semesters of Advanced or Honors Math Courses Taken in High School (MATH_SEM)	Self-Explanatory	Discrete Variable (1 = 6 or less) (2 = 7 – 8) (3 = 9 – 10) (4 = 11 – 12) (5 = 13 or more)	PHQ
Mother's Education Level (MAGRADED)	Self-Explanatory	Discrete Variable (1 = High School Graduate or less) (2 = Postsecondary School other than College) (3 = Some College) (4 = College Degree) (5 = Some Graduate School or Graduate Degree)	PHQ
Father's Education Level (PAGRADED)	Self-Explanatory	Same as Mother's Education Level	PHQ

Table 9. Description of Variables (Cont.)

Variable Name	Variable Description	Measurement Level	Data Source
Graduate in Engineering (GRAD_ENG)	Dependent Variable. Graduate with an Engineering Degree	Dummy Variable (1 = Graduate with Engineering Degree) (0 = All Else)	IR
Engineering with QPR >= 3.30 (OVER3.30)	Dependent Variable. Graduate with an Engineering Degree and QPR >= 3.30	Dummy Variable (1 = QPR >= 3.30 and GRAD_ENG = 1) (0 = All Else)	IR
Cumulative QPR from Engineering Degree (CQPR_ENG)	Dependent Variable. QPR Upon Graduation with Engineering Degree	Continuous Variable (0.00 to 4.00)	IR

1. Independent Variables

Independent variables, or explanatory variables, are those that help to predict a given dependent variable. The selection of these independent variables was determined from previous research. Basic descriptions of each are below.

FEMALE This is a dummy variable delineating the female midshipmen. The females are coded 1, the males 0. There are 176 females in the set, representing 10.7 percent of the total sample. The literature states that females perform lower than males in engineering, therefore the expected sign for this coefficient in all analyses is negative.

BLACK, ASIAN, OTHERACE These demographic variables are dummy variables derived from raw ethnicity data as described in Table 10.

Table 10. Descriptive Ethnicity Data

Ethnicity	Frequency	Percent	Cumulative Percent
African American	84	5.10	5.10
Asian American	43	2.61	7.71
Caucasian	1381	83.80	91.51
Filipino	29	1.76	93.27
Hispanic	83	5.04	98.31
Native American	15	0.90	99.21
Other	1	0.06	99.27
Puerto Rican	12	0.73	100.00
Total	1648	100.00	

Ethnicity was recoded into the dummy variables BLACK, ASIAN, and OTHERACE for analysis purposes. BLACK midshipmen are those who were listed as African American in admissions data. ASIAN midshipmen are a combination of those listed as Asian and those listed as Filipino. White midshipmen are those listed as Caucasian. OTHERACE includes midshipmen of all other racial types. These groupings were made to improve significance in the statistical analysis of the study as well as to explore the effect that race has on academic success. The breakdown of these demographics is shown in Table 11.

Table 11. Racial Analysis Descriptive Data

Racial Grouping	Frequency	Percent	Cumulative Percent
Black	84	5.10	5.10
Asian	72	4.37	9.47
White	1381	83.80	93.26
OTHERACE	111	6.73	100.00
Total	1648	100.00	

The expected signs for the coefficients for BLACK and OTHERACE are negative. The expected sign for the coefficient for ASIAN is positive.

HS_RANK This variable describes the midshipman's academic standing in high school. It is a standardized variable with a range of 400 to 800 used by the Naval Academy Admissions Board to equate all high school ranking regardless of the size of a high school's graduating class. The expected sign for this coefficient is positive.

SATMAVG and **SATVAVG** These are both numeric variables representing the average score received by the midshipman on the Scholastic Achievement Test (SAT) prior to entering the Naval Academy. The expected sign on the SATMAVG coefficient is positive. The expected sign on the SATVAVG coefficient is unknown.

PRE_CALC This is the score that the midshipman received on the Pre-Calculus Examination. This test is administered by the Mathematics Department and taken by all midshipmen during their Plebe Summer. It is used by the Mathematics Department to determine who is eligible to take a Calculus I placement examination for the purpose of validating the Calculus I course. The expected sign on the coefficient is positive.

TOT_TOL This is the score that the midshipman received on the Toledo Examination. The Toledo Examination is administered by the Chemistry Department and taken by all midshipmen during their plebe summer; it is used as a placement tool in chemistry. The expected sign on the coefficient is positive.

TISSTD This variable is the Technical Interest score from the Strong-Interest Inventory (SII), formerly known as the Strong-Campbell Interest Inventory (SCII), taken as part of the admissions process to the Naval Academy. The Technical Interest score was developed specifically by the Naval Academy. Whereas the SII reports the respondent's scores on four scales (General Occupation, Basic Interest, Occupational, and Personal Style), USNA has developed an alternate scoring of the test that reports the respondent's scores on three different scales: Technical Interest, Humanities Interest, and Career Retention (Sheppard, 2001). The Technical Interest score is used by USNA in its admissions process. The expected sign on this coefficient is positive.

MAT1PERF This is a numeric variable representing the midshipman's performance in his or her first semester mathematics course. The value of this variable was calculated from a combination of the mathematics course taken and the grade received in that course.

Table 12 lists each mathematics course available to a first or second semester plebe during the class years included in the study. The final column of Table 12 is a Difficulty Rating assigned to that course. The Difficulty Rating was determined through consultation with the Mathematics Department; it is a comparative variable used to rank all mathematics courses in order of difficulty. As shown in Table 12, the Difficulty Rating is not meant to imply that SM122 is three times as hard as SM005 or that SM221 is twice as difficult as SM121, but merely rank the different courses by difficulty. The inclusion of this variable introduces a certain amount of measurement error into any models that make use of it. This error was deemed sufficiently low in light of the value of equating numerous math courses with one variable.

Table 12. Mathematics Courses Difficulty Ratings

Course Code	Course Name	Difficulty Rating
SM005	Precalculus Mathematics	1
SM121	Calculus & Analytical Geometry I	2
SM121A	Analytical Geometry, Calc., & Trigonomentry	2
SM121R	Calculus I	2
SM131	Calculus I	2
SM161	Calculus & Computers I	2
SM122	Calculus II	3
SM122A	Calculus II	3
SM122D	Calculus II	3
SM122R	Calculus II	3
SM122S	Calculus II	3
SM162	Calculus & Computers II	3
SM221	Calculus III	4
SM221P	Calculus III	4
SM221S	Calculus III	4
SM251	Calculus & Computers III	4
SM212	Differential Equations	5
SM212M	Differential Equations	5
SM212P	Differential Equations	5

Note: In equations on the following pages, the Difficulty Rating is represented as MAT1DIFF or MAT2DIFF for 1st or 2nd Semester Mathematics Course Difficulty Rating, respectively.

Letter grades (A, B, C, D or F) at the Naval Academy are assigned Quality Point Equivalents (QPE) of 4.0, 3.0, 2.0, 1.0, and 0.0, respectively. Grades in the data set, MAT1GRAD or MAT2GRAD, are their numerical Quality Point Equivalents. MAT1PERF is the product of the Difficulty Rating and the QPE of the plebe's grade in that course and is calculated as follows:

$$\text{MAT1PERF} = \text{MAT1GRAD} * \text{MAT1DIFF} \tag{3-1}$$

For example, consider Midshipman X who received an A in SM131:

MAT1PERF (Mid X) = 4 (the QPE of an A) * 2 (MAT1DIFF of SM131) = 8.

Consider also Midshipman Y who received a C in SM212:

MAT1PERF (Mid Y) = 2 (the QPE of a C) * 5 (MAT1DIFF of SM212) = 10.

The expected sign on the coefficient of MAT1PERF is positive.

CHM1PERF This is a numeric variable representing the midshipman's performance in his or her first semester chemistry course. The value of this variable was calculated from a combination of the chemistry course taken and the grade received in that course.

Table 13 lists each chemistry course available to a first or second semester plebe during the class years included in the study. The final column of Table 13 is a Difficulty Rating assigned to that course. The Difficulty Rating was determined through consultation with the Chemistry Department; it is a comparative variable used to rank all chemistry courses in order of difficulty. As shown in Table 12, the Difficulty Rating is not meant to imply that SC151 is three times as hard as SY100 or that SC111 is twice as difficult as SY100, but merely rank the courses by difficulty. The inclusion of this variable introduces a certain amount of measurement error into any models that make use of it. This error was deemed sufficiently low in light of the value of equating numerous chemistry courses with one variable.

Table 13. Chemistry Courses Difficulty Ratings

Course Code	Course Name	Difficulty Rating
SY100	Fundamentals of Science	1
SC111	Foundations of Chemistry I	2
SC111Y	Elements of Chemistry I	2
SC112	Elements of Chemistry II	3
SC122	Chemistry II	3
SC151	Modern Chemistry	3

Note: In equations below, the Difficulty Rating is represented as CHM1DIFF or CHM2DIFF for 1st or 2nd Semester Chemistry Course Difficulty Rating, respectively.

Grades in the data set, CHM1GRAD or CHM2GRAD, are their numerical Quality Point Equivalents. CHM1PERF is the product of the Difficulty Rating and the QPE of the plebe's grade in that course and is calculated as follows:

$$\text{CHM1PERF} = \text{CHM1GRAD} * \text{CHM1DIFF} \tag{3-2}$$

Calculation of CHM1PERF follows the examples of MAT1PERF above. The expected sign on the coefficient of CHM1PERF is positive.

SEM1AQPR This variable represents the Academic Quality Point Rating earned by the midshipman during the first semester of his or her plebe year. The AQPR is the semester average of all grades received by that midshipman during that semester; it is computed by multiplying the QPE corresponding to the letter grade received in each course by the semester hours of credit for the course and dividing the sum of these products by the total number of semester hours represented by all of the courses taken. The expected sign on this coefficient is positive.

HARDWORK This is a numeric variable representing a midshipman's Work Ethic as self-reported on the Personal History Questionnaire. See Appendix A for the determination of this value. The expected sign on the coefficient is positive.

MIL_APT This is a numeric variable representing a midshipman's level of Military Aptitude as self-reported on the Personal History Questionnaire. See Appendix A for the determination of this value. The expected sign on the coefficient is positive.

AC_PREP This is a numeric variable representing a midshipman's level of Academic Preparation as self-reported on the Personal History Questionnaire. See

Appendix A for the determination of this value. The expected sign on the coefficient is positive.

MATHSEM This is a numeric variable representing a midshipman's level of high school mathematics achievement as self-reported on the Personal History Questionnaire. The expected sign on this coefficient is positive.

MAGRADED and **PAGRADED** These are numeric variables representing the level of education received by the midshipman's mother and father, respectively, as reported by the candidate on the Personal History Questionnaire. The expected sign on these coefficients is positive.

2. Dependent Variables

The dependent variables in the study were chosen to reflect 'Academic Success.' Three different measures of success are defined for the purposes of this study. They are described here:

GRAD_ENG This dependent variable will be analyzed using logistic regression; it describes one level of academic success. The most basic level of success as an engineer is to graduate with a degree in an engineering major. The dependent variable GRAD_ENG takes on a value of zero for those individuals who initially chose an engineering major and either 1) failed to complete the course of studies or 2) switched to a Division II or III (non-engineering) major sometime prior to graduation. The variable GRAD_ENG takes on a value of one for those individuals who initially chose an engineering major and graduated with an engineering degree. For the sake of the analysis, a degree in General Engineering, although not an accredited engineering major, is considered an engineering degree. The data indicated that 82.0 percent of those who initially chose an engineering major graduated with an engineering degree while 18.0 percent left the Naval Academy or graduated with a Division II or III degree.

OVER3.30 This dependent variable will also be analyzed using logistic regression; it further quantifies academic success. The second level of success to be analyzed is that which would allow a Naval Academy graduate to enter a top-tier graduate program in engineering. Twenty members of the engineering faculty were surveyed to determine their opinion of the level of undergraduate achievement in terms of

Cumulative Quality Point Rating (CQPR) necessary to obtain entrance into a top-tier engineering graduate program. The mean of their replies was a CQPR of 3.30. The dependent variable OVER3.30 takes on a value of one for those individuals whose GRAD_ENG = 1 and who graduated with greater than or equal to a 3.30 CQPR. It takes on a value of zero otherwise. The data indicate that 24.0 percent of those who initially chose an engineering major graduated with an engineering degree and a CQPR of greater than 3.30 while 76.0 percent did not.

CQPR_ENG This dependent variable will be analyzed using linear regression techniques. CQPR_ENG is the midshipman's CQPR upon the completion of an engineering degree at the Naval Academy; it is computed by multiplying the QPE corresponding to the letter grade received in each academic course by the semester hours of credit for that course and dividing the sum of these products by the total number of semester hours represented by all of the courses taken by the midshipman during his four years at the Academy.

In order to graduate, a midshipman must successfully complete or validate a minimum of 140 semester hours, including a minimum of 90 semester hours in the core program, with a cumulative CQPR of at least 2.00. As stated above, only 82.0 percent of the initial sample actually graduated with an engineering degree, therefore the linear regression analysis will only cover 1,351 cases of the total.

D. DATA CLEANING

As described above, 58 of the 1,648 cases (not including those 277 cases which were missing all PHQ data) were missing one or more variables. Two distinct methods were employed to alleviate this problem. The first method, Performance Extrapolation, was used to supply data for those cases missing a value for MAT1PERF or CHM1PERF. The second method, Mean Insertion, was used for all others.

1. Performance Extrapolation

This method was used to insert data for MAT1PERF or CHM1PERF where none was available. Twenty cases required data insertion. In all of these cases, data insertion was necessary, because that individual did not take a chemistry or mathematics course in his or her first semester at the Academy. The individuals in question were able to validate a semester or more of chemistry or calculus and opted to not take a chemistry or calculus course during their first semester.

For these cases, a 2nd Semester Mathematics or Chemistry Performance score (MAT2PERF or CHM2PERF) was calculated using the available data. The following paragraphs describe the method by which a first semester performance score (MAT1PERF or CHM1PERF, respectively) was derived from a second semester performance score (MAT2PERF or CHM2PERF, respectively). It should be noted that those eighteen cases that were dropped from the initial data set were dropped due to the fact that they possessed no first or second semester calculus or chemistry data thereby eliminating the possibility of generating a MAT1PERF or CHM1PERF value.

It is assumed that from one semester to the next, the Difficulty Rating of the course taken by a midshipman in a particular subject would increase by one, i.e., the Difficulty Rating of a midshipman moving from a Calculus I course to a Calculus II course would jump from 2 to 3. This assumption holds true for all midshipmen with the exception of those who failed the first semester course and had to repeat it. Conversely, it can be assumed that from a later semester to an earlier one, the Difficulty Rating would decrease by one. For example, a midshipman who validated Calculus I and took Calculus II during the second semester, which has a Difficulty Rating of 3, would have taken a course *had he or she not validated Calculus I*, with a Difficulty Rating of 2 during the first semester. It is further assumed that a midshipman would receive *at least* a comparable grade in a lower Difficulty Rating course compared to the higher Difficulty Rating course that was taken.

Using the above assumptions, the equations for MAT1PERF and CHM1PERF are listed below:

$$\text{MAT1PERF(Missing)} = \text{MAT2GRAD} * (\text{MAT2DIFF} - 1) \tag{3-3}$$

$$\text{CHM1PERF(Missing)} = \text{CHM2GRAD} * (\text{CHM2DIFF} - 1) \tag{3-4}$$

For example, consider the midshipman who validated Calculus I and II and chose to take Calculus III during the second semester and received a B. This midshipman's MAT1PERF is calculated below:

MAT1PERF(Missing) = 3 (the QPE of an A) * (4 – 1) = 9.

This value is then inserted into the data set as the MAT1PERF.

2. Mean Insertion

This method was used to insert data for the remaining thirty-eight cases. In each of these cases, data were missing for up to three, but not all, PHQ variables. To correct for these cases, the means of all variables were computed by gender and race. The mean value for each variable was then inserted into the missing case in accordance with the appropriate gender and race.

Appendix B contains tables that summarize, by gender and race, the mean and other descriptive statistics for each variable prior to mean insertion.

E. CHAPTER SUMMARY

Data from three sources were collated into a master data set that was cleaned and readied for analysis (N=1,648). Each case contains nineteen independent variables and three dependent variables. Table 14 provides the Descriptive Statistics for the resultant data set. This data set was used in all analyses.

Table 14. Descriptive Statistics for Analysis Data Set

	N	Minimum	Maximum	Mean	Std Dev
FEMALE	1648	0	1	0.107	0.3089
BLACK	1648	0	1	0.051	0.2200
ASIAN	1648	0	1	0.044	0.2045
OTHERACE	1648	0	1	0.067	0.2507
HS_RANK	1648	400	800	587.068	105.8917
SATMAVG	1648	415	800	651.292	61.9507
SATVAVG	1648	200	790	559.397	81.2914
PRE_CALC	1648	28	100	80.025	11.5281
TOT_TOL	1648	23	60	47.273	5.9016
TISSTD	1648	262	747	529.078	86.4892
MAT1PERF	1648	0	20	7.061	2.8237
CHM1PERF	1648	0	12	5.400	2.1419
SEM1AQPR	1648	0.00	4.00	2.842	0.5708
HARDWORK	1371	26	42	36.689	2.8154
MIL_APT	1371	3	15	12.130	2.4244
AC_PREP	1371	6	20	16.109	2.4885
MATH_SEM	1371	1	5	3.340	1.2992
MAGRADED	1371	1	5	3.509	1.3415
PAGRADED	1371	1	5	3.817	1.2886
GRAD_ENG	1648	0	1	0.820	0.3845
OVER3.30	1648	0	1	0.240	0.4270
CQPR_ENG	1351	2.03	4.00	3.005	0.4733
Valid N (listwise)	1351				

Note: Valid N (listwise) is listed for the CQPR_ENG analysis. The Valid N (listwise) for the GRAD_ENG and OVER3.30 analyses is 1,648 for those analyses that do not take into account PHQ data. The Valid N (listwise) for the GRAD_ENG and OVER3.30 analyses is 1,371 for those analyses that do take into account PHQ data.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. METHODOLOGY

A. INTRODUCTION

The primary aim of this study is to identify those variables that significantly affect academic success for midshipmen in engineering majors. A sample of 1,648 midshipmen from the classes of 1997 through 2000 form the data set for this analysis. Because this study is focused on academic success for engineers, only those midshipmen who initially chose an engineering major were included in the study.

Fifteen intellective and non-intellective variables in addition to four demographic variables were utilized in an attempt to predict academic success for engineering students. Two measures of success have been developed: (1) Graduation with an engineering degree (GRAD_ENG), and (2) Graduation with a CQPR of 3.30 or higher and an engineering degree (OVER3.30). Logistic regression models that attempt to predict these measures of success were estimated and tested.

The independent variables were also used to attempt to predict a midshipman's level of academic achievement, as measured by that midshipman's CQPR (CQPR_ENG). Linear regression models that attempt to predict academic achievement separately from the above measures of academic success were estimated and tested.

A secondary aim of this study is to examine the utility of the Personal History Questionnaire (PHQ) in the prediction of academic success and achievement. This will be accomplished via simultaneous analysis of each of the dependent variables with and without PHQ data. In this way, the marginal effects of the PHQ data on the models can be determined.

B. LOGISTIC REGRESSION MODELS

The dependent variables, GRAD_ENG and OVER3.30 will be analyzed using logistic regression techniques. Logistic regression is used because it predicts discrete outcomes from continuous, discrete, and dichotomous variables. The outcome variable of a logistic regression, Y, is the probability that an individual will be a member of a given group versus not being a member of that group based on a nonlinear function of the best linear combination of predictors:

$$Y_i = e^u / (1 + e^u) = 1 / (1 + e^{-u}) \qquad (4 – 1)$$

Y_i is the estimated probability that the *i*th individual will be a member of a given group, *e* is the base of the natural logarithm, and *u* is the linear regression equation:

$$u = A + B_1X_1 + B_2X_2 + \ldots + B_kX_k \qquad (4 – 2)$$

with constant *A*, logistic coefficients B_j, and predictors X_j for k predictors (j = 1, 2, …, k).

The linear regression equation creates the log of the odds:

$$\ln (Y / (1 – Y)) = u = A + \Sigma B_jX_{ij} \qquad (4 – 3)$$

The linear regression equation becomes the natural log of the probability of being a member of one group divided by the probability of not being in the group. For instance, the logistic regression of GRAD_ENG will yield the log odds that an individual will graduate with an engineering degree as opposed to not graduating with an engineering degree.

Equation (4 – 1), rewritten, reads: $P(\text{outcome})_i = f (X_{ij}) + \text{Constant} \qquad (4 – 4)$

$P(\text{outcome})_i$ is the probability of a given outcome for the *i*th individual. $f (X_{ij})$ is the measure of the *j*th independent variable, X, for the *i*th individual.

Once the logistic regression models are finalized, the decimal probability of each outcome for the case where each independent variable has its mean as its value will be presented. At the same time, the marginal effects of each independent variable on the outcome will be calculated and discussed. The marginal effect of an independent variable is calculated:

$$\text{Marginal Effect} = B_j * P(Y=1) * P(Y=0) \qquad (4 – 5)$$

B_j is the logistic coefficient, *P(Y=1)* is the probability that a given outcome will occur, and *P(Y=0)* is the probability that a given outcome will not occur. The marginal effect reports the change in the probability of the outcome per unit of change of each independent variable.

1. Models including PHQ Data

a. GRAD_ENG

Following Equation (4 – 4), the model specification that describes the determinants of a midshipman graduating with an engineering degree reads:

$$P\ (GRAD\_ENG\ (w/PHQ))_i = f\ (FEMALE_i, BLACK_i, ASIAN_i, OTHERACE_i, HS\_RANK_i, SATMAVG_i, SATVAVG_i, TISSTD_i, SEM1AQPR_i, MAT1PERF_i, CHM1PERF_i, PRE\_CALC_i, TOT\_TOL_i, AC\_PREP_i, HARDWORK_i, MIL\_APT_i, MATH\_SEM_i, MAGRADED_i, PAGRADED_i) + Constant \qquad (4 – 6)$$

$P\ (GRAD\_ENG\ (w/PHQ))_i$ is the probability that the *i*th individual will graduate with an engineering degree while taking into account PHQ data. (N=1,371)

b. OVER3.30

Following Equation (4 – 4), the model specification that describes the determinants of a midshipman graduating with an engineering degree and a CQPR of 3.30 or greater reads:

$$P\ (OVER3.30\ (w/PHQ))_i = f\ (FEMALE_i, BLACK_i, ASIAN_i, OTHERACE_i, HS\_RANK_i, SATMAVG_i, SATVAVG_i, TISSTD_i, SEM1AQPR_i, MAT1PERF_i, CHM1PERF_i, PRE\_CALC_i, TOT\_TOL_i, AC\_PREP_i, HARDWORK_i, MIL\_APT_i, MATH\_SEM_i, MAGRADED_i, PAGRADED_i) + Constant \qquad (4 – 7)$$

$P\ (OVER3.30\ (w/PHQ))_i$ is the probability that the *i*th individual will graduate with an engineering degree and a CQPR of 3.30 or greater while taking into account PHQ data. (N=1,371)

2. Models excluding PHQ Data

a. GRAD_ENG

In order to test the marginal effects that the PHQ data have on each model, the models will be analyzed a second time while excluding the PHQ data from the analysis. Following Equation (4 – 4), the model specification that describes the determinants of a midshipman graduating with an engineering degree then reads:

P (GRAD_ENG (w/o PHQ))$_i$ = f (FEMALE$_i$, BLACK$_i$, ASIAN$_i$, OTHERACE$_i$, HS_RANK$_i$, SATMAVG$_i$, SATVAVG$_i$, TISSTD$_i$, SEM1AQPR$_i$, MAT1PERF$_i$, CHM1PERF$_i$, PRE_CALC$_i$, TOT_TOL$_i$) + Constant (4 – 8)

P (GRAD_ENG (w/o PHQ))$_i$ is the probability that the *i*th individual will graduate with an engineering degree while excluding PHQ data. (N=1,648)

b. OVER3.30

Following Equation (4 – 4), the model specification that describes the determinants of a midshipman graduating with an engineering degree and a CQPR of 3.30 or greater reads:

P (OVER3.30 (w/o PHQ))$_i$ = f (FEMALE$_i$, BLACK$_i$, ASIAN$_i$, OTHERACE$_i$, HS_RANK$_i$, SATMAVG$_i$, SATVAVG$_i$, TISSTD$_i$, SEM1AQPR$_i$, MAT1PERF$_i$, CHM1PERF$_i$, PRE_CALC$_i$, TOT_TOL$_i$) + Constant (4 – 9)

P (OVER3.30 (w/o PHQ))$_i$ is the probability that the *i*th individual will graduate with an engineering degree and a CQPR of 3.30 or greater while excluding PHQ data. (N=1,648)

3. Effects of Personal History Questionnaire Data

The total sample includes 1,648 cases. Of those, 277 cases have no PHQ data. This explains the difference in sample size between those analyses that take into account PHQ data and those that exclude PHQ data. A comparison of these models will allow the effects of the PHQ data to be examined.

4. Finalized Logistic Model Specification

For each of the four models specified above, the following methodology will be used to arrive at the final specification:

1. The initial model with nineteen or thirteen variables, for PHQ or non-PHQ, respectively, will be estimated and analyzed.
2. Those variables whose statistical significance is greater than 0.3 will be removed from the model. The demographic variables (FEMALE, BLACK, ASIAN, and OTHERACE) will not be removed.

3. The analysis will be repeated as necessary, at each step eliminating those variables whose significance is greater than 0.3 until no variables meet this criterion.

4. The analysis will then be conducted and those variables whose significance is greater than 0.15 will be removed. This will be repeated until there are no variables, with the exception of the four demographic variables, whose significance is greater than 0.15.

5. Each model will be analyzed using STEPWISE variable entry with the SPSS 10.0 Analysis software. Stepwise variable entry allows the determination of the order of importance of each variable to the total model.

6. The logistic coefficients will then be used to present a decimal probability for each outcome variable. At the same time, the marginal effects of the independent variables on the outcome will be discussed.

C. LINEAR REGRESSION MODELS

The dependent variable, CQPR_ENG, will be analyzed using linear regression techniques. The outcome of such a regression will be an equation of coefficients that can be used to directly predict CQPR_ENG.

1. CQPR_ENG with PHQ Data

The model specification used to describe the determinants of a midshipman's academic QPR upon graduating with an engineering degree takes the following form:

$$\text{CQPR\_ENG (w/PHQ)}_i = \beta_0 + \beta_1\text{FEMALE}_i + \beta_2\text{BLACK}_i + \beta_3\text{ASIAN}_i + \beta_4\text{OTHERACE}_i + \beta_5\text{HS\_RANK}_i + \beta_6\text{SATMAVG}_i + \beta_7\text{SATVAVG}_i + \beta_8\text{TISSTD}_i + \beta_9\text{SEM1AQPR}_i + \beta_{10}\text{MAT1PERF}_i + \beta_{11}\text{CHM1PERF}_i + \beta_{12}\text{PRE\_CALC}_i + \beta_{13}\text{TOT\_TOL}_i + \beta_{14}\text{AC\_PREP}_i + \beta_{15}\text{HARDWORK}_i + \beta_{16}\text{MIL\_APT}_i + \beta_{17}\text{MATH\_SEM}_i + \beta_{18}\text{MAGRADED}_i + \beta_{19}\text{PAGRADED}_i \qquad (4 – 10)$$

CQPR_ENG (w/PHQ))$_i$ is the cumulative QPR at graduation with an engineering degree for the *i*th individual while taking into account PHQ data. (N=1,371)

2. CQPR_ENG without PHQ Data

In order to test the marginal effects that the PHQ data have on CQPR_ENG, it will be analyzed a second time excluding the PHQ data from the analysis. The model specification used to describe the determinants of a midshipman's academic QPR upon graduating with an engineering degree then takes the following form:

$$\text{CQPR\_ENG (w/o PHQ)}_i = \beta_0 + \beta_1\text{FEMALE}_i + \beta_2\text{BLACK}_i + \beta_3\text{ASIAN}_i + \beta_4\text{OTHERACE}_i + \beta_5\text{HS\_RANK}_i + \beta_6\text{SATMAVG}_i + \beta_7\text{SATVAVG}_i + \beta_8\text{TISSTD}_i + \beta_9\text{SEM1AQPR}_i + \beta_{10}\text{MAT1PERF}_i + \beta_{11}\text{CHM1PERF}_i + \beta_{12}\text{PRE\_CALC}_i + \beta_{13}\text{TOT\_TOL}_i \quad (4-11)$$

$\text{CQPR\_ENG (w/o PHQ)}_i$ is the cumulative QPR at graduation with an engineering degree for the *i*th individual while excluding PHQ data. (N=1,648)

As described above for the logistic regression analyses, the difference in sample size is due to the inclusion or exclusion of those cases that possess PHQ data. A comparison of these models allows the effects of the PHQ data to be explored.

3. Finalized Linear Model Specification

For the linear regression analyses, the methodology will follow that outlined above for the logistic analyses with two exceptions. Decimal probability is not applicable to linear regression and will not be determined. The concept of marginal effects of independent variables is not applicable to the linear model and will not be discussed.

D. CHAPTER SUMMARY

Six models were developed to analyze the three dependent variables first including and then excluding the Personal History Questionnaire data. Each model was estimated based on the methodology described above so that only those variables that are statistically significant (Sig. < 0.15) together with the demographic variables remain. The variables in each model are then ranked according to their individual importance to each model. The results of these analyses follow in Chapter V.

V. RESULTS

A. INTRODUCTION

Four models have been estimated to predict determinants of academic success for engineering majors at the United States Naval Academy. Two models have been estimated to predict determinants of academic achievement. Both logistic and linear regression analyses were conducted using the SPSS 10.0 software. The results of those analyses are described in this chapter.

For all analyses, the data set contains only those midshipmen who initially chose an engineering major. It must be kept in mind that the results of these analyses would be different if the data set included all midshipmen in a given year group.

1. Statistical Comparison between Models

In reviewing the results of these analyses, it is important to establish significance level thresholds with respect to the various independent variables. This allows determination of the importance of an independent variable to the model. Table 15 details coefficient significance levels for use with this study.

Table 15. Range of Significance for Regression Coefficients

Range of Significance	Predictive Value
0.000 – 0.049	Highly Significant
0.050 – 0.099	Significant
0.100 – 0.149	Marginally Significant
0.150 and higher	Not Significant

It is equally important to determine guidelines for comparison between models. One such metric for comparison is the chi-square test, which tests for independence between two discrete variables. In chi-square analysis, the null hypothesis generates expected frequencies of variables based on a random distribution. The expected frequencies are compared against observed frequencies. If the observed frequencies closely match expected frequencies, the chi-square is low, and the two variables are independent, i.e. not related to each other. If observed and expected frequencies do not match, the chi-square is high, and one concludes that the two variables are related. The chi-square allows comparison between two or more logistic regression models.

To compare linear regression models, other metrics are used. These include the coefficient of determination (R^2) and the F statistic. R^2 is a goodness of fit statistic that describes how well the model equation fits the sample. The value of R^2 is always positive and ranges from zero to one. A value of one indicates that the equation perfectly fits the sample data, whereas a value of zero indicates that the equation does not fit the sample data at all. The F statistic is a measure of the overall significance of the equation. It indicates the predictive power of an equation. The higher the value of the F statistic, the more predictive the equation.

2. Chapter Organization

This chapter consists of six sections. Following this introduction, the results of the analysis for each dependent variable are described. Then, two notional midshipmen are described to demonstrate the utility of the regression models. Finally, the chapter is briefly summarized.

B. GRADUATION WITH AN ENGINEERING MAJOR

In this study, the initial criterion for academic success in the field of engineering is to graduate with an engineering degree. The dependent variable GRAD_ENG has been used to model this level of success. Two models of GRAD_ENG were developed—the first included Personal History Questionnaire data whereas the second did not. As described in Chapter IV, these models are the product of an iterative process designed to maximize the predictive capability of each model while at the same time maximizing the statistical significance of the individual variables included therein. A comparison of the finalized model for each shows that PHQ data do not contribute to the prediction of GRAD_ENG and actually inhibits its prediction. Therefore, only the model excluding PHQ data is presented. Appendix C fully describes the iteration process for GRAD_ENG and more fully compares the two models.

1. Final Logistic Regression of GRAD_ENG (Excluding PHQ Data)

The final regression model for the dependent variable GRAD_ENG included eleven of the thirteen independent variables initially considered for inclusion. There were nine significant predictor variables. These variables were (listed in order from largest to smallest contribution to the total chi-square): 1) First Semester Academic QPR (SEM1AQPR), 2) Pre-Calculus Examination Score (PRE_CALC), 3) Strong Campbell Interest Inventory Technical Interest Score (TISSTD), 4) First Semester Mathematics Performance (MAT1PERF), 5) Gender - Female (FEMALE), 6) Average Verbal Scholastic Aptitude Test Score (SATVAVG), 7) Average Math Scholastic Aptitude Test Score (SATMAVG), 8) Race - Other (OTHERACE), and 9) Race – Asian (ASIAN). First Semester Chemistry Performance (CHM1PERF) and Race – Black (BLACK) are included in the final model but are not highly significant.

These findings bear further discussion. The single greatest predictor of whether or not a midshipman will graduate with an engineering degree is how that midshipman performs in the classroom during the first semester at the Naval Academy. Typically, this is the semester that the plebe is struggling most to adapt to the military lifestyle and deal with plebe responsibilities. The second best predictor is performance on the Pre-Calculus placement exam, which is given within the first week of a plebe reporting to the Academy. This is closely followed in importance by the Technical Interest score from the Strong Campbell Interest Inventory, a test taken during the admissions process. The fourth most powerful predictor is a plebe's performance in the first semester mathematics course; this logically supports the above findings for two reasons. First, the calculus grade contributes to the total semester AQPR, which is the single greatest predictor. Secondly, the mathematics course taken by a plebe is directly contingent on the plebe's performance on the Pre-Calculus examination. Gender, race, and performance on the SAT test round out the significant predictors.

Some interesting results were encountered with respect to signs of the variable coefficients. It was hypothesized that all intellective variables would have positive coefficients, with the possible exception of SATVAVG. In fact, the coefficient for SATVAVG is negative, which implies that the higher score one received on the Verbal section of the Scholastic Aptitude Test the lower the chances that the individual would graduate with an engineering major. It was hypothesized that all demographic variables, which denote membership in a minority group, would have negative coefficients, with the possible exception of ASIAN. In fact, all demographic variables including ASIAN had negative coefficients.

Table 16 summarizes the results from the Finalized Logistic Model. The independent variables are listed in order of their contribution to the model chi-square and R^2.

Table 16. Finalized Logistic Regression Model for Graduating with an Engineering Degree

			Chi-Square Contribution		R^2 Contribution	
Independent Variable	B	Sig	Additional	Cumulative	Additional	Cumulative
SEM1AQPR	0.634	0.014	155.373	155.373	0.1473	0.1473
PRE_CALC	0.024	0.003	50.755	206.129	0.0452	0.1925
TISSTD	0.004	0.000	42.928	249.057	0.0372	0.2297
MAT1PERF	0.151	0.002	14.965	264.021	0.0127	0.2424
FEMALE	-0.612	0.003	10.600	274.621	0.0089	0.2513
SATVAVG	-0.004	0.001	4.947	279.568	0.0042	0.2555
SATMAVG	0.004	0.010	9.933	289.501	0.0083	0.2638
OTHERACE	-0.527	0.038	3.503	293.004	0.0029	0.2667
ASIAN	-0.666	0.029	3.813	296.817	0.0032	0.2699
CHM1PERF	0.109	0.079	3.221	300.038	0.0027	0.2725
BLACK	-0.121	0.675	0.174	300.212	0.0001	0.2727
Constant	-6.310	0.000				
-2 Log Likelihood	1254.592					
Model Significance	0.000					

Note: Value given for R^2 is the Nagelkerke R^2.
Note: This model excludes PHQ data. (N=1,648)

From Table 16, the chi-square and the -2 log likelihood were 300.212 and 1254.592 for this model, respectively. Both demonstrate a high goodness-of-fit for the model and show strong statistical reliability. As can be seen in Appendix C, both statistics are higher than for those of the GRAD_ENG model including PHQ data, which shows that PHQ data inhibits the prediction of engineering graduation.

2. Probability of GRAD_ENG (Excluding PHQ Data) and Marginal Effects

The logistic regression model can be used to calculate the decimal probability that a given outcome will occur. A midshipman who possessed the mean value for each of the eleven independent variables in the final GRAD_ENG model has an 80.4 percent chance of graduating with an engineering degree. Of greater interest are the marginal effects of each independent variable on this probability. Table 17 summarizes the marginal effects of each variable, which are the decimal changes in the probability that a midshipman will graduate with an engineering degree if the value of that independent variable changes by one standard deviation from its mean. For instance, the midshipman who earns a first semester AQPR one standard deviation higher than the mean has a 5.7 percent greater chance of graduating with an engineering degree. These results are similar in magnitude to one standard deviation improvements on the pre-calculus examination score, technical interest measured by the SCII, and performance in the first semester mathematics course. The increase in the probability of graduating as an engineer for a one standard deviation improvement on the SAT mathematics test (3.9 percent) is outweighed by the decrease in probability of graduating as an engineer caused by scoring one standard deviation higher on the SAT verbal test (5.1 percent).

Of particular interest are the marginal effects of membership in one of the demographic minorities. In the class years 1997 through 2000, being female decreased one's probability of graduating as an engineer by 9.6 percent. Being Asian decreased one's probability by 10.5 percent. Being of a minority other than Black or Asian decreased one's probability by 8.3 percent. According to the model, being Black decreased one's probability of graduating an engineer 1.9 percent; this number, however, is not statistically significant and few conclusions can be drawn from it.

Table 17. Marginal Effects for Graduating with an Engineering Degree

Marginal Effects at Mean Values:				
Independent Variable	Std. Dev.	Mean (X_{ij})	Marginal Effect	Logistic Coefficient (B_j)
SEM1AQPR	0.5708	2.842	0.057	0.634***
PRE_CALC	11.5281	80.025	0.044	0.024***
TISSTD	86.4892	529.078	0.054	0.004***
MAT1PERF	2.8237	7.061	0.067	0.151***
FEMALE	-	0.107	-0.096	-0.612***
SATVAVG	81.2914	559.397	-0.051	-0.004***
SATMAVG	61.9507	651.292	0.039	0.004***
OTHERACE	-	0.067	-0.083	-0.527***
ASIAN	-	0.044	-0.105	-0.666***
CHM1PERF	2.1419	5.400	0.037	0.109**
BLACK	-	0.051	-0.019	-0.121
Constant		1		-6.310***

$u = \Sigma B_j X_{ij}$
u = 1.414863
$P = 1/(1+e^{-u})$ = Probability of Graduating with an Engineering Degree
P = 0.804531834

Note: *** is Highly Significant
** is Significant
* is Marginally Significant

C. GRADUATION IN ENGINEERING HAVING ACHIEVED SUPERIOR ACADEMICS

The second criterion for academic success in the field of engineering is to graduate with an engineering degree while achieving superior academics, which for the purposes of this study means achieving a CQPR >= 3.30. The dependent variable OVER3.30 has been used to model this level of success. As with GRAD_ENG, two models of OVER3.30 have been developed, one with and one without PHQ, respectively; each of the models are the product of an iterative process designed to maximize the predictive power of each model while at the same time maximizing the statistical significance of the individual variables included therein. Comparison of the finalized model for each shows that the PHQ adds no value to the prediction of OVER3.30. Therefore, only the model that excludes PHQ data is presented. Appendix D fully describes the iteration process for OVER3.30 and more fully compares the two models.

1. Final Logistic Regression of OVER3.30 (Excluding PHQ Data)

The final regression model for the dependent variable OVER3.30 consisted of twelve of the thirteen independent variables initially considered for inclusion, seven of which were significant. These variables were (listed in order from largest to smallest contribution to the total chi-square): 1) First Semester Academic QPR (SEM1AQPR), 2) First Semester Mathematics Performance (MAT1PERF), 3) First Semester Chemistry Performance (CHM1PERF), 4) High School Ranking (HS_RANK), 5) Pre-Calculus Examination Score (PRE_CALC), 6) Strong Campbell Interest Inventory Technical Interest Score (TISSTD), and 7) Average Math Scholastic Aptitude Test Score (SATMAVG). Toledo Examination Score (TOT_TOL), Race – Black (BLACK), Race – Asian (ASIAN), Race - Other (OTHERACE), and Gender – Female (FEMALE) are included in the final model but are not highly significant.

The single greatest predictor of whether or not a midshipman will excel academically in an engineering major is that midshipman's performance in the classroom during the first semester. This is followed by mathematics and chemistry performance during that same semester. The fourth strongest predictor is academic performance during high school, followed by the score on the Pre-Calculus placement examination. Strong Campbell Technical Interest and SAT Math are the final two significant predictors.

The signs of coefficients yielded results that are mostly consistent with the previously discussed GRAD_ENG model. The coefficients of the intellective variables were all positive with one notable exception. The coefficient for the Toledo Examination Score was negative. Why this was the case is not known at this time. The demographic variable coefficients were all negative, indicating that membership in any of the minority groups in the study lessened one's chances of achieving a CQPR of 3.30 or greater.

Table 18 summarizes the observed results from the Finalized Logistic Model. The independent variables are listed in order of their contribution to the model chi-square and R^2.

Table 18. Finalized Logistic Regression Model for Engineering and a CQPR >= 3.30

			Chi-Square Contribution		R^2 Contribution	
Independent Variable	B	Sig	Additional	Cumulative	Additional	Cumulative
SEM1AQPR	3.126	0.000	757.117	757.117	0.5517	0.5517
MAT1PERF	0.198	0.000	62.654	819.771	0.0353	0.5870
CHM1PERF	0.254	0.000	26.392	846.162	0.0145	0.6015
HS_RANK	0.003	0.000	20.016	866.179	0.0108	0.6123
PRE_CALC	0.033	0.007	11.959	878.138	0.0064	0.6187
TISSTD	0.003	0.006	9.516	887.653	0.0051	0.6238
SATMAVG	0.004	0.052	4.255	891.908	0.0023	0.6261
TOT_TOL	-0.033	0.100	2.813	894.722	0.0015	0.6275
BLACK	-1.035	0.234	1.567	896.289	0.0008	0.6284
ASIAN	-0.483	0.287	1.138	897.426	0.0006	0.6290
OTHERACE	-0.294	0.497	0.475	897.901	0.0003	0.6292
FEMALE	-0.067	0.824	0.049	897.950	0.0000	0.6292
Constant	-20.804					
-2 Log Like	917.209					
Model Significance	0.000					

Note: Value given for R^2 is the Nagelkerke R^2.
Note: This model excludes PHQ data. (N=1,648)

From Table 18, the chi-square and the -2 log likelihood were 897.950 and 917.209 for this model, respectively. Both demonstrate a high goodness-of-fit for the model and show strong statistical reliability. As can be seen in Appendix D, both statistics are higher than for those of the OVER3.30 model including PHQ data, showing that the PHQ does not contribute to this analysis.

2. Probability of OVER3.30 (Excluding PHQ Data) and Marginal Effects

The logistic regression model can be used to calculate the decimal probability that a given outcome will occur. A midshipman who possessed the mean value for each of the twelve independent variables in the final OVER3.30 model has a 9.8 percent chance of graduating with an engineering degree and a CQPR >= 3.30. Of greater interest are the marginal effects of each independent variable on this probability. Table 19 summarizes the marginal effects of each variable, which are the decimal changes in the probability that a midshipman with graduate an engineer with a CQPR >= 3.30 if the value of that independent variable changes by one standard deviation from its mean. A midshipman who earns a first semester AQPR one standard deviation higher than the mean has a 15.8 percent greater chance of receiving a CQPR >= 3.30 with an engineering degree. A one standard deviation improvement in performance in the first semester math or chemistry course increases the probability by approximately five percent. A one standard deviation improvement in performance on the pre-calculus examination score and a one standard deviation improvement in high school ranking each increase probability by approximately three percent. The same increase in performance on the math SAT or technical interest score increases the probability of obtaining a 3.30 CQPR and an engineering degree by approximately two percent.

The marginal effects of membership in the demographic minorities are of interest but are not statistically significant in the OVER3.30 model. The decrease in probability for membership in a minority group ranges from nine percent for being African American to less than one percent for being female.

Table 19. Marginal Effects for Graduating with an Engineering Degree and CQPR >= 3.30

Marginal Effects at Mean Values:				
Independent Variable	Std. Dev.	Mean (X_{ij})	Marginal Effect	Logistic Coefficient (B_j)
SEM1AQPR	0.5708	2.842	0.158	3.126***
MAT1PERF	2.8237	7.061	0.050	0.198***
CHM1PERF	2.1419	5.4	0.048	0.254***
HS_RANK	105.8917	587.068	0.028	0.003***
PRE_CALC	11.5281	80.025	0.034	0.033***
TISSTD	86.4892	529.078	0.023	0.003***
SATMAVG	61.9507	651.292	0.022	0.004**
TOT_TOL	5.9016	47.273	-0.017	-0.033*
BLACK	-	0.051	-0.092	-1.035
ASIAN	-	0.044	-0.043	-0.483
OTHERACE	-	0.067	-0.026	-0.294
FEMALE	-	0.107	-0.006	-0.067
Constant		1		-20.804***

$u = \Sigma B_j X_{ij}$
$u = -2.216712$
$P = 1/(1+e^{-u})$ = Probability of Graduating with Engineering Degree and CQPR>=3.30
$P = 0.098259752$

Note: *** is Highly Significant
** is Significant
* is Marginally Significant

D. ACADEMIC ACHIEVEMENT IN ENGINEERING MEASURED BY CUMULATIVE QPR

Academic achievement for those studying engineering can be measured by CQPR at graduation. The dependent variable CQPR_ENG has been used to model this level of achievement. As with the previously discussed dependent variables, two models of CQPR_ENG have been developed, one contains PHQ data while the second does not. As described in Chapter IV, these models are the product of an iterative process designed to maximize the predictive power of each model while at the same time maximizing the statistical significance of the individual variables included therein. A comparison of the F Statistic and R^2 value for each finalized model shows that the model excluding PHQ data has higher predictive power. However, the finalized model including PHQ data does contain a PHQ variable, whereas no other finalized model including PHQ data did so. Appendix E fully describes the iteration process for CQPR_ENG.

The final linear regression model for CQPR_ENG included twelve of the thirteen independent variables initially considered for inclusion. There were ten significant predictor variables. These variables (listed in order from largest to smallest contribution to the total R^2) were: 1) First Semester Academic QPR (SEM1AQPR), 2) Average Math Scholastic Aptitude Test Score (SATMAVG), 3) First Semester Chemistry Performance (CHM1PERF), 4) High School Ranking (HS_RANK), 5) First Semester Mathematics Performance (MAT1PERF), 6) Race – African American (BLACK), 7) Race – Asian (ASIAN), 8) Pre-Calculus Examination Score (PRE_CALC), 9) Toledo Examination Score (TOT_TOL), and 10) Strong Campbell Interest Inventory Technical Interest Score (TISSTD). Race – Other (OTHERACE) and Gender – Female (FEMALE) are included in the final model but are not significant.

The signs of the regression coefficients proved interesting in two cases. Eleven of the twelve coefficients had the same sign as in previous analyses. Seven of the eight intellective variables had positive signs, and three of the four demographic variables had negative coefficients. As in previous analyses, the Toledo Examination Score (TOT_TOL) has a negative coefficient. The coefficient for Gender – Female is positive, which is different from all other analyses. However, due to its statistical insignificance, this result is effectively meaningless.

Table 20 summarizes the results from the Finalized Linear Model. The independent variables are listed in order of their contribution to the model R^2. From Table 20, the following linear equation is developed to predict a midshipman's CQPR upon graduation from an engineering degree:

$$CQPR\_ENG_i = 0.47571 + 0.01476(FEMALE_i) – 0.12157(BLACK_i) – 0.09973(ASIAN_i) - 0.04076(OTHERACE_i) + 0.00055(HS\_RANK_i) + 0.00080(SATMAVG_i) + 0.39402(SEM1AQPR_i) + 0.1941(MAT1PERF_i) + 0.04246(CHM1PERF_i) + 0.00311(PRE\_CALC_i) – 0.0444(TOT\_TOL_i) \quad (5 – 1)$$

$CQPR\_ENG_i$ is the predicted CQPR for the *i*th midshipman. (X_i) is the value of the dependent variable, X, for the *i*th midshipman.

Table 20. Finalized Linear Regression Model for Cumulative QPR at Graduation

	R^2 Contribution						
Independent Variable	Additional	Cumulative	B	B SE	β	t	Sig.
SEM1AQPR	0.5751	0.5751	0.39402	0.025	0.453	16.002	0.000
SATMAVG	0.0406	0.6157	0.00080	0.000	0.100	4.750	0.000
CHM1PERF	0.0167	0.6324	0.04246	0.006	0.187	7.135	0.000
HS_RANK	0.0127	0.6451	0.00055	0.000	0.124	6.983	0.000
MAT1PERF	0.0081	0.6532	0.01941	0.004	0.114	4.716	0.000
BLACK	0.0022	0.6554	-0.12157	0.040	-0.052	-3.069	0.002
ASIAN	0.0016	0.6569	-0.09973	0.039	-0.041	-2.531	0.012
PRE_CALC	0.0016	0.6585	0.00311	0.001	0.070	3.166	0.002
TOT_TOL	0.0017	0.6602	-0.00444	0.002	-0.053	-2.609	0.009
TISSTD	0.0012	0.6614	0.00020	0.000	0.036	2.221	0.026
OTHERACE	0.0004	0.6617	-0.04076	0.034	-0.020	-1.197	0.231
FEMALE	0.0001	0.6618	0.01476	0.027	0.009	0.544	0.587
Constant			0.47571	0.128		3.713	0.000
Std. Error of the Estimate	0.2765						
Adj. R^2 for the Model	0.6588						
F Statistic	218.2						

Note: This model excludes PHQ data. (N=1,648)

E. NOTIONAL MIDSHIPMEN

In order to demonstrate the utility of the final models for each dependent variable, this section presents two notional midshipmen. A notional midshipman is a collection of characteristics, defined by values for the thirteen independent variables included in the final model for each dependent variable. These values will be inserted into the final models and the results discussed.

In discussing the marginal effects for the logistic regression analyses, a notional midshipman was created that possessed the mean values for each independent variable. This is effective in describing the marginal effects of the various independent variables but is not realistic in describing a true person, because of the presence of dummy variables in these analyses. The mean of a dummy variable (which, in actuality, can only be a one or zero) equals the percentage of cases whose value is one. The following notional midshipmen are analyzed in terms of the three dependent variables.

1. Midshipman Dick

Consider Midshipman Dick. He isn't very interested in matters technical, as determined by his Strong Campbell Interest Inventory results, but scored extremely well on his pre-calculus examination and was then able to validate two semesters of calculus. He received an A in his calculus III course. He did not fare as well in his chemistry course and only received a C. He did not do very well in his other courses as evidenced by his semester AQPR. It is two weeks before he has to decide his major and is agonizing over choosing an engineering major or not. Table 21 more fully describes Midshipman Dick.

Table 21. Midshipman Dick

Gender:	Male
Race:	Asian
High School Ranking:	640
Average Math Scholastic Aptitude Test Score:	720
Average Verbal Scholastic Aptitude Test Score:	510
Pre-Calculus Examination Score:	93
Toledo Examination Score:	33
Technical Interest Score from SCII:	440
First Semester Mathematics Performance:	16
First Semester Chemistry Performance:	4
First Semester Academic QPR:	2.45

Utilizing the final models for the three dependent variables, the following predictions can be made concerning Midshipman Dick's possible success as an engineering student at the United States Naval Academy:

Midshipman Dick has a 90.56 percent chance of graduating with an engineering degree. He has a 20.68 percent chance of graduating with an engineering degree and obtaining a CQPR >= 3.30.

Using Equation 5 – 1, Midshipman Dick's predicted CQPR at graduation with an engineering degree is:

$$CQPR\_ENG = 0.47571 + 0.01476(0) – 0.12157(0) – 0.09973(1) - 0.04076(0) + 0.00055(640) + 0.00080(720) + 0.39402(2.45) + 0.1941(16) + 0.04246(4) + 0.00311(93) – 0.0444(33) = 2.98$$

2. Midshipman Jane

Consider Midshipman Jane. She has always been interested in technical issues and is looking forward to choosing an engineering major. Her pre-calculus examination score was average; she was unable to validate any mathematics courses. She received a B in Calculus I and an A her first semester chemistry course. Her first semester AQPR shows that she did well in her other courses. Table 22 more fully describes Midshipman Jane.

Table 22. Midshipman Jane

Gender:	Female
Race:	White
High School Ranking:	560
Average Math Scholastic Aptitude Test Score:	600
Average Verbal Scholastic Aptitude Test Score:	580
Pre-Calculus Examination Score:	72
Toledo Examination Score:	46
Technical Interest Score from SCII:	595
First Semester Mathematics Performance:	6
First Semester Chemistry Performance:	8
First Semester Academic QPR:	3.15

Utilizing the final models for the three dependent variables, the following predictions can be made concerning Midshipman Jane's possible success as an engineering student at the United States Naval Academy:

Midshipman Jane has a 73.90 percent chance of graduating with an engineering degree. She has a 25.33 percent chance of graduating with an engineering degree and obtaining a CQPR >= 3.30.

Using Equation 5 – 1, Midshipman Jane's predicted CQPR at graduation with an engineering degree is:

$$CQPR\_ENG = 0.47571 + 0.01476(1) - 0.12157(0) - 0.09973(1) - 0.04076(0) + 0.00055(560) + 0.00080(600) + 0.39402(3.15) + 0.1941(6) + 0.04246(8) + 0.00311(72) - 0.0444(46) = 3.11$$

F. CHAPTER SUMMARY

Of the six models estimated for the three dependent variables, three models were chosen for discussion in this chapter. As the finalized models, they maximize predictive power and the significance of the included variables. None of the three finalized models contained PHQ data. In all cases, the comparison between the model that did possess PHQ data and the model that did not showed that the models that excluded PHQ data were more powerful.

Notional midshipmen were then developed to demonstrate the utility of these models in predicting possible future success for engineering students. The models could serve as the basis for a computer application that midshipmen could use in order to prepare to choose an academic major.

VI. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

A. SUMMARY

This study explored the relationship between various predictor variables and engineering success at the United States Naval Academy. These variables were drawn from admissions (pre-enrollment) records as well as academic performance during the first semester of the freshman year. The findings of this thesis suggest that those midshipmen who perform well during their first semester, particularly in their calculus and chemistry courses, have the ability to succeed as engineering students at USNA.

Chapter I introduced the issues faced by USNA in attempting to provide the United States Navy an officer corps possessing a firm background in matters technical. It also provided the five research questions that this thesis examines: (1) What admissions and plebe year variables significantly affect graduation with an engineering degree at the United States Naval Academy? (2) What admissions and plebe year variables significantly affect superior academic performance in engineering majors at the United States Naval Academy? (3) What admissions and plebe year variables significantly affect Cumulative Quality Point Rating (CQPR) for engineers at the United States Naval Academy? (4) Does the Personnel History Questionnaire (PHQ) contribute to the prediction of academic success for engineers at the United States Naval Academy? (5) Can a model be devised to assist midshipmen in choosing an academic major?

Chapter II described in detail several studies performed at other colleges and universities that explored these same issues. Focus was placed on the quantitative research that could be used in a model on which to base this effort. Also discussed was the general need for improved academic counseling at the university level.

Chapter III described the study's data. This description included the sources of data, the specific variables used in the analyses, and the data cleaning techniques used to ensure accurate analysis. The data set was composed of nineteen independent or explanatory variables used to predict the three dependent variables. The study attempted to use both intellective and non-intellective variables in its analysis, however, as will be described below, the non-intellective variables did not prove to be statistically significant.

Chapter IV described the statistical analyses used to answer the research questions posed in the Introduction. Logistic regression was used to predict academic success in the form of graduation with an engineering degree and graduation with an engineering degree and superior academic performance, while linear regression was used to predict academic achievement in the form of cumulative quality point rating. For each analysis, the final model was estimated through an iterative process that eliminated those variables that were not statistically significant. Each model was estimated including, and then excluding, Personal History Questionnaire data in order to test its utility for predicting success in engineering majors.

Chapter V presented the results of the analyses described in Chapter IV. Each analysis was presented in a manner that showed the contribution of each variable to the total model. The logistic regressions were also presented showing the marginal effects of each variable to the percentage probability of the outcome.

Chapter I presented the five research questions that this thesis investigated. The answers to each of those questions are discussed here.

1. Graduation with an Engineering Degree

What admissions and plebe year variables significantly affect graduation with an engineering degree at the United States Naval Academy? By far, the single greatest predictor of graduating with an engineering degree is first semester academic performance. It is during this semester that a midshipman is first subjected to the rigors of Naval Academy life; the ability to perform academically during this turbulent time is extremely important for would-be engineers. This is no surprise; it matches the findings of other researchers (Benefield et al., 1996; Fletcher, Halpin, and Halpin, 1999; and Pascarella, Duby, Miller, and Rasher, 1981). Following in importance is the score received by a midshipman on the pre-calculus placement exam given by the Mathematics Department during plebe summer. Technical interest is the third strongest predictor, followed by first semester mathematics performance.

Performance on the Scholastic Aptitude Test (SAT) only minimally predicts graduation as an engineer; this again is in keeping with the literature, which states that the

SATs are most predictive of early college performance. Their utility for prediction diminishes the further one moves into a collegiate career.

Gender and race play a role in engineering success. Women are only ninety percent as likely to graduate with an engineering degree as their male counterparts. Racial minorities face a handicap of similar magnitude as women. The reasons for this are not apparent from this study. These reasons most surely lie in the realm of non-intellective factors, which, for reasons discussed below, were dropped from the analysis.

2. Superior Academics with an Engineering Degree

What admissions and plebe year variables significantly affect superior academic performance in engineering majors at the United States Naval Academy? Again, the single greatest predictor of achieving superior academics in an engineering degree was academic performance during the first semester. Chemistry and math grades during the first semester were the second and third strongest predictors, respectively. This supports the findings of Levin and Wychoff (1990) who showed that the best predictors of success were performance in the most recent math and science courses taken. High school class standing, performance on the pre-calculus placement examination, and technical interest rounded out the highly significant contributors.

Gender and race were not statistically significant to the models predicting superior academics. There are two possible reasons for this: 1) gender and race are important in this area but, due to the small numbers of these individuals who attained superior academics, the results are not statistically significant, or 2) gender and race do not play a role at the highest levels of academic achievement and are eclipsed by other factors that are gender and race neutral.

3. Cumulative Quality Point Rating for Engineers

What admissions and plebe year variables significantly affect Cumulative Quality Point Rating (CQPR) for engineers at the United States Naval Academy? In a linear regression model predicting CQPR, the strongest predictor is first semester academic performance. Seven other intellective variables are highly significant to the model; they are average math SAT score, performance in the first semester chemistry course, high school class standing, performance in the first semester mathematics course, pre-calculus placement examination score, Toledo examination performance, and technical interest.

The intellective variables all positively affected CQPR with the exception of the Toledo Examination results. One possible explanation is that those midshipmen who do well on the chemistry placement test tend to choose Division II majors to take advantage of their aptitude in math and the sciences. Belonging to Black or Asian minority also significantly and negatively affected CQPR, however being female or a member of any other racial minority did not.

4. The Personal History Questionnaire

Does the Personnel History Questionnaire contribute to the prediction of academic success for engineers at the United States Naval Academy? The findings of this research are that the PHQ does not contribute to the prediction of academic success. That is not to say that the PHQ is not valuable, merely that the method of employment of PHQ data in this study proved to be insignificant for academic prediction. The PHQ may very well prove valuable if analyzed by other means or via other methodologies.

The PHQ was examined because of its status as one of the few sources of non-intellective data that is quantifiable. Non-intellective data must be considered when predicting academic success, especially when the aim is to improve counseling for students. To ignore this type of data is to ignore those qualities that make us human: motivation and desire. And it is because of the existence of these qualities in students that purely quantitative analyses cannot be used solely for academic counseling.

5. Prediction Models for Academic Success

Can a prediction model be devised to assist midshipmen in choosing an academic major? Yes, a prediction model or models can and should be implemented to aid midshipmen. The models estimated as part of this study are similar to those estimated by Levin and Wychoff (1987 and 1990) who state:

> The outcome models of this study are uniquely suited for advising purposes because of the following attributes: 1) predictive statements can be made for students on an individual basis because individual student characteristics are analyzed by the models; 2) students and advisors together can examine the likelihood of a variety of predictive outcomes

> depending on the relevancy of the outcome to the student; 3) the models provide results that are easily interpreted by advisors and understood by students.
>
> To further elaborate on the usefulness of the models, it should be noted that the models will allow students, via the advising process, to understand the extent of risk involved in their educational plans, and to make decisions regarding risk levels that may be personally acceptable. This is possible because students will be able to identify the way their personal characteristics contribute to the predicted outcome. (p. 40)

The power of the model predicting Cumulative Quality Point Rating (CQPR) in particular (R^2 = 0.659) shows its utility in counseling. The model accounts for nearly sixty-six percent of the variance in CQPR. It is possible that this value approaches the limit of variance for which a model of this type can account. Variables that have not and possibly cannot be accounted for such as motivation make up the other thirty-four percent.

B. CONCLUSIONS

The tacit aim of this thesis was to examine academic success as measured by the completion of an engineering curriculum and academic achievement in the form of a quality point rating. In actuality, this study examined both of these issues. Moreover, those midshipmen who graduated with engineering degrees did so for two reasons: 1) they overcame the academic challenges inherent to an engineering curriculum, and 2) they chose to remain in the engineering track through to graduation. Although this study did not specifically account for it, many of those midshipmen who left engineering were successful: They graduated in their new program of study.

Logistic and linear analyses were used to estimate success, achievement, and persistence; the results of all analyses showed strong agreement with the literature. The assumption stated in the Summary of Chapter II that there is little difference between midshipmen and university students studying engineering appears to have been valid. This study supports the use of statistical modeling in academic counseling, especially in regard to choosing an academic major.

It must be stressed that the utility of such prediction models are not as stand-alone tools. Because of the fact that such models are predominantly quantitative and intellective in nature, they can and should provide only one data point for midshipmen attempting to choose their academic major. It is believed, however, that the use of such models by plebe year academic counselors during their interactions with plebes would enhance the advising process, at least in regards to the question of whether or not to choose an engineering major.

C. RECOMMENDATIONS

1. To the United States Naval Academy

Several members of the faculty of the Naval Academy assisted this effort, both in offering suggestions and providing data. Commonly, the faculty expressed an interest in this type of research and a desire to conduct similar analyses within their own academic sphere. The prime inhibitor to such efforts is a lack of understanding of statistical analyses and methodologies. The first recommendation of this thesis to the Naval Academy, and to all institutions of higher learning, is to expand the educational opportunities for faculty in terms of statistical analysis. At the Naval Academy, two organizations have statistical expertise; they are the Office of Institutional Research, Planning, and Assessment, and the Economics Department. The Naval Academy could improve its ability to fulfill its mission by promoting and supporting faculty education in the realm of statistical assessment.

In gathering data for this thesis, two opportunities were missed due to the difficulty in properly formatting and assembling the pertinent information. The first opportunity was to include in this study LASSI data as analyzed by Watson (2001). Prior to his study, LASSI data were only available in the form of several thousand hardcopy test answer sheets kept by the Academic Center. Watson spent a considerable amount of time scanning these sheets to obtain computer data files to be used in his analyses. In light of his efforts, Watson (2001) states that "the Naval Academy Academic Center has taken steps to ensure that a consistent record is being maintained so that further analysis may be conducted if necessary" (p. 91). However, upon request, the Naval Academy Academic Center was unable to provide a copy of the database that Watson constructed for his thesis. The author believes that, due to the significance of his findings, Watson's data would have been invaluable to this thesis.

The second missed opportunity involved data from the Physics Department's Physics Diagnostic Test (PDT), which is a slightly revised version of the Force Concept Inventory (FCI), a diagnostic test used by colleges and universities throughout the nation to test students' understanding of physics (Hestenes, Wells, and Swackhammer, 1992). Levin and Wychoff (1990) conclude that an understanding of physics, based on grades received in physics courses, predicts future success in engineering studies. An understanding of the basic concepts that underlie physics, which the PDT and FCI test, should also contribute to the prediction of engineering success. These data were unavailable for this study due to time constraints; the outdated format and the difficulties in conversion and organization prevented the data from being supplied and subsequently used in this study.

The second recommendation of this thesis to the Naval Academy is to require all Departments that give standardized tests to maintain the data in organized files that use the most current data format. Maintaining data files in formats that are no longer in general use greatly hampers their usefulness. The problem is particularly acute due to the fact that the Naval Academy recently updated from and eliminated its legacy computer system, the Naval Academy Time Share (NATS). Large amounts of data were lost in this transition. In other cases, the data were maintained but are kept in out-of-date formats that are disjointed or inaccessible. The Naval Academy should make every effort to retrieve, reorganize, or reconstruct whatever data it can; this will facilitate robust statistical analysis in all areas and improve the Academy's ability to fulfill its mission.

The final research question for this study involved prediction models and their use by midshipmen. The author has created a model, in Excel spreadsheet format, that could be used by academic counselors to assist plebes choosing a major. As stated earlier, such a model is only a tool to be used as part of a more complete counseling regimen, but it can supply information that forms the basis for objective and reasoned discussion. The final recommendation of this thesis for the Naval Academy is to adopt some form of model to aid in academic counseling.

In making these recommendations, the author understands that hard data and statistical analysis cannot answer every question or solve every problem. Further, the questions they do answer only describe one aspect of the issue. In light of the findings of this thesis, the author fully supports their use in tackling those issues that may warrant their use.

2. For Further Research

This study has focused on the academic prediction of success in Division I (engineering) majors at the United States Naval Academy. Another equally valuable avenue of research would be to conduct similar analyses for the Division II (math and sciences) and Division III (humanities and social sciences) majors. The analysis of Division I and II majors may be identical in terms of variables used due to the similarities in subject matter. The analysis for the Division III majors might rely on other variables, i.e., performance in plebe year English courses and the U.S. Government or Naval History courses as opposed to performance in calculus or chemistry, or rely on the same variables with very different expected outcomes, i.e., the Strong Interest Inventory technical interest scale. These research efforts could lead to computer models as described above for each of the majors groups, which, in turn, would lead to a greater improvement in the academic counseling of plebes choosing a major.

In a follow-on study, a researcher could conduct similar analyses that take into account a wider range of variables in order to improve the predictive capabilities of the estimated models. LASSI data, if retrieved, could be used as additional independent variables for those year groups that took the LASSI. It is no longer administered by the Academic Center, but including data similar to the LASSI from other surveys conducted by the Academic Center may likewise improve the prediction of academic performance. The Physics Diagnostic Test could be included as an independent variable. A third source of data is the Candidate Information System (CIS) on the Naval Academy website. The CIS is a collection of online surveys used by the Academy and several different academic departments to gather data about incoming freshman; all incoming plebes are required to submit information on these surveys. Much of the data have direct utility for academic prediction in general, and for engineering in particular. The CIS is a good source of non-intellective data of the sort received from the Personal History Questionnaire.

Appendix B provides descriptive statistics for the entire data set used in this study, which consists of the members of the classes of 1997 through 2000 who initially chose an engineering major. These statistics are organized according to gender and race and contain a wealth of untapped information. A future researcher could conduct One Way Analysis of Variance on these data to accurately determine the differences in engineering students between demographic groupings.

Still other types of research in this area are needed. A qualitative study involving interviewing engineering students, ex-engineering students, and others, similar to the research efforts of MacGuire and Halpin (1995), could shed light on those issues that most directly affect the majors decision for plebes. Interviewing midshipmen who change to a non-engineering major could shed light on their reasons for doing so. Another area of research that has important consequences for the Naval Academy and the United States Navy is minority performance in engineering majors. This study has shown that minorities, both gender and racial, perform at a level below their majority peers in engineering academics but has done nothing to determine the reasons for this troubling finding. Until these reasons are known and understood, no steps can be taken to address this problem.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A – PERSONAL HISTORY QUESTIONNAIRE

A. INTRODUCTION

The Personal History Questionnaire is one of a limited number of data sources that is able to provide non-intellective data concerning midshipmen. As such, the role that it may be able to play in predicting academic success is interesting and worth analyzing. A total of six variables from the PHQ were used in the analysis. Three were taken directly from individual questions in the survey; they are Semesters of Advanced or Honors Math Courses Taken in High School (MATHSEM), Mother's Education Level (MAGRADED), and Father's Education Level (PAGRADED). The other three variables for analysis were developed from the PHQ; they are the Hardwork Score (HARDWORK), the Military Aptitude Score (MIL_APT), and the Academic Preparation Score (AC_PREP). Their development is described below. Further, a copy of the actual PHQ is enclosed.

B. BACKGROUND

The Personal History Questionnaire is a survey consisting of eighty-five questions pertaining to candidate's families, interests, and experiences. Prepared by the Navy Personnel Research and Development Center (NPRDC) in the mid-nineteen eighties, the PHQ was to serve as a replacement to the Strong-Campbell Interest Inventory should it be necessary. The PHQ was administered to the entering USNA classes of 1988 through 2000. Though no longer administered by USNA, IR has maintained the results of this survey on file.

NPRDC developed the PHQ by borrowing from items existing in the public domain, modifying items found in other commercially available questionnaires, and independently authoring several questions in-house. There exists no documentation pertaining to its effectiveness with respect to predicting academic success for engineers. One aim of this study is to determine its utility in this area.

Two distinct issues call into question the accuracy of the PHQ in truly assessing candidate suitability. These issues must be kept in mind when using data from this questionnaire. First, the PHQ contains questions that ask candidates to describe their own quality. Candidates hoping to be admitted to the Academy may feel pressured to "game" the survey to improve their chances. Second, the PHQ was not administered in a standard manner across all candidates; it was mailed to each candidate as part of the admissions packet. In spite of the problems inherent in the survey, the PHQ does provide a type of data not available through other means. For this reason, data from the PHQ have been used.

The following three sections describe how the independent variables HARDWORK, MIL_APT and AC_PREP were prepared.

C. DEVELOPMENT OF HARDWORK

The HARDWORK variable was created by assigning a score to the answer of each of nine questions from the PHQ and summing those scores. Table 23 summarizes the values for each answer to the nine questions that make up HARDWORK.

Table 23. HARDWORK Point Values

	Answer				
Question	A	B	C	D	E
# 3	5	4	3	2	1
# 5	1	2	3	4	5
# 12	5	4	3	2	1
# 22	5	4	3	2	1
# 47	5	4	3	2	1
# 54	3	2	2	1	-
# 59	4	3	2	1	-
# 61	1	2	3	4	5
# 65	5	4	3	2	1

HARDWORK = Score (#3) + Score (#5) + Score (#12) + Score (#22)

+ Score (#47) + Score (#54) + Score (#59) + Score (#61)

+ Score (#65)

D. DEVELOPMENT OF MILITARY APTITUDE

The MIL_APT variable was created by assigning a score to the answer of each of three questions from the PHQ and summing those scores. Table 24 summarizes the values for each answer to the three questions that make up MIL_APT.

Table 24. MIL_APT Point Values

	Answer				
Question	A	B	C	D	E
# 6	5	4	3	2	1
# 21	5	4	3	2	1
# 26	5	4	3	2	1

MIL_APT = Score (#6) + Score (#21) + Score (#26)

E. DEVELOPMENT OF AC_PREP

The AC_PREP variable was created by assigning a score to the answer of each of four questions from the PHQ and summing those scores. Table 25 summarizes the point values for each answer to the four questions that make up AC_PREP.

Table 25. AC_PREP Point Values

	Answer				
Question	A	B	C	D	E
# 67	5	4	3	2	1
# 71	1	2	3	4	5
# 72	5	4	3	2	1
# 74	5	4	3	2	1

AC_PREP = Score (#67) + Score (#71) + Score (#72) + Score (#74)

F. PERSONAL HISTORY QUESTIONNAIRE FORM

PERSONAL HISTORY QUESTIONNAIRE

FORM PHQ-84A

THIS FORM SUBJECT TO THE PRIVACY ACT OF 1974. SEE PAGE 8.

This survey consists of a number of questions about you, your family, your interests, and your experiences.

Read each question and all of its possible answers carefully. Being as frank as you can, quickly select the one answer that is most appropriate for you. Then mark the circle corresponding to that answer in the space provided in this questionnaire booklet. Be sure to select one answer — and only one answer — for every question.

Please mail your completed booklet promptly to the Naval Academy in one of the envelopes provided.

GENERAL INSTRUCTIONS

USE NO. 2 PENCIL ONLY

- DO NOT USE FELT TIP, BALLPOINT OR INK PENS.
- COMPLETELY BLACKEN CIRCLE.
- MAKE CLEAN ERASURES.
- MAKE NO STRAY MARKS.
- DO NOT WRINKLE BOOKLET.
- DO NOT SEPARATE BOOKLET PAGES.

CORRECT MARKS

●●●●●

INCORRECT MARKS

Now turn the page and fill in your name and other identifying information.

DO NOT MARK IN THIS AREA

234290

*U.S. GOVERNMENT PRINTING OFFICE: 1996—412-151

RECORD YOUR IDENTIFYING INFORMATION IN THE BOXES AND CIRCLES PROVIDED BELOW.

- **NAME:** Print as much as possible of your last name and your first and middle initial.
- **SOCIAL SECURITY NUMBER.**
- **TODAY'S DATE:** Year, month, and day.
- **BIRTHDATE:** Year and month only.
- **SEX.**
- **HIGHEST GRADE PASSED:** Mark the highest grade passed as of today's date. If G.E.D., fill in the H.S. GRAD circle.

LAST NAME | FI | MI

SOCIAL SECURITY NUMBER

TODAY'S DATE: YR. | MO. | DAY

BIRTH DATE: YR. | MO.

SEX

MALE

FEMALE

HIGHEST GRADE PASSED

10 OR LESS

11

12 H.S. GRAD

13

14 J.C. GRAD

15

16 AB-BS

MA-MS

Now go on to the next page and answer the survey questions.

1. When I first earned money on a regular basis (other than from members of my family), my age was:
   - Ⓐ 12 or younger.
   - Ⓑ 13–14.
   - Ⓒ 15.
   - Ⓓ 16 or older.
   - Ⓔ I haven't had a paid regular job.

2. How many hours a week have you usually worked on paying jobs since the beginning of the 11th grade? (Do not include summer jobs.)
   - Ⓐ None.
   - Ⓑ 1 to 10 hours.
   - Ⓒ 11 to 15 hours.
   - Ⓓ 16 to 20 hours.
   - Ⓔ More than 20 hours.

3. I usually do:
   - Ⓐ Much more than I resolved to do.
   - Ⓑ A bit more than I resolved to do.
   - Ⓒ Never less than what I resolved to do.
   - Ⓓ A little less than I resolved to do.
   - Ⓔ Much less than I resolved to do.

4. I would find a life in which one wouldn't have to work at all:
   - Ⓐ Very pleasant.
   - Ⓑ Pleasant.
   - Ⓒ Somewhat unpleasant.
   - Ⓓ Unpleasant.
   - Ⓔ Very unpleasant.

5. When doing something difficult:
   - Ⓐ I give up very quickly.
   - Ⓑ I give up rather quickly.
   - Ⓒ I give up somewhat quickly.
   - Ⓓ I don't give up too soon.
   - Ⓔ I usually see it through.

6. When I became interested in a military career, my age was
   - Ⓐ 12 or younger.
   - Ⓑ 13 to 14.
   - Ⓒ 15.
   - Ⓓ 16.
   - Ⓔ 17 or older.

7. During the last couple of years, the part of my own support that I personally earned was approximately,
   - Ⓐ Less than 10%.
   - Ⓑ 10% to 30%.
   - Ⓒ More than 30% but less than 60%.
   - Ⓓ 60% to 90%.
   - Ⓔ More than 90%.

8. The amount of influence the members of my family exercised on my career choice was:
   - Ⓐ A great deal; they exerted considerable pressure on me to accept their choice.
   - Ⓑ Some; they exerted moderate pressure on me to accept their choice.
   - Ⓒ A little; they encouraged me generally but left the choice up to me.
   - Ⓓ A great deal, but they did not try to pressure me to accept their choice.

9. Without any false modesty, I believe that the highest rank that I could reach in the course of a Navy career is:
   - Ⓐ Lieutenant.
   - Ⓑ Commander.
   - Ⓒ Captain.
   - Ⓓ Admiral.
   - Ⓔ I don't know.

10. Other than those required for school, how many books do you usually read?
    - Ⓐ At least one a week.
    - Ⓑ Two or three a month.
    - Ⓒ About one a month.
    - Ⓓ About one every 6 months.
    - Ⓔ One or less a year.

11. Indicate the total number of semesters of advanced and honors mathematics courses, such as those listed below, that you expect to complete in junior and senior high school. Include all courses you have taken since beginning the seventh grade (summer courses also), but count each course only once.
    - advanced algebra
    - advanced mathematics
    - analytic or coordinate geometry
    - trigonometry
    - independent study, seminars, or special clusters in mathematics
    - precalculus
    - calculus
    - computer programming

    - Ⓐ 6 or less.
    - Ⓑ 7–8.
    - Ⓒ 9–10.
    - Ⓓ 11–12.
    - Ⓔ 13 or more.

12. How often do you accomplish what you must do without having to be pushed to do it (by others)?
    - Ⓐ Always.
    - Ⓑ Very often.
    - Ⓒ Often.
    - Ⓓ Sometimes but not often.
    - Ⓔ Rarely.

TURN PAGE AND CONTINUE

13. Since I started high school, my money for recreation (or "extras") usually came from:

Ⓐ Allowance and gifts from family.
Ⓑ Mostly allowance and gifts, some my own earnings.
Ⓒ Mostly my own earnings, some from the family.
Ⓓ All from my own earnings.

14. If I wake up in the morning feeling a little "out of sorts" but don't feel really ill, I:

Ⓐ Don't go to school or work because it's possible that I might be coming down with something serious.
Ⓑ Go to school or work but take medicine "just in case" or let everyone know just how bad I feel.
Ⓒ Go to school or work without any unnecessary complaining but consider going home if I get noticeably worse.
Ⓓ Go to school or work without hesitation because I consider that my responsibilities come first.

LISTED BELOW ARE SOME ACTIVITIES AND CONCERNS OF FIRST YEAR NAVAL ACADEMY STUDENTS. USING THE FOLLOWING SCALE, INDICATE HOW YOU ESTIMATE EACH ITEM WILL APPLY TO YOU DURING YOUR FIRST YEAR AT THE ACADEMY.

During my first year at the Academy I expect I will:	HIGHLY PROBABLE	SOMEWHAT PROBABLE	SOMEWHAT IMPROBABLE	HIGHLY IMPROBABLE
15. Earn academic honors.	Ⓐ	Ⓑ	Ⓒ	Ⓓ
16. Earn honors in athletics.	Ⓐ	Ⓑ	Ⓒ	Ⓓ
17. Need tutoring in one course.	Ⓐ	Ⓑ	Ⓒ	Ⓓ
18. Earn military honors.	Ⓐ	Ⓑ	Ⓒ	Ⓓ
19. Have difficulties with studies or concentration.	Ⓐ	Ⓑ	Ⓒ	Ⓓ
20. Seek vocational or individual counseling.	Ⓐ	Ⓑ	Ⓒ	Ⓓ

21. Most Navy officers' jobs fall into one of the following categories: Surface Warfare, Submarine Warfare, Naval Aviation, and "Other" (such as Supply, Naval Intelligence, Public Affairs). Do you already know which of these areas you would like to go into?

Ⓐ Yes, I have known for a while.
Ⓑ Yes, I am almost sure.
Ⓒ Yes, but very tentatively.
Ⓓ No, I like them all equally.
Ⓔ No, I still don't know enough to decide.

22. If I have not attained my goal and have not done a task well then:

Ⓐ I continue to do my best to attain the goal.
Ⓑ I exert myself once again to attain the goal.
Ⓒ I find it difficult not to lose heart.
Ⓓ I'm inclined to give up.
Ⓔ I usually give up.

23. Of the following I feel that the thing I would like most in a job would be:

Ⓐ Promotion and pay according to ability.
Ⓑ Satisfactory vacations.
Ⓒ Good supervision.
Ⓓ Freedom to make decisions.
Ⓔ Working for myself.

24. The factor that was most responsible for my interest in a military career was:

Ⓐ Admiration for military heroes.
Ⓑ Educational benefits or job security.
Ⓒ The influence of close friends.
Ⓓ Advice from parents or guardians.
Ⓔ Personal preference over other careers.

25. The number of times per week I usually go out socially is:

Ⓐ 1 or less.
Ⓑ 2.
Ⓒ 3.
Ⓓ 4.
Ⓔ 5 or more.

26. Compared to most people my age, I think I will get used to military life:

Ⓐ Much more easily.
Ⓑ A little more easily.
Ⓒ About as easily.
Ⓓ A little less easily.
Ⓔ Less easily.

27. By the end of my first semester at the Academy I expect my grades will be:

Ⓐ Good in some courses, low in others, with an overall average high enough to stay in school.
Ⓑ Average or better in every course.
Ⓒ Very good in every course except possibly one.
Ⓓ Excellent in every course.

28. By the end of my first year at the Naval Academy I expect my military performance will be:

Ⓐ Good in some areas, low in others, with an overall average high enough to stay in the service.
Ⓑ Average or better in every aspect
Ⓒ Very good in almost every aspect.
Ⓓ Excellent in every aspect.

CONTINUE TO NEXT PA

29. In the past, how have you reacted to competition?

Ⓐ I have always done my best in competitive situations.
Ⓑ I have usually done my best in competitive situations.
Ⓒ I have done all right, but haven't liked it.
Ⓓ I have been unaffected by it.

LISTED BELOW ARE SOME OF THE TRADITIONAL ADVANTAGES OF A MILITARY CAREER AND A NAVAL ACADEMY EDUCATION. USING THE FOLLOWING SCALE, INDICATE HOW IMPORTANT EACH ITEM WAS IN YOUR DECISION TO CONSIDER A MILITARY CAREER.

	EXTREMELY IMPORTANT	VERY IMPORTANT	IMPORTANT	RATHER UNIMPORTANT	VERY UNIMPORTANT
30. Opportunities to direct others, have responsibilities and authority.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
31. Disciplined life style.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
32. Financial and tuition benefits.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
33. Promotion opportunities.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
34. Lack of adventure in civilian jobs.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
35. Economic security.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
36. Novelty of experience.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ

USING THE FOLLOWING SCALE, INDICATE HOW IMPORTANT IT IS TO YOU PERSONALLY, TO PURSUE EACH OF THE GOALS LISTED BELOW.

	EXTREMELY IMPORTANT	VERY IMPORTANT	SOMEWHAT IMPORTANT	SOMEWHAT UNIMPORTANT	VERY UNIMPORTANT
37. An exciting life.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
38. A sense of accomplishment.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
39. Family security.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
40. Social recognition.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
41. Financial success.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ

42. In comparison with most of the people I know, able to give a talk before a group:

Ⓐ Much more easily.
Ⓑ Somewhat more easily.
Ⓒ Just as easily.
Ⓓ A little less easily.
Ⓔ Much less easily.

43. Compared to others my age, I think my athletic abilities are:

Ⓐ In the top 1%.
Ⓑ In the top 5%.
Ⓒ In the top 25%.
Ⓓ Average.
Ⓔ Below average.

44. Naval Academy students sometimes leave befor receiving their commission. If this should happe to you, which of the following do you think wou be the MOST LIKELY cause?

Ⓐ Not applicable, I am absolutely certain I will obtain a commission.
Ⓑ Change to a major not offered at the Naval Academy.
Ⓒ Lack of ability for military service.
Ⓓ Lack of academic ability or necessary study skills.
Ⓔ Other.

45. What kind of upbringing did you have?

Ⓐ Strict but fair.
Ⓑ Strict but unfair.
Ⓒ Inconsistent.
Ⓓ Not very strict but fair.
Ⓔ Not very strict but unfair.

46. When growing up, how often, compared to other your age, were you allowed to make your own decisions?

Ⓐ Much more often.
Ⓑ Somewhat more often.
Ⓒ About as often.
Ⓓ Somewhat less often.
Ⓔ Much less often.

47. How often do you help with chores and tasks around the home, the yard, or a family business?

Ⓐ Very often; I have jobs assigned to me and a regular schedule to do them.
Ⓑ Often, but not regularly.
Ⓒ Sometimes, when I am asked.
Ⓓ Sometimes; my parents complain a great deal but they rarely make me help.
Ⓔ Rarely or never; I am not required to.

TURN PAGE AND CONTINUE

48. I think of myself as a shy person:
- (A) Never.
- (B) Hardly ever.
- (C) Sometimes.
- (D) Often.
- (E) Almost always.

The following two questions must be answered in coordination. Please read all choices carefully before answering.

49. Which one of the following recreational activities do you engage in most often?
- (A) Participating in competitive team sports.
- (B) Participating in cómpetitive individual sports.
- (C) Noncompetitive jogging, swimming, or other physical activity.
- (D) Social relaxation with others, such as parties, dances, etc.
- (E) None of the above. (CHOOSE A, B, C, or D in the NEXT QUESTION.)

50. Which one of the following recreational activities do you engage in most often? (ANSWER "E" IF YOU CHOSE A, B, C or D IN THE PREVIOUS QUESTION.)
- (A) Reading, listening to records or other solitary activities.
- (B) Attending or participating in plays, concerts, or other artistic activities.
- (C) Working on cars, bikes, models, or electronics.
- (D) Sailing, hiking, camping, or horseback riding.
- (E) None of the above; I chose A, B, C, or D in the previous question.

51. How often have you changed your mind about future career plans since you entered high school?
- (A) I have not changed my plans.
- (B) Only once.
- (C) Two or three times.
- (D) Four or more times.

52. When I had my first evening date my age was:
- (A) 13 or younger.
- (B) 14.
- (C) 15.
- (D) 16 or older.
- (E) I haven't had an evening date.

53. How well do you do most things you have decided to do?
- (A) I almost always succeed in the things I attempt and do them better than most people do.
- (B) I often find that I have bitten off more than I can chew and have to give up.
- (C) I usually get the things done that I attempt, but I seldom do them as well as I want to.
- (D) I find that I do most things as well as other people do.
- (E) I seldom get the things done that I attempt, but I usually do them better than other people.

54. Which of the following is most typical of your study habits?
- (A) I work quite regularly.
- (B) I usually get to work when deadlines get close.
- (C) I usually have to be in the mood.
- (D) I work quite irregularly.

55. In comparison with most of the people I know, I am able to make new friends:
- (A) Much more easily.
- (B) A little more easily.
- (C) As easily as other people.
- (D) A little less easily.
- (E) Less easily.

56. How many nonfiction magazines do you read each month?
- (A) None.
- (B) 1 or 2.
- (C) 3 or 4.
- (D) 5 or 6.
- (E) More than 6.

57. When faced with an unpleasant situation, I usually:
- (A) Try to react immediately and figure out the best solution.
- (B) Put if off for a little while so I can think it over.
- (C) Put it off for quite a while so I can think of a better solution.
- (D) Don't worry about it; things tend to take care of themselves.

58. During high school, I have been a leader in my group of friends:
- (A) Almost always.
- (B) Very often.
- (C) Often.
- (D) Sometimes, but not often.
- (E) Never.

59. How hard do you usually work at getting good grades in high school?

Ⓐ I work very hard.
Ⓑ I could work a little harder.
Ⓒ I could work a lot harder.
Ⓓ I don't have to work hard. I get good grades easily.

60. How many really close friends do you have?

Ⓐ I have a lot of them.
Ⓑ I have a few of them.
Ⓒ I have one really close friend.
Ⓓ I don't have any really close friends.

61. I find myself putting things off until the last minute:

Ⓐ Almost always.
Ⓑ Often.
Ⓒ Sometimes, but not often.
Ⓓ Rarely.
Ⓔ Never.

62. What is the highest level of formal education obtained by your mother or female guardian?

Ⓐ High school graduate or less.
Ⓑ Postsecondary school other than college.
Ⓒ Some college.
Ⓓ College degree.
Ⓔ Some graduate school or graduate degree.

63. What is the highest level of formal education obtained by your father or male guardian?

Ⓐ High school graduate or less.
Ⓑ Postsecondary school other than college.
Ⓒ Some college.
Ⓓ College degree.
Ⓔ Some graduate school or graduate degree.

USING THE FOLLOWING SCALE, INDICATE THE EXTENT TO WHICH EACH OF THE STATEMENTS BELOW APPLIES TO YOU.

	VERY CHARACTERISTIC OF ME	SOMEWHAT CHARACTERISTIC OF ME	SLIGHTLY CHARACTERISTIC OF ME	SOMEWHAT UNCHARACTERISTIC OF ME	VERY UNCHARACTERISTIC OF ME
64. When I believe strongly in something, I act on it.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
65. I meet my obligations on time.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
66. I believe a person's day should be planned ahead each morning.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
67. I feel that people who can't meet deadlines just aren't organized enough.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
68. While I was growing up, I was encouraged to continue my education beyond high school.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
69. I believe that getting together with friends to "party" is one of life's important pleasures.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
70. Criticism makes me very uncomfortable.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
71. I think it's useless to plan too far ahead because things hardly ever come out the way you planned anyway.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
72. I complete projects on time by making steady progress.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
73. When an opportunity arises to have a good time, I take it and don't worry about the consequences.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
74. When I want to achieve something, I set subgoals and consider specific means for achieving those goals.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
75. It's hard for me to resist temptations.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ
76. I take risks to put excitement in my life.	Ⓐ	Ⓑ	Ⓒ	Ⓓ	Ⓔ

TURN PAGE AND CONTINUE

The following questions refer to the way in which certain important events in our society affect different people. Each item consists of a pair of alternatives lettered A or B. Please select the one statement of each pair (and only one) which you more strongly believe to be the case as far as you're concerned.

This is a measure of personal belief: Obviously there are no right or wrong answers. Try to respond to each item independently when making your choice; do not be influenced by your previous choices.

77. Ⓐ Many of the unhappy things in people's lives are partly due to bad luck.

Ⓑ People's misfortunes result from the mistakes they make.

78. Ⓐ In the long run, people get the respect they deserve in this world.

Ⓑ Unfortunately, an individual's worth often passes unrecognized no matter how hard he or she tries.

79. Ⓐ Without the right breaks, one cannot be an effective leader.

Ⓑ Capable people who fail to become leaders have not taken advantage of their opportunities.

80. Ⓐ Becoming a success is a matter of hard work, luck has little or nothing to do with it.

Ⓑ Getting a good job depends mainly on being in the right place at the right time.

81. Ⓐ The average citizen can have an influence in government decisions.

Ⓑ This world is run by the few people in power, and there is not much the little guy can do about it.

82. Ⓐ Most people don't realize the extent to which their lives are controlled by accidental happenings.

Ⓑ There is really no such thing as (good or bad) "luck".

83. Ⓐ In the case of the well prepared student there is rarely, if ever, such a thing as an unfair test.

Ⓑ Many times exam questions tend to be so unrelated to course work that studying is really useless.

84. Ⓐ It is hard to know whether or not a person really likes you.

Ⓑ How many friends you have depends on how nice a person you are.

85. Ⓐ In the long run the bad things that happen to us are balanced by the good ones.

Ⓑ Most misfortunes are the result of lack of ability, ignorance, laziness, or all three.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B – DESCRIPTIVE STATISTICS FOR MEAN INSERTION

Tables 26 – 33 summarize, by gender and race, the mean and descriptive statistics of all variables. These means were inserted into the data set for the thirty-eight cases that were missing variables.

Table 34 summarizes those cases that received data by way of mean insertion. It lists by gender, race, and class year each of the thirty-eight cases that were filled in this manner.

Table 26. Female Descriptives (SATVAVG, SATMAVG, TISSTD)

Variable	Statistics	White	Asian	Black	Other
satvavg	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	577.85	600.25	557.57	528.07
	Median	572.00	575.00	590.00	518.00
	Variance	4078.279	6911.643	5225.619	8584.225
	Std. Dev.	63.86	83.14	72.29	92.65
	Min	457	535	470	371
	Max	760	790	660	688
satmavg	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	654.34	681.38	573.86	613.79
	Median	652.00	683.00	599.00	603.50
	Variance	3192.39	1921.982	2625.143	4533.104
	Std. Dev.	56.5	43.84	51.24	67.33
	Min	500	625	500	505
	Max	800	740	630	737
tisstd	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	478.72	458.25	450.14	509.36
	Median	484.00	463.50	427.00	537.50
	Variance	6293.819	6207.929	6404.476	8783.016
	Std. Dev.	79.33	78.79	80.03	93.72
	Min	295	344	336	369
	Max	657	583	575	640

Table 27. Female Descriptives (CHM1PERF, MAT1PERF, SEM1AQPR, HS_RANK)

Variable	Statistics	White	Asian	Black	Other
chm1perf	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	5.20	5.50	4.00	5.00
	Median	4.00	5.00	4.00	4.00
	Variance	4.502	3.143	1.333	7.846
	Std. Dev.	2.12	1.77	1.15	2.8
	Min	0	4	2	2
	Max	12	8	6	12
mat1perf	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	6.70	6.00	5.43	7.21
	Median	6.00	6.00	6.00	6.00
	Variance	8.06	6.286	0.952	15.412
	Std. Dev.	2.84	2.51	0.98	3.93
	Min	0	3	4	2
	Max	16	9	6	16
sem1aqpr	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	2.87	2.71	2.45	2.62
	Median	2.88	2.49	2.44	2.38
	Variance	0.331	0.331	0.067	0.493
	Std. Dev.	0.5753	0.5754	0.2597	0.7023
	Min	1.38	2.06	1.94	1.87
	Max	4.00	3.60	2.81	4.00
hs_rank	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	635.65	659.38	592.86	572.00
	Median	627.00	675.00	625.00	550.00
	Variance	9888.324	16774.55	7648.81	11440.31
	Std. Dev.	99.44	129.52	87.46	106.96
	Min	401.00	450	450.00	401
	Max	800.00	800.00	700	775.00

Table 28. Female Descriptives (AC_PREP, HARDWORK, MATH_SEM, MIL_APT)

Variable	Statistics	White	Asian	Black	Other
ac_prep	Valid N	118	8	7	9
	Missing N	29	0	0	5
	Total N	147	8	7	14
	Mean	16.69	15.25	15.14	16.67
	Median	17.00	16.00	16.00	17.00
	Variance	4.901	9.071	8.810	11.000
	Std. Dev.	2.21	3.01	2.97	3.32
	Min	10	11	11	11
	Max	20	20	20	20
hardwork	Valid N	118	8	7	9
	Missing N	29	0	0	5
	Total N	147	8	7	14
	Mean	37.87	36.75	35.57	37.44
	Median	38.00	37.00	35.00	38.00
	Variance	5.531	4.786	7.286	4.528
	Std. Dev.	2.35	2.19	2.70	2.13
	Min	30	33	31	34
	Max	42	39	39	41
math_sem	Valid N	118	8	7	10
	Missing N	29	0	0	4
	Total N	147	8	7	14
	Mean	3.73	2.88	2.57	2.50
	Median	4.00	3.00	2.00	3.00
	Variance	1.191	2.411	2.952	1.389
	Std. Dev.	1.09	1.55	1.72	1.18
	Min	1	1	1	1
	Max	5	5	5	4
mil_apt	Valid N	118	8	7	10
	Missing N	29	0	0	4
	Total N	147	8	7	14
	Mean	11.51	10.63	11.71	9.90
	Median	12.00	11.50	12.00	10.50
	Variance	7.551	7.982	4.905	2.544
	Std. Dev.	2.75	2.83	2.21	1.60
	Min	3	6	8	7
	Max	15	15	14	12

Table 29. Female Descriptives (MAGRADED, PAGRADED, PRE_CALC, TOT_TOL)

		White	Asian	Black	Other
magraded	Valid N	117	8	7	10
	Missing N	30	0	0	4
	Total N	147	8	7	14
	Mean	3.79	3.38	3.00	3.70
	Median	4.00	3.50	3.00	4.00
	Variance	1.337	1.982	1.000	1.344
	Std. Dev.	1.16	1.41	1.00	1.16
	Min	1	1	1	1
	Max	5	5	4	5
pagraded	Valid N	118	8	7	10
	Missing N	29	0	0	4
	Total N	147	8	7	14
	Mean	4.08	3.50	3.29	4.00
	Median	4.00	4.00	4.00	4.00
	Variance	1.224	1.429	2.905	0.444
	Std. Dev.	1.11	1.20	1.70	0.67
	Min	1	1	1	3
	Max	5	5	5	5
pre_calc	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	81.26	83.88	68.86	72.57
	Median	82.00	87.00	65.00	71.00
	Variance	114.672	71.839	267.476	131.802
	Std. Dev.	10.71	8.48	16.35	11.48
	Min	43	67	49	58
	Max	100	93	91	93
tot_tol	Valid N	147	8	7	14
	Missing N	0	0	0	0
	Total N	147	8	7	14
	Mean	46.35	49.63	46.43	44.57
	Median	47.00	50.00	47.00	43.00
	Variance	36.187	28.554	33.619	44.725
	Std. Dev.	6.02	5.34	5.80	6.69
	Min	30	41	38	34
	Max	60	56	56	59

Table 30. Male Descriptives (SATVAVG, SATMAVG, TISSTD)

Variable	Statistics	White	Asian	Black	Other
satvavg	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	565.41	545.52	486.97	522.90
	Median	567.00	545.00	473.00	520.00
	Variance	6502.886	6842.444	5519.236	5847.135
	Std. Dev.	80.64	82.72	74.29	76.47
	Min	200	386	333	337
	Max	780	740	667	703
satmavg	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	660.81	635.17	573.06	606.85
	Median	662.50	630.00	580.00	608.00
	Variance	3229.58	3817.414	4170.114	3810.111
	Std. Dev.	56.83	61.79	64.58	61.73
	Min	450	515	415	435
	Max	800	800	700	730
tisstd	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	536.26	529.59	528.74	528.31
	Median	542.00	537.50	525.00	525.00
	Variance	7325.759	7854.531	5864.247	7357.278
	Std. Dev.	85.59	88.63	76.58	85.77
	Min	262	311	361	320
	Max	747	698	690	731

Table 31. Male Descriptives (CHM1PERF, MAT1PERF, SEM1AQPR, HS_RANK)

Variable	Statistics	White	Asian	Black	Other
chm1perf	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	5.61	5.50	3.81	4.41
	Median	6	6.00	4.00	4.00
	Variance	4.482	3.016	3.922	4.016
	Std. Dev.	2.12	1.74	1.98	2.00
	Min	0	2	0	0
	Max	12	12	12	12
mat1perf	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	7.33	6.61	5.36	5.99
	Median	6.00	6.00	6.00	6.00
	Variance	8.05	6.083	3.919	6.406
	Std. Dev.	2.84	2.47	1.98	2.53
	Min	0	2	0	0
	Max	20	16	12	12
sem1aqpr	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	2.90	2.76	2.38	2.56
	Median	2.94	2.80	2.33	2.69
	Variance	0.308	0.242	0.268	0.344
	Std. Dev.	0.5552	0.4918	0.5176	0.5861
	Min	0.00	1.56	1.00	0.88
	Max	4.00	3.81	3.75	3.75
hs_rank	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	586.53	580.69	525.95	568.74
	Median	575.00	587.50	525.00	575.00
	Variance	10970.01	10568.98	11684.29	9550.61
	Std. Dev.	104.74	102.81	108.09	97.73
	Min	400.00	400	400.00	400
	Max	800.00	800	800	800

Table 32. Male Descriptives (AC_PREP, HARDWORK, MATH_SEM, MIL_APT)

Variable	Statistics	White	Asian	Black	Other
ac_prep	Valid N	1022	57	58	86
	Missing N	212	7	19	11
	Total N	1234	64	77	97
	Mean	16.06	16.54	15.12	16.43
	Median	16.00	17.00	15.00	17.00
	Variance	6.073	5.538	7.862	7.189
	Std. Dev.	2.46	2.35	2.80	2.68
	Min	6	10	10	8
	Max	20	20	20	20
hardwork	Valid N	1012	57	60	86
	Missing N	222	7	17	11
	Total N	1234	64	77	97
	Mean	36.59	36.70	35.92	36.72
	Median	37.00	37.00	37.00	37.00
	Variance	7.943	6.142	13.400	8.345
	Std. Dev.	2.82	2.48	3.66	2.89
	Min	26	31	27	28
	Max	42	42	41	42
math_sem	Valid N	1019	56	59	86
	Missing N	215	8	18	11
	Total N	1234	64	77	97
	Mean	3.39	2.91	2.80	3.06
	Median	4.00	3.00	3.00	3.00
	Variance	1.618	1.756	2.165	2.22
	Std. Dev.	1.27	1.32	1.47	1.49
	Min	1	1	1	1
	Max	5	5	5	5
mil_apt	Valid N	1022	57	60	86
	Missing N	212	7	17	11
	Total N	1234	64	77	97
	Mean	12.28	12.28	11.23	12.20
	Median	13.00	13.00	12.00	13.00
	Variance	5.287	5.777	10.385	6.349
	Std. Dev.	2.3	2.4	3.22	2.52
	Min	5	6	4	7
	Max	15	15	15	15

Table 33. Male Descriptives (MAGRADED, PAGRADED, PRE_CALC, TOT_TOL)

Variable	Statistics	White	Asian	Black	Other
magraded	Valid N	1021	57	60	86
	Missing N	213	7	17	11
	Total N	1234	64	77	98
	Mean	3.55	3.46	3.40	2.74
	Median	4.00	4.00	4.00	3.00
	Variance	1.793	1.895	1.837	2.004
	Std. Dev.	1.34	1.38	1.36	1.42
	Min	1	1	1	1
	Max	5	5	5	5
pagraded	Valid N	1023	57	59	86
	Missing N	211	7	18	11
	Total N	1234	64	77	98
	Mean	3.92	3.65	3.02	2.98
	Median	4.00	4.00	3.00	3.00
	Variance	1.567	1.732	1.914	1.976
	Std. Dev.	1.25	1.32	1.38	1.41
	Min	1	1	1	1
	Max	5	5	5	5
pre_calc	Valid N	1234	64	77	97
	Missing N	0	0	0	0
	Total N	1234	64	77	97
	Mean	80.83	79.09	72.26	76.28
	Median	83.00	80.00	73.00	78.00
	Variance	126.913	107.166	129.142	171.015
	Std. Dev.	11.27	10.35	11.36	13.08
	Min	28	55	50	34
	Max	100	96	97	97
tot_tol	Valid N	1233	63	77	95
	Missing N	1	1	0	2
	Total N	1234	64	77	97
	Mean	47.80	46.98	44.27	44.79
	Median	49.00	48.00	44.00	45.00
	Variance	33.559	34.855	34.069	32.806
	Std. Dev.	5.79	5.9	5.84	5.73
	Min	23	26	30	29
	Max	59	56	57	59

Table 34. Mean Insertion Summary

Case	Gender	Race	ac _prep	Hard work	mil _apt	math _sem	ma graded	pa graded	tot _tol
1	male	white	1997						
2			1997						
3			1998						
4				1997					
5				1997					
6				1998	1998				
7				1999					
8				1999					
9				1999					
10				2000					
11				2000					
12				2000					
13				2000					
14				2000					
15				2000	2000				
16				2000	2000	2000			
17						1997			
18						1998			
19						1998			
20						1999			
21						2000			
22							1997		
23							1997		
24							1998	1998	
25							1999		
26								2000	
27									1998
28		asian				1998			
29									1998
30		black	1998						
31			1999						
32						2000			
33								2000	
34		otherace							1999
35									1999
36	female	white					1998		
37		otherace		1999					
38			2000						

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C – MODEL DEVELOPMENT FOR GRAD_ENG

A. INTRODUCTION

As described in Chapter IV, each model was estimated using an iterative process to finalize model specification. This Appendix summarizes that process for the dependent variable GRAD_ENG, the log odds that a midshipman will graduate with an engineering degree from the Naval Academy having initially chosen an engineering major.

The model with Personal History Questionnaire (PHQ) data is presented first, followed by the model without. The two models are compared in terms of variables included and goodness of fit. The comparison shows that PHQ data do not add to the predictability of GRAD_ENG. It is for this reason that the Finalized Model presented in Chapter V does not include PHQ data.

B. GRAD_ENG WITH PHQ DATA

This model specification includes data from the Personal History Questionnaire. A total of nineteen independent variables are included in the Initial Model as shown in Table 35, which summarizes the results for all iterations of the GRAD_ENG with PHQ data model specification.

From Table 35, examination of the initial model allowed the removal of HS_RANK, TOT_TOL, and all of the PHQ variables with the exception of MATH_SEM. Examination of the first iteration then removed MATH_SEM, completely eliminating all PHQ variables from the GRAD_ENG model.

Eleven variables remain in the final model. BLACK is not statistically significant but will be kept in the model due to its importance as a demographic variable. SEM1AQPR is only marginally significant. All other variables are highly significant.

Chi-square values for each model show that each is statistically reliable, and the increase in the -2 Log Likelihood from the initial model to the final model shows an increase in the model's predictive power.

Table 35. GRAD_ENG Model Iteration (N=1,371)

Independent Variable	Initial Model B	Initial Model Sig	First Iteration B	First Iteration Sig	Final Model B	Final Model Sig
FEMALE	-0.635	0.006	-0.649	0.003	-0.665	0.003
BLACK	-0.137	0.676	-0.096	0.763	-0.078	0.806
ASIAN	-0.689	0.033	-0.699	0.029	-0.667	0.037
OTHERACE	-0.652	0.018	-0.688	0.011	-0.681	0.012
HS_RANK	-0.001	0.542	-	-	-	-
SATMAVG	0.004	0.017	0.005	0.010	0.004	0.014
SATVAVG	-0.003	0.010	-0.003	0.012	-0.003	0.012
PRE_CALC	0.019	0.044	0.019	0.031	0.018	0.039
TOT_TOL	0.002	0.888	-	-	-	-
TISSTD	0.004	0.000	0.004	0.000	0.004	0.000
MAT1PERF	0.202	0.000	0.204	0.000	0.203	0.000
CHM1PERF	0.134	0.060	0.137	0.039	0.141	0.033
SEM1AQPR	0.488	0.091	0.452	0.110	0.437	0.122
HARDWORK	-0.029	0.405	-	-	-	-
MIL_APT	-0.018	0.595	-	-	-	-
AC_PREP	0.019	0.620	-	-	-	-
MATH_SEM	-0.074	0.243	-0.081	0.188	-	-
MAGRADED	0.062	0.344	-	-	-	-
PAGRADED	0.014	0.844	-	-	-	-
Constant	-5.171	0.003	-6.241	0.000	-6.255	0.000
Chi-Square	265.107	(df = 19)	262.111	(df = 12)	260.367	(df = 11)
-2 Log Like	1055.000		1057.995		1059.739	
Model Sig	0.000		0.000		0.000	

Table 36 summarizes the observed results from the Final Model of Table 35. It is presented in comparison to Table 16 in Chapter V. The independent variables are listed in order of their contribution to the total model, as determined by each variable's contribution to the model chi-square and R^2.

Table 36. Finalized Logistic Regression Model for Graduating with an Engineering Degree

			Chi-Square Contribution		R^2 Contribution	
Independent Variable	B	Sig	Additional	Cumulative	Additional	Cumulative
MAT1PERF	0.203	0.000	160.124	160.124	0.1783	0.1783
TISSTD	0.004	0.000	32.849	192.973	0.0341	0.2124
CHM1PERF	0.141	0.033	27.240	220.213	0.0276	0.2400
SATMAVG	0.004	0.014	12.842	233.055	0.0128	0.2529
FEMALE	-0.665	0.003	7.910	240.966	0.0079	0.2607
SATVAVG	-0.003	0.012	4.673	245.638	0.0046	0.2653
OTHERACE	-0.681	0.012	4.815	250.453	0.0047	0.2701
ASIAN	-0.667	0.037	4.406	254.859	0.0043	0.2744
PRE_CALC	0.018	0.039	2.982	257.841	0.0029	0.2773
SEM1AQPR	0.437	0.122	2.466	260.307	0.0024	0.2797
BLACK	-0.078	0.806	0.060	260.367	0.0001	0.2798
Constant	-6.255	0.000				
-2 Log Likelihood	1059.739					
Model Significance	0.000					

Note: Value given for R^2 is the Nagelkerke R^2.

Note: This model includes PHQ data. (N=1,371)

C. GRAD_ENG WITHOUT PHQ DATA

This model specification does not include data from the Personal History Questionnaire. Table 37 summarizes the results for all iterations of the GRAD_ENG without PHQ data model specification.

From Table 37, examination of the initial model allowed the removal of HS_RANK and TOT_TOL. The remaining variables are identical to those of the final model with PHQ data.

Eleven variables remain in the final model. BLACK is not statistically significant but will be kept in the model due to its importance as a demographic variable. CHM1PERF is significant at ten percent. All other variables are highly significant.

Chi-square values for each model show that each is statistically reliable, and the increase in the -2 Log Likelihood from the initial model to the final model shows an increase in the model's predictive power.

Table 37. GRAD_ENG Model Iteration (N=1,648)

Independent Variable	Initial Model B	Initial Model Sig	Final Model B	Final Model Sig
FEMALE	-0.595	0.004	-0.612	0.003
BLACK	-0.119	0.680	-0.121	0.675
ASIAN	-0.659	0.031	-0.666	0.029
OTHERACE	-0.522	0.041	-0.527	0.038
HS_RANK	0.000	0.555	-	-
SATMAVG	0.004	0.009	0.004	0.010
SATVAVG	-0.004	0.002	-0.004	0.001
PRE_CALC	0.024	0.003	0.024	0.003
TOT_TOL	-0.003	0.834	-	-
TISSTD	0.004	0.000	0.004	0.000
MAT1PERF	0.151	0.002	0.151	0.002
CHM1PERF	0.114	0.084	0.109	0.079
SEM1AQPR	0.647	0.014	0.634	0.014
Constant	-6.124	0.000	-6.310	0.000
Chi-Square	300.596	(df = 13)	300.212	(df = 11)
-2 Log Like	1254.208		1254.592	
Model Sig	0.000		0.000	

D. COMPARISON OF FINAL GRAD_ENG LOGISTIC REGRESSIONS

Two models of GRAD_ENG have been estimated. The first model included PHQ data while the second did not. Of the initial independent variables in each estimation, the same eleven independent variables make up the final model estimations. None of these eleven variables were of the six PHQ variables. In the equation estimating the log odds ratio of graduating with an engineering degree as opposed to not graduating with an engineering degree, the Personal History Questionnaire adds no value.

The chi-square and –2 log likelihood of the model that initially excluded PHQ data were 300.212 and 1254.592, respectively. The chi-square and –2 log likelihood for the model including PHQ data were 260.367 and 1059.739, respectively. The model excluding PHQ data is more predictive than the model that includes the data.

APPENDIX D – MODEL DEVELOPMENT FOR OVER3.30

A. INTRODUCTION

As described in Chapter IV, each model was estimated using an iterative process to finalize model specification. This Appendix summarizes that process for the dependent variable OVER3.30, the log odds that a midshipman will graduate with an engineering degree and a CQPR >= 3.30 from the Naval Academy.

The model with Personal History Questionnaire (PHQ) data is presented first, followed by the model without. The two models are compared in terms of variables included and goodness of fit. The comparison shows that PHQ data do not add to the predictability of OVER3.30.

B. OVER3.30 WITH PHQ DATA

This model specification includes data from the Personal History Questionnaire. A total of nineteen independent variables are included in the Initial Model as shown in Table 38, which summarizes the results for all iterations of the OVER3.30 with PHQ data model specification.

From Table 38, examination of the initial model allowed the removal of SATVAVG, TOT_TOL, and all of the PHQ variables with the exception of MIL_APT and PAGRADED. Examination of the first iteration allowed the removal of SATMAVG and MIL_APT. Following the second iteration, PAGRADED was removed. All PHQ variables were eliminated from the OVER3.30 model.

Ten variables remain in the final model. None of the four demographic variables is statistically significant, however, each will remain within the model. The remaining six variables prove to be highly significant.

Chi-square values for each model show that each is statistically reliable, and the increase in the -2 Log Likelihood from the initial model to the final model shows an increase in the model's predictive power.

Table 38. OVER3.30 Model Iteration (N=1,371)

Independent Variable	Initial Model B	Initial Model Sig	First Iteration B	First Iteration Sig	Second Iteration B	Second Iteration Sig	Final Model B	Final Model Sig
FEMALE	-0.154	0.676	-0.117	0.748	-0.051	0.884	-0.076	0.828
BLACK	-1.813	0.168	-1.843	0.155	-1.882	0.147	-1.844	0.159
ASIAN	-0.646	0.190	-0.632	0.198	-0.592	0.218	-0.575	0.233
OTHERACE	-0.423	0.390	-0.396	0.418	-0.444	0.358	-0.326	0.492
HS_RANK	0.003	0.002	0.003	0.002	0.003	0.000	0.003	0.001
SATMAVG	0.004	0.073	0.003	0.191	-	-	-	-
SATVAVG	-0.001	0.525	-	-	-	-	-	-
PRE_CALC	0.024	0.093	0.021	0.120	0.027	0.027	0.027	0.028
TOT_TOL	-0.006	0.807	-	-	-	-	-	-
TISSTD	0.003	0.006	0.003	0.007	0.003	0.006	0.003	0.005
MAT1PERF	0.221	0.000	0.226	0.000	0.236	0.000	0.232	0.000
CHM1PERF	0.252	0.001	0.251	0.000	0.260	0.000	0.264	0.000
SEM1AQPR	3.495	0.000	3.431	0.000	3.384	0.000	3.338	0.000
HARDWORK	0.037	0.386	-	-	-	-	-	-
MIL_APT	-0.060	0.147	-0.055	0.178	-	-	-	-
AC_PREP	0.030	0.522	-	-	-	-	-	-
MATH_SEM	-0.030	0.716	-	-	-	-	-	-
MAGRADED	-0.039	0.652	-	-	-	-	-	-
PAGRADED	-0.106	0.230	-0.129	0.103	-0.109	0.161	-	-
Constant	-23.377	0.000	-21.203	0.000	-20.699	0.000	-20.957	0.000
Chi-Square	763.082	(df = 19)	759.572	(df = 13)	756.018	(df = 11)	754.067	(df = 10)
-2 Log Like	707.520		711.030		714.585		716.536	
Model Sig	0.000		0.000		0.000		0.000	

Table 39 summarizes the results from the Final Model of Table 38. It is presented in comparison to Table 18 in Chapter V. The independent variables are listed in order of their contribution to the total model, as determined by each variable's contribution to the model chi-square and R^2.

Table 39. Finalized Logistic Regression Model for Engineering and a CQPR >= 3.30

			Chi-Square Contribution		R^2 Contribution	
Independent Variable	B	Sig	Additional	Cumulative	Additional	Cumulative
SEM1AQPR	3.338	0.000	644.610	644.610	0.5702	0.5702
MAT1PERF	0.232	0.000	51.843	696.452	0.0352	0.6054
CHM1PERF	0.264	0.000	24.958	721.410	0.0165	0.6219
HS_RANK	0.003	0.001	14.499	735.909	0.0094	0.6314
TISSTD	0.003	0.005	8.211	744.120	0.0053	0.6367
PRE_CALC	0.027	0.028	5.385	749.505	0.0035	0.6401
BLACK	-1.844	0.159	2.580	752.085	0.0017	0.6418
ASIAN	-0.575	0.233	1.452	753.537	0.0009	0.6427
OTHERACE	-0.326	0.492	0.483	754.019	0.0003	0.6430
FEMALE	-0.076	0.828	0.047	754.067	0.0000	0.6430
Constant	-20.957	0.000				
-2 Log Like	716.536					
Model Significance	0.000					

Note: Value given for R^2 is the Nagelkerke R^2.

Note: This model includes PHQ data. (N=1,371)

C. OVER3.30 WITHOUT PHQ DATA

This model specification does not include data from the Personal History Questionnaire. Table 40 summarizes the results for all iterations of the OVER3.30 without PHQ data model specification.

From Table 40, examination of the initial model allowed the removal of SATVAVG. Twelve variables remain in the final model. None of the four demographic variables are statistically significant, however, they will be left in the model. TOT_TOL is significant at ten percent; all other variables are highly significant.

Chi-square values for each model show that each is statistically reliable, and the increase in the -2 Log Likelihood from the initial model to the final model shows an increase in the model's predictive power.

Table 40. OVER3.30 Model Iteration (N=1,648)

Independent Variable	Initial Model		Final Model	
	B	Sig	B	Sig
FEMALE	-0.076	0.800	-0.067	0.824
BLACK	-1.092	0.212	-1.035	0.234
ASIAN	-0.485	0.284	-0.483	0.287
OTHERACE	-0.320	0.461	-0.294	0.497
HS_RANK	0.003	0.000	0.003	0.000
SATMAVG	0.005	0.025	0.004	0.052
SATVAVG	-0.001	0.214	-	-
PRE_CALC	0.032	0.008	0.033	0.007
TOT_TOL	-0.030	0.138	-0.033	0.100
TISSTD	0.003	0.009	0.003	0.006
MAT1PERF	0.192	0.000	0.198	0.000
CHM1PERF	0.249	0.000	0.254	0.000
SEM1AQPR	3.193	0.000	3.126	0.000
Constant	-20.734	0.000	-20.804	0.000
Chi-Square	899.496	(df = 13)	897.950	(df = 12)
-2 Log Like	915.664		917.209	
Model Sig	0.000		0.000	

D. COMPARISON OF FINAL OVER3.30 LOGISTIC REGRESSIONS

Two models of OVER3.30 have been estimated. The first model included PHQ data while the second did not. Of the initial independent variables in each estimation, ten variables were common to both models while the model not including PHQ data possessed two more. Of these ten and twelve variables, respectively, none were of the six PHQ variables. In the equation estimating the log odds ratio of graduating with an engineering degree and a CQPR >= 3.30 as opposed to not graduating with an engineering degree and a CQPR >= 3.30, the Personal History Questionnaire adds no value.

The chi-square and –2 log likelihood of the model that initially excluded PHQ data were 897.950 and 917.209, respectively. The chi-square and –2 log likelihood for the model including PHQ data were 754.067 and 716.536, respectively. The model excluding PHQ data is more predictive than the model that includes the data.

APPENDIX E – MODEL DEVELOPMENT FOR CQPR_ENG

A. INTRODUCTION

As described in Chapter IV, each model was estimated using an iterative process to finalize model specification. This Appendix summarizes that process for the dependent variable CQPR_ENG, which is the predicted value for a midshipman's QPR upon graduation with an engineering major.

The model with Personal History Questionnaire (PHQ) data is presented first, followed by the model without. The two models are compared in terms of variables included and goodness of fit. The comparison shows that PHQ data do not add to the predictability of CQPR_ENG.

B. CQPR_ENG WITH PHQ DATA

This model specification includes data from the Personal History Questionnaire. Table 41 summarizes the results for all iterations of the CQPR_ENG with PHQ data model specification.

From Table 41, examination of the initial model allowed the removal of four PHQ variables: HARDWORK, AC_PREP, MAGRADED, and PAGRADED. Examination of the first iteration then removed SATVAVG, TOT_TOL, and MATH_SEM. Of the PHQ variables, MIL_APT remains in the final model and is highly significant.

Twelve variables remain in the Final Model. FEMALE is statistically insignificant but will be kept in the model due to its importance as a demographic variable. OTHERACE is only marginally significant but will also remain. All other variables are highly significant.

The R^2 value indicates that each model accounts for sixty-six percent of the variance in the data. In other words, each model fairly accurately predicts the CQPR of a graduating engineer. The R^2 value did not increase as insignificant variables were removed from the model indicating that the predictive power of the model did not improve appreciably. The F Statistic did, however, improve from the initial to the final model, indicating that the significance of the model as a whole improved as variables were removed.

Table 41. CQPR_ENG Model Iteration (N=1,371)

Independent Variable	Initial Model B	Initial Model Sig	First Iteration B	First Iteration Sig	Final Model B	Final Model Sig
FEMALE	-0.002	0.952	0.000	0.993	0.000	0.991
BLACK	-0.156	0.001	-0.161	0.000	-0.162	0.000
ASIAN	-0.113	0.008	-0.112	0.008	-0.110	0.009
NOT_WAB	-0.054	0.145	-0.054	0.141	-0.055	0.138
HS_RANK	0.001	0.000	0.001	0.000	0.001	0.000
SATMAVG	0.001	0.002	0.001	0.003	0.001	0.001
SATVAVG	0.000	0.255	0.000	0.259	-	-
PRE_CALC	0.003	0.011	0.003	0.013	0.002	0.032
TOT_TOL	-0.003	0.164	-0.003	0.156	-	-
TISSTD	0.000	0.007	0.000	0.008	0.000	0.012
MAT1PERF	0.022	0.000	0.022	0.000	0.021	0.000
CHM1PERF	0.040	0.000	0.041	0.000	0.037	0.000
SEM1AQPR	0.385	0.000	0.385	0.000	0.394	0.000
HARDWORK	0.001	0.791	-	-	-	-
MIL_APT	-0.008	0.029	-0.007	0.036	-0.008	0.023
AC_PREP	0.002	0.625	-	-	-	-
MATH_SEM	-0.007	0.265	-0.007	0.284	-	-
MAGRADED	0.001	0.940	-	-	-	-
PAGRADED	0.000	0.970	-	-	-	-
Constant	0.505	0.010	0.586	0.000	0.542	0.000
R^2	0.660		0.660		0.659	
Adjusted R^2	0.655		0.656		0.655	
F Statistic	112.089		142.378		177.456	

C. CQPR_ENG WITHOUT PHQ DATA

This model specification does not include data from the Personal History Questionnaire. Table 42 summarizes the observed results for all iterations of the CQPR_ENG without PHQ data model specification.

From Table 42, examination of the initial model allowed the removal of SATVAVG leaving twelve variables in the final model. FEMALE and OTHERACE are not statistically significant but will be kept in the model due to their importance as demographic data. All other variables are highly significant.

The R^2 indicates that each model accounts for sixty-six percent of the variance in the data. Similar to the model including PHQ data, the R^2 value did not increase when the insignificant variable SATVAVG was removed, which indicates that the predictive power of the model did not improve appreciably. The F Statistic did improve, indicating that the significance of the values obtained from the model also improved.

Table 42. CQPR_ENG Model Iteration (N=1,648)

Independent Variable	Initial Model B	Initial Model Sig	Final Model B	Final Model Sig
FEMALE	0.014	0.609	0.015	0.587
BLACK	-0.120	0.003	-0.122	0.002
ASIAN	-0.099	0.012	-0.100	0.012
OTHERACE	-0.039	0.255	-0.041	0.231
HS_RANK	0.001	0.000	0.001	0.000
SATMAVG	0.001	0.001	0.001	0.002
SATVAVG	0.000	0.210	-	-
PRE_CALC	0.003	0.000	0.003	0.000
TOT_TOL	0.005	0.006	-0.004	0.009
TISSTD	0.000	0.019	0.000	0.026
MAT1PERF	0.020	0.000	0.019	0.000
CHM1PERF	0.043	0.000	0.042	0.000
SEM1AQPR	0.389	0.000	0.394	0.000
Constant	0.464	0.000	0.476	0.000
R^2	0.662		0.662	
Adjusted R^2	0.659		0.659	
F Statistic	201.623		218.200	

D. COMPARISON OF FINAL CQPR_ENG LINEAR REGRESSIONS

Two models of CQPR_ENG have been estimated. The first included PHQ data; the second did not. The R^2 and F Statistic of the model including PHQ data were 0.659 and 177.456, respectively. The R^2 and F Statistic of the model excluding PHQ data were 0.662 and 218.200, respectively. The model excluding PHQ data is the more predictive of the two.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Benefield, L. D., Walker, W. F., Halpin, G., Halpin, G., & Trentham, L. (1996). Assessment and Quality in Higher Education: A Model with Best Practice. Paper presented at the Conference on Assessment and Quality of the American Association for Higher Education, June.

Candidate Information System (2002, April 17). Retrieved May 27, 2002 from http://candidatedata.usna.edu

Castenada, G. G., & Winer, J. L. (1985). Psychological Models of Engineering Careers: Academic Prediction. Paper presented at the Annual Meeting of the Southwestern Psychological Association, April.

Durio, H. F., Kildow, C. A., & Slover, J. A. T. (1980). Ethnicity and Sex Differences in Use of College Entrance Examinations, Mathematics Achievement, and High School Rank as Predictors of Performance and Retention Among Engineering Students. Paper presented at the Annual Meeting of the American Educational Research Association, April.

Fletcher, J. T., Halpin, G., & Halpin, G. (1999). High School and College Grades: Is Past Performance a Predictor of Future Performance? Paper presented at the Annual Meeting of the Mid-South Educational Research Association, November.

Hayden, D. C., & Halloway, E. L. (1985). A Longitudinal Study of Attrition Among Engineering Students. *Engineering Education, Vol. 75, No. 7,* pp. 664-668.

Hestenes, D., Wells, M., and Swackhammer, G. (1992). Force Concept Inventory. *The Physics Teacher, Vol. 30,* pp. 141-157.

LeBold, W. K. (1958). Intellectual and Non-Intellectual Factors Involved in Predicting Engineering Success. *Journal of Engineering Education, Vol. 48, No. 7,* pp. 514-519.

Levin, J., & Wychoff, J. (1987). *Identification of Predictors of Persistence and Success in Baccalauareate Engineering: Implications for Academic Advising.* Pennsylvania State University, University Park. Division of Undergraduate Studies.

Levin, J., & Wychoff, J. (1990). *Identification of Student Characteristics that Predict Persistence and Success in an Engineering College at the End of Sophomore Year: Informing the Practice of Academic Advising.* Pennsylvania State University, University Park. Division of Undergraduate Studies.

MacGuire, S., & Halpin, G. (1995). Factors Related to Persistence in Engineering: Results of a Qualitative Study. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, November.

Muchinsky, P. M., & Hoyt, D. P. (1973). Predicting College Grades of Engineering Graduates from Selected Personality and Aptitude Variables. *Education and Psychological Measurement, Vol. 33, No. 4,* pp. 935-937.

O'Connor, C. A., & McAnulty, B. H. (1981). The Predictive Ability of ACT Scores for Student Success at an Engineering School. *Measurement and Evaluation in Guidance, Vol. 14, No. 2*, pp. 54-60.

Pascarella, E. T., Duby, P. B., Miller, V. A., & Rasher, S. P. (1981). Preenrollment Variables and Academic Performance as Predictors of Freshman Year Persistence, Early Withdrawal, and Stopout Behavior in an Urban, Nonresidential University. *Research in Higher Education, Vol. 15,* pp. 329-349.

Roueche, J. E., & Sims, D. M. (1968). Open-Door College or Open-Door Curriculums. *Junior College Journal, Vol. 38, No. 5, February.*

Sheppard, T. A. (2001). A Validation of the Strong Campbell Interest Inventory as Part of the Admissions Process at the United States Naval Academy (Masters thesis, Naval Postgraduate School, Monterey, CA).

Shoemaker, J. S. (1986). Predicting Cumulative and Major GPA of UCI Engineering and Computer Science Majors. Paper presented at the Annual Meeting of the American Educational Research Association, April.

U.S. Naval Academy, Academic Dean Instruction 1531.202, *The Majors Program: Class of 2002*, W. C. Miller, 01 September 1999.

Wankat, P. C. (1986). Current Advising Practices and How to Improve Them. *Engineering Education, Vol. 76*, pp. 213-216.

Watson, A. W. (2001). Identifying Midshipmen for Academic Assistance Using Entry Variables (Masters thesis, Naval Postgraduate School, Monterey, CA).

Weinstein, C. E., Palmer, D. R., & Schulte, A. C. (1987). Learning and Studies Strategies Inventory. Clearwater, FL: H&H Publishing Company, Inc.

Yess, J. P. (1979). *Predicting the Academic Success of Community College Students in Specific Programs of Study*. Unpublished doctoral dissertation, Boston University.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Nimitz Library
   United States Naval Academy
   Annapolis, MD

4. Superintendent
   U. S. Naval Academy
   Annapolis, MD

5. Office of Institutional Research, Planning, and Assessment
   U. S. Naval Academy
   Annapolis, MD

6. Professor Alice Crawford
   Naval Postgraduate School
   Monterrey, CA

7. Professor Roger Little
   U. S. Naval Academy
   Annapolis, MD

8. LT Nicholas A. Kristof
   Chester, MD

9. Mr. and Mrs. Zoltan J. Kristof
   Pittsburgh, PA