

PREDICTING ATTRITION IN THE MILITARY OCCUPATIONAL SPECIALTY
PHASE OF AN ARMY SPECIALTY-TRAINING PROGRAM

by

Amanda R. Webb

Doctoral dissertation submitted to the Faculty of the
Medical and Clinical Psychology Graduate Program
Uniformed Services University of the Health Sciences
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy, 2018



UNIFORMED SERVICES UNIVERSITY, SCHOOL OF MEDICINE GRADUATE PROGRAMS
Graduate Education Office (A 1045), 4301 Jones Bridge Road, Bethesda, MD 20814



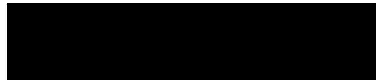
APPROVAL OF THE DOCTORAL DISSERTATION IN THE DEPARTMENT OF
MEDICAL AND CLINICAL PSYCHOLOGY

Title of Dissertation: "Predicting Attrition in the MOS Phase of an Army Specialty-Training Program"

Name of Candidate: Amanda R. Webb
Doctor of Philosophy Degree
August 24, 2017

DISSERTATION AND ABSTRACT APPROVED:

DATE:



8/29/17

Tracy Sbrocцо, PhD
DEPARTMENT OF MEDICAL & CLINICAL PSYCHOLOGY
Committee Chairperson



8/29/17

Neil E. Grunberg, PhD
DEPARTMENT OF MILITARY & EMERGENCY MEDICINE
Dissertation Advisor



8/29/17

Marjan G. Holloway, PhD
DEPARTMENT OF MEDICAL & CLINICAL PSYCHOLOGY
Committee Member



8/29/17

Cara H. Olsen, DrPH
DEPARTMENT OF PREVENTIVE MEDICINE & BIOSTATISTICS
Committee Member

ACKNOWLEDGMENTS

I wish to acknowledge several individuals who contributed their time and effort to the completion of this research project. First and foremost, I would like to offer sincere appreciation to my research advisor, Dr. Neil Grunberg, for his mentorship, guidance, and support. His confidence and trust in me to pursue my military and academic passions have allowed me to identify areas of research that I am truly excited about. I also am deeply appreciative of the opportunity to participate in Operation Bushmaster and the training of military medical students. Those experiences have helped me cultivate my professional identity and for the first time since coming to USUHS, I can envision the type of leader and clinician I want to be for the rest of my career.

I would like to extend genuine gratitude to Drs. Tracy Sbrocco and Marjan Holloway for their insightful feedback and continuous support. You both have served as important role models during my time at USUHS. A special thanks to Dr. Cara Olsen who taught me to embrace statistics and appreciate the challenges and rewards of good statistical analyses. To our laboratory biostatistician Erin Barry, for providing statistical expertise and methodological support to this project.

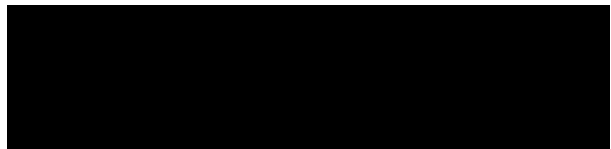
I also wish to recognize my mother, Marjorie Luczak, who always knows the perfect thing to say when inspiration or encouragement is needed. I am reminded of one particular statement, “when the going gets tough...suck it up,” that motivated me during many trying times. I will forever be grateful to her for holding me to high standards and for encouraging me to achieve more than I thought possible. To my husband Robert who inspired this work, our marriage and friendship remind me daily what it means to be part of a team. Your unwavering support, patience, and encouragement have been invaluable

over the last 15 years. You make the challenges bearable and the rewards infinitely sweeter.

To my USUHS cohort, thank you for your friendship and fellowship. Now that we've emerged from the trenches, I look forward to working with you downrange. A special thanks to the U.S. Army who entrusted me with the important responsibility of being a leader, officer, clinician, and researcher. My deepest gratitude to 1st Special Warfare Training Group, in particular, LTC Art Finch, MAJ Edwin Deagle, Dr. Luanne Kent, and the brave soldiers that participated in the study who provided me with the opportunity to engage in operational experiences and research that I am genuinely passionate about. Finally, to USUHS and the Department of Medical and Clinical Psychology, thank you for providing me with the opportunity to learn and grow as a military psychologist.

COPYRIGHT STATEMENT

The author hereby certifies that the use of any copyrighted material in the dissertation manuscript entitled: *Predicting Attrition in the Military Occupational Specialty Phase of an Army Specialty-Training Program* is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.



Amanda R. Webb

May 18, 2018

ABSTRACT

Title of Dissertation: Predicting Attrition in the Military Occupational Specialty Phase of an Army Specialty-Training Program

Amanda R. Webb, Doctor of Philosophy, 2018

Thesis directed by: Neil E. Grunberg, Professor, Military and Emergency Medicine

Identifying personnel with the right attributes to perform high-risk specialized military duties is critical for individual, team, and unit effectiveness. A large body of research has identified a core set of characteristics common among high-risk personnel and special operators who successfully complete rigorous assessment and selection (A&S) courses (79). However, the attributes required to succeed at later stages of training have not been identified. The present study examined whether demographic, performance, and psychological data collected during A&S could predict training outcome (*i.e.*, graduation and attrition) during the military occupational specialty (MOS) phase of a specialized training course. To address this research gap, data from 1,460 male, active duty, enlisted Army candidates who attended an Army specialty-training program between fiscal years 2014 and 2017 were analyzed using multinomial logistic regression and classification and regression tree (CART) analyses. Study findings revealed that the modeling strategies: [a] identified different variables for each MOS; [b] produced different findings between models for each MOS; and [c] had high rates of

misclassification. There are several possible explanations for these findings including: [a] a need for more specific and sensitive assessment instruments, [b] attributes other than those examined are associated with MOS training outcome, and/or [c] the unit's current MOS classification procedures are as effective as possible. The present study also identified the need for thorough and careful data management strategies. Dedicated database managers, quality assurance procedures, and appropriate statistical analyses are important for meaningful interpretation of the vast amounts of data collected from military personnel.

TABLE OF CONTENTS

LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER 1: Introduction	1
Purpose.....	1
Background	4
Academic Institutions – Admissions Procedures.....	4
Colleges and Universities	4
Military Academies.....	6
History of Military Assessment, Selection, and Classification.....	9
World War I (WWI).....	10
World War II (WWII).....	11
Post-Vietnam War.....	13
Projects A and B.....	15
Military Selection, Assessment, and Classification in the 21st century	16
Selection and Classification from Recruiter to Military Entrance Processing Station (MEPS)	17
The Introduction of Non-cognitive Measures.....	18
Maximizing Noncommissioned Officer (NCO) Performance for the 21st Century and Army Class Effort	20
Attrition	22
Talent Management/Competency.....	24
Assessing High-Risk Personnel	27
Cognitive and Non-cognitive Factors	27
Present Project: Predicting Attrition in an Army Specialty-Training Program	30
CHAPTER 2: Specific Aims and Hypotheses	32
Overview and Specific Aims	32
Specific Aim 1	32
Hypothesis 1a.....	32
Hypothesis 1b.....	32
Hypothesis 1c.....	32
Specific Aim 1 Rationale	33
Specific Aim 2	33
Hypothesis 2.....	34
Specific Aim 2 Rationale	34
CHAPTER 3: Methods	35
Overview.....	35
Research Design and Procedures	35
Data Management	35
Participants.....	36

Procedures	36
Assessment Measures	37
Outcome Variables.....	41
Data Analytic Plan	42
Missing Data and Software Analysis.....	42
Data Management	42
Sample and Power.....	42
Specific Aim 1	43
Specific Aim 2	44
CHAPTER 4: Results	47
Data Management Results	47
Data Cleaning.....	47
Entrance General Technical (GT) Scores	48
Age.....	49
Years of Education (YRS ED).....	49
Multidimensional Aptitude Battery (MAB).....	49
Training MOS and Training Outcome	50
Army Physical Fitness Test (APFT).....	50
Quality Assurance.....	50
Missing Data, Skewness, Kurtosis, and Multicollinearity	51
Variable Selection.....	53
Intelligence and Education Variables	53
Personality Variables	54
Specific Aim 1	55
Descriptive Data Analyses.....	56
Sample Characteristics.....	56
Weapons Experts.....	58
Engineer Experts	58
Medical Experts.....	59
Communications Experts	60
Summary of Descriptive Statistical Analysis Findings	60
Model Building	60
Weapons Expert.....	62
Engineer Expert	62
Medical Expert.....	62
Communications Expert.....	63
Model Validation	63
Weapons Expert.....	63
Engineer Expert	64
Medical Expert.....	64
Communications Expert.....	65
Specific Aim 1 - Multinomial Logistic Regression Summary.....	65
Hypothesis 1a	66
Hypothesis 1b	66
Hypothesis 1c	66

Specific Aim 2	66
Classification and Regression Tree Development	67
Weapons Experts	68
Engineer Experts	69
Medical Experts	70
Communications Experts	70
Additional Analyses	70
Specific Aim 2 – CART Summary	72
Hypothesis 2	72
CHAPTER 5: Discussion.....	73
Purpose.....	73
Variables of Interest.....	74
Study Findings	74
Multinomial Logistic Regression.....	75
Classification and Regression Tree (CART)	77
Summary of Study Findings	78
Possible Explanation Of Findings.....	80
Current Instruments	80
Intelligence Quotient (IQ) and Personality	80
Motivation.....	81
Unassessed Attributes	83
Emotional/Social Intelligence.....	83
Peer Evaluations.....	84
Instructor Evaluations	86
Current Assignment Procedures are as Effective as Possible.....	86
Summary	87
Study Strengths, Limitations And Future Directions.....	87
Strengths	87
Limitations and Future Directions	88
CHAPTER 6: Summary.....	93
CHAPTER 7: Conclusion.....	95
APPENDIX A: Tables	96
APPENDIX B: Figures	144
APPENDIX C: Glossary.....	154
REFERENCES	159

LIST OF TABLES

Table 1. The Armed Services Vocational Aptitude Battery (ASVAB) test	96
Table 2. Army ASVAB Line Scores (aptitude composites)	97
Table 3. Predictor Variables	98
Table 4. Outcome Variables	98
Table 5. 60% and 40% Subsample Counts	99
Table 6. Missing Data by Variable: Weapons Experts	99
Table 7. Missing Data by Variable: Engineer Experts	100
Table 8. Missing Data by Variable: Medical Experts	101
Table 9. Missing data by Variable: Communications Experts	102
Table 10. Skewness and Kurtosis: Demographic and Performance Variables	103
Table 11. Skewness and Kurtosis: Psychological Variables	104
Table 12. Skewness and Kurtosis: Psychological Variables	105
Table 13. Correlational Analyses: Cognitive/Intelligence Tests	106
Table 14. Correlational Analyses: Multidimensional Aptitude Battery (MAB-II)	107
Table 15. Correlational Analyses: Level of Education	108
Table 16. Correlational Analyses: Performance Variables	108
Table 17. Correlational Analyses: Personality Variables	109
Table 18. Final Predictor Variables	110
Table 19. Outcomes: All Military Occupational Specialties (MOS)	111
Table 20. Outcomes: Medical Experts	111
Table 21. Demographic Continuous Variables (All MOS)	112
Table 22. Demographic Categorical Variables (All MOS)	112
Table 23. Performance Variables (All MOS)	112
Table 24. Psychological Variables	112
Table 25. Psychological Variables	112
Table 26. Age (<i>Mean, SD</i>)	113
Table 27. Active Duty Time in Service (ADTIS; <i>Mean, SD</i>)	114
Table 28. Grade/Rank (<i>Mode, Valid%</i>)	115
Table 29. Previous MOS (PrevMOS; <i>Mode, Valid%</i>)	116
Table 30. MOS Preferences (MOSPref; <i>Mode, Valid%</i>)	116
Table 31. Years of Education (YRS ED; <i>Mean, SD</i>)	117
Table 32. Army Physical Fitness Test score (APFT; <i>Mean, SD</i>)	118
Table 33. Ruck/Forced March (<i>Mean, SD</i>)	119
Table 34. Land Navigation (Land NAV; <i>Mean, SD</i>)	120
Table 35. Verbal IQ (VIQ; <i>Mean, SD</i>)	121
Table 36. Performance IQ (PIQ; <i>Mean, SD</i>)	122
Table 37. Emotional/Internalizing Dysfunction (EID; <i>Mean, SD</i>)	123
Table 38. Thought Dysfunction (THD; <i>Mean, SD</i>)	124
Table 39. Behavioral/Externalizing Dysfunction (BXD; <i>Mean, SD</i>)	125
Table 40. Somatic Complaints (RC1; <i>Mean, SD</i>)	126
Table 41. Cynicism (RC3; <i>Mean, SD</i>)	127
Table 42. Ideas of Persecution (RC6; <i>Mean, SD</i>)	128
Table 43. Aggressiveness - revised (AGGRr; <i>Mean, SD</i>)	129

Table 44. Introversion/Low Positive Emotionality - revised (INTRr; <i>Mean, SD</i>)	130
Table 45. Demographic Variables (Medical Experts)	131
Table 46. Performance Variables (Medical Experts; <i>Mean, SD</i>).....	131
Table 47. Cognitive/Intellectual (Medical Experts; <i>Mean, SD</i>)	132
Table 48. Personality Variables (Medical Experts, <i>Mean, SD</i>)	132
Table 49. Multinomial Logistic Regression Models: Weapons Experts	133
Table 50. Multinomial Logistic Regression Models: Engineer Experts.....	134
Table 51. Multinomial Logistic Regression Models: Medical Experts	135
Table 52. Multinomial Logistic Regression Models: Communications Experts	136
Table 53. Validation Model Classification Accuracy: Weapons Experts.....	137
Table 54. Validation Model Classification Accuracy: Engineer Experts	137
Table 55. Validation Model Classification Accuracy: Medical Experts	137
Table 56. Validation Model Classification Accuracy: Communications Experts	137
Table 57. CART Test Model: Weapons Experts	138
Table 58. CART Test Model: Engineer Experts.....	138
Table 59. CART Test Model: Medical Experts	138
Table 60. CART Test Model: Communications Experts.....	139
Table 61. CART Modified Priors: Weapons Experts	139
Table 62. CART Modified Priors: Engineer Experts.....	140
Table 63. CART Modified Priors: Medical Experts.....	141
Table 64. CART Modified Priors: Communications Experts.....	142
Table 65. Random Forests: Weapons Experts	142
Table 66. Random Forests: Engineer Experts.....	143
Table 67. Random Forests: Medical Experts.....	143
Table 68. Random Forests: Communications Experts.....	143

LIST OF FIGURES

Figure 1. Institutional Review Board (IRB) approval letter	144
Figure 2. Womack Army Medical Center IRB exemption letter	145
Figure 3. Data sharing agreement	146
Figure 4. Classification and Regression Tree (CART): Weapons experts	147
Figure 5. CART: Engineer expert.....	148
Figure 6. CART: Medical experts.....	150
Figure 7. CART: Communications experts	151

CHAPTER 1: Introduction

PURPOSE

Special operations forces (SOF) are the epitome of United States (U.S.) military. They are elite units charged with executing complex and dangerous operations across land, sea, and air to further the nation's operational and strategic objectives (33). Members of these elite units are adept at applying specialized skills across a broad range of environments with little to no external support earning them the reputation of being President Obama's, "military tool of choice" (92). The current operational tempo has drastically expanded the employment of SOF across the globe (3; 4; 63; 84). Approximately 8,000 SOF forces are deployed or forward stationed in 80 countries around the world highlighting their versatility and range (3; 4). Identifying the right personnel, and subsequently assigning them to the military occupational specialty (MOS) that best fit their individual attributes, is an important task.

A motto of Special Operations Command (SOCOM), *tip of the spear*, captures the role of these warfighters compared to conventional military forces. SOF missions are high-risk and failure can have serious national security implications. As a result, operators are considered the most technically and tactically proficient personnel in the military and their specialty training far surpasses their conventional force peers (79). Not only must candidates successfully pass the military's most challenging "selection" programs (*e.g.*, Basic Underwater Demolitions/SEAL [BUD/S] and Special Forces Assessment and Selection [SFAS]) (50), they also must complete a lengthy, highly specialized, and intense training pipeline before they are awarded an official badge of

completion (*e.g.*, Navy SEAL trident or Army Green Beret). From accession to graduation, production of a single operator can take as long as 66 weeks (more if a candidate recycles a phase of training), depending on the service branch. Consequently, high rates of attrition resulting from training failures or recycles are costly from both a time and financial perspective.

In SOF training pipelines, candidates undergo rigorous team and individualized training; in some cases, candidates are specially trained in a unique MOS (*e.g.*, medic). Given the time, resources, and rigor of these training courses, optimal fit between a candidate and job specialty is vital. Unlike conventional military forces, SOF units are typically task organized into smaller elements with some teams as small as 12 personnel. In these organizations, there is little redundancy; each operator becomes an authority for a specific skillset. For example, when there is only one school-trained medic on a 12-man team conducting covert operations in hostile territory, the role of the medic becomes paramount. Assigning candidates to an MOS for which they are not best suited can lead to high rates of attrition, dissatisfaction, and poor performance; in extreme cases, person-job mismatch can compromise national security missions and unnecessarily endanger personnel. For those reasons, optimizing MOS classification is a critical necessity.

The rising number of recycle and relief rates in an Army specialty-training program have prompted military leaders to examine the current MOS assignment process. **This doctoral research project addressed the organization's need to improve MOS assignment by identifying the variables that best predict successful graduation versus attrition (*i.e.*, recycle and relief) in an Army specialty-training program. Additionally, this project sought to develop an MOS assignment algorithm (*i.e.*, a**

classification tree) to aid Army specialty unit leaders with classification decisions.

Improved classification to a particular MOS should increase first time graduation rates and reduce overall attrition. Ultimately, improved MOS classification will optimize stewardship. The present project used data from 1,460 candidates in a specialty-training program to identify a classification and statistical regression model that best predicted MOS graduation and attrition.

A large number of civilian and military organizations utilize assessment and selection (A&S) models to identify personnel best suited for employment, admissions, and/or classification (14; 78; 79; 81; 83). In some cases, the methods employed by these organizations are applicable to the current project; in other cases, they are not. The Introduction summarizes the A&S methods and predictor variables used by these organizations with an emphasis on what procedures are pertinent to the current research project and what research gaps remain. The Introduction is divided into the following sections: [a] admissions procedures for civilian and military academic institutions, [b] military assessment, selection, and classification, and [c] A&S of high-risk military occupations. The Specific Aims section lists the aims and hypotheses examined by this research project. The Methods section presents the methodological processes and the data analytic plan. The Results section provides information about the sample characteristics and findings from the statistical regression and classification tree analyses. The Discussion section addresses findings from the statistical analyses, outlines study limitations, and provides recommendations for future directions. Then, the Summary and Conclusion are presented.

BACKGROUND

Academic Institutions – Admissions Procedures

Colleges and Universities

Selecting a diverse group of applicants with the highest likelihood of graduating has been an important goal of colleges and universities for many years. Although the type of training completed by military applicants differs substantially from the education and training offered by a four-year degree program, the goal of identifying the applicants with the highest likelihood of success is identical. College students are expected to demonstrate competency in broadly applicable core subjects as well as in their academic majors. Similarly, military recruits must demonstrate proficiency in military core competencies (*e.g.*, physical aptitude and general military education) as well as competency in their MOS (*i.e.*, their military academic major). Because the current project was conducted to try to improve assessment and assignment to MOSs, a review of admissions practices in academic institutions was a useful starting point.

Extensive research has examined factors most predictive of success in colleges and universities (14; 36; 94; 107). Admissions criteria vary by institution, but the scholastic aptitude test (SAT), American College Testing (ACT), and high school grade point average (HSGPA) are common elements reviewed by admissions committees (14; 70). Despite variability in the quality and academic rigor of different high schools, a large body of research indicates that HSGPA is the strongest and most consistent predictor of first year college GPA (14; 36; 107). Geiser and Santelices (36) examined the predictive validity of HSGPA, SAT, and ACT scores on four-year college outcomes using a sample of nearly 80,000 participants from the University of California from 1996 to 1999. The authors reported three key findings. First, HSGPA was the strongest

predictor of four-year outcomes, regardless of academic concentration. Second, HSGPA accounted for a greater proportion of variance in cumulative fourth-year grades compared to first-year grades. Third, socioeconomic status (SES) did not alter the predictive value of HSGPA.

Similarly, a study based on 50 universities and 189,612 college students from 2000 to 2006 found that ACT composite scores and HSGPA were highly associated with first-year college performance. First year performance was the strongest predictor of second- and third-year success (107). SES did not emerge as a predictor of academic success. There are no definitive explanations why HSGPA has been a stronger predictor of academic success than the SAT and ACT, but researchers speculate that HSGPA provides information about an individual's performance over time, whereas the SAT and ACT reflect a student's performance in a single 3 to 4 hour testing session (14; 36; 107).

Despite abundant literature examining the predictive validity of various academic, aptitude, and achievement markers, the military does not consider HSGPA, SAT, or ACT scores in enlistment decisions. In fact, although preferred, a high school diploma is not a requirement for military enlistment (71). Instead, the military primarily relies on scores from the Armed Services Vocational Aptitude Battery (ASVAB) to make enlistment and classification decisions. The ASVAB assesses a candidate's likelihood for success in the military whereas the SAT and ACT assess the likelihood of success in college (73). A more detailed account of the ASVAB is provided in the section below, "History of Military Assessment, Selection, and Classification." The approach used to determine entry into academic institutions, therefore, may not seem relevant to military MOS assignment; however, it highlights two important points: [a] identifying metrics to assess

candidate potential is important, and [b] the variables most indicative of success should not be assumed but should be assessed and validated. Consistent with these points, the present project was designed to identify the variables most predictive of success or relief in an Army specialty-training program. Another similar, and perhaps more comparable population, consists of applicants seeking entry into military educational institutions. Therefore, it is useful to examine the literature regarding selection at the U.S. Military Academies.

Military Academies

Graduation from a four-year civilian academic institution is only partially comparable to graduation from a military occupational training course. Perhaps more analogous are the admissions procedures for military academies. The U.S. Military Academies (Army, Navy, Air Force, Coast Guard, and Merchant Marine) are considered among the most prestigious institutions for producing well-educated and first-rate military officers. Acceptance to a military academy is highly competitive and requires an appointment by the President of the United States. As part of the application process, candidates must obtain either a Congressional nomination (*i.e.*, nomination by the Vice President, members of the U.S. Senate, or members of the House of Representatives) or a service connected/military affiliated nomination (*e.g.*, nominations for familial relationships with a service member) (2). Competition is steep and candidates are evaluated on several domains: character, scholarship, leadership, physical aptitude, medical fitness, and motivation. Collectively, these domains are referred to as the “whole person” (2). According to the current admissions webpage for West Point, the U.S. Army Military Academy, only 1,257 of 13,827 applicants were admitted in the last academic

year. Of the admitted applicants, nearly 92% had earned varsity letters (almost 67% of varsity athletes were varsity team captains), 70% were ranked in the top fifth of their high school class, and 85% had earned an academic honor (*e.g.*, valedictorian, salutatorian, National Merit scholarship, National Honor Society) (9). Admissions class profiles for the four other military academies reported similar admissions statistics (5-8).

Military academies are invested in training future generations of military leaders committed to a full and successful career in the military. The high cost of \$153,000 to \$228,500 to train a single academy officer (100) underscores the importance of selecting applicants who are likely to graduate and are committed to military service. To that end, the Undersecretary of Defense (Personnel and Readiness) requested the RAND National Defense Research Institute examine the admissions procedures for the military academies (40). At present, only findings from the Army's West Point are available.

Hanser and Oguz (40) examined the primary variables used by West Point to rank applicants. Using four models (three logistic regression models and a boosted logistic regression model), the authors assessed the relative predictive power of the following variables: **academic ability** (*i.e.*, college entrance exam rank [CEER] calculated from ACT/SAT scores and high school rank [HSRCS]), **leadership** (*i.e.*, athletic activities score [AAS], extracurricular activities score [EAS], and faculty appraisal score [FAS]), and **physical fitness** (CFA). Taken together, these variables represent the whole candidate score (WCS). Although essays also are evaluated as part of the admissions decision, they were not included in this research project.

The first model used logistic regression to examine the strength of WCS in predicting graduation. The second model examined the predictive validity of the CEER

separate from the other WCS components. The third model examined all WCS components separately; that is, CEER was broken down into ACT/SAT and HSRCS. The fourth model was a boosted regression, which used the same variables as model three. Using a boosted regression model to examine the variables allowed researchers to examine possible interaction effects (40).

The overarching goals of the project were to determine whether the WCS or components of the WCS were predictive of graduation, early promotion to the rank of O-4 (Major), or promotion to O-5 (Lieutenant Colonel). Early promotion to O-4 was defined as promotion within 10 years of commissioning. The sample consisted of the graduating classes from 1992 through 2001. The average graduation rate for the ten years studied was 76%. Results from model one indicated the WCS was significantly able to predict graduation rates such that every 1 standard deviation increase in WCS increased the probability of graduation by approximately 4.7%. Regression models examining the individual WCS components found the following variables significantly predicted graduation: CFA, AAS, EAS, FAS, and CEER. Model three examined the CEER components separately and only SAT/ACT emerged as significant. The boosted logistic regression model, included for its ability to identify interactions and create a better fitting model, indicated SAT/ACT scores, FAS, and HSRCS were most responsible for increasing the likelihood of graduation. The authors explain that HSRCS emerged as a strong predictor in model four, and not model three, because the logistic regression model was able to uncover an interaction effect (40).

To determine whether the WCS or components of the WCS could also predict performance, the authors used the same four regression models to examine early

promotion to O-4 and promotion to O-5. The regression models did not show any notable findings for promotion to O-4; however, the WCS emerged as significantly predictive of promotion to O-5. Overall, the authors concluded that the WCS was significantly associated with graduation and promotion to O-5 (40).

The RAND study offers evidence for the use of the WCS in admissions decisions. Components of the WCS that emphasize academic ability, leadership, and physical fitness accounted for 26% of the variance predicting successful graduation from West Point and 25% of the variance predicting subsequent promotion to O-5 (40). Given the success of WCS variables to predict officer performance, it is possible that similar variables may be equally effective to predict enlisted MOS attrition and graduation. Therefore, the present project included predictor variables that assess similar domains.

Although a review of the literature examining civilian and officer-producing institutions provided valuable information about the utility of certain predictor variables, the procedures were not directly comparable to enlisted selection, assessment, classification, and training programs. To address this discrepancy, the next section briefly reviews the history of military A&S and summarizes findings most relevant to the current project.

History of Military Assessment, Selection, and Classification

The A&S process differs among units and groups within the military; however, two of the central domains assessed by selection programs are **psychological** and **physical** functioning (35; 79; 82; 85; 102). Psychological and physical assessments are used for two main reasons: [a] to determine whether a candidate meets the minimum qualification standards to be “selected” for service (62), and [b] to “classify” candidates

in a particular MOS based on their demonstrated aptitude across a variety of domains (62). Ultimately, the goal of selection and classification is to identify personnel who meet the minimum standards and then assign them to the MOS for which they are best-suited and are most likely to succeed (85).

Military A&S has a long and storied history. This section reviews the history of military A&S and is organized chronologically beginning with World War I, followed by World War II, and finally the post-Vietnam War era. Important developments during these time periods are highlighted. The section concludes with a review of military A&S in the 21st century and emphasizes current procedures and research developments. The present project follows from this long tradition and was designed to contribute to the developing body of research geared to improve MOS classification.

World War I (WWI)

Psychological assessment, selection, and classification have played an instrumental role in the U.S. Army. The history of military testing can be traced back to WWI when psychological testing was used in an effort to reduce psychiatric casualties and improve assignment of recruits to military roles/duties (82; 85). Robert Yerkes (co-author of Yerkes-Dodson arousal-performance curve) is credited with directing the development of the Army Alpha and Beta test batteries, which were administered to more than 1.7 million men during the war (85). The Army Alpha assessed verbal abilities (*e.g.*, verbal skills, information/knowledge, and numerical abilities) and the Army Beta assessed non-verbal abilities (*e.g.*, picture completion) (73). Although these tests assessed different cognitive domains, the Beta version was designed for use with non-English speaking recruits and recruits with limited language skills. This subset of

candidates often lacked a formal education and experienced difficulty reading and writing, which precluded the use of the Army Alpha test (73). The Alpha and Beta tests represent the first efforts to psychologically screen military personnel and are considered among the earliest attempts at cognitive testing (85).

World War II (WWII)

There were many important developments in psychological assessment during WWII. Using the Army Alpha and Beta tests as a template, the Personnel Research Section developed the Army General Classification Test (AGCT) to assist with selection and classification of recruits (35; 85). Additionally, the Army developed tests to assess candidate aptitude in particular functional areas (*e.g.*, mechanical and clerical speed). Ten of these aptitude tests eventually became the Army Classification Battery (ACB) and were used to supplement the AGCT (80). In addition to the development of psychological tests, this era also gave rise to the first U.S. psychological assessment center, which was created to screen candidates for the Office of Strategic Services (OSS) (15). Experimental social psychologist Kurt Lewin, a pioneer in leadership and group dynamics research, was a consultant on the project (75).

The OSS grew out of an organization known as the Coordinator of Information (COI) with the prescribed mission to collect, analyze, and disseminate intelligence in support of the ongoing war effort (101). In 1942, the COI evolved into the OSS, a covert agency dedicated to intelligence gathering and paramilitary operations (82; 101). Members of the OSS were tasked with a multitude of high-stress missions that involved intelligence collection, subversion, espionage, and psychological warfare with some operations conducted in enemy-occupied territories (82; 101). OSS operations informed

the development of the Central Intelligence Agency (CIA) and Army Special Forces (SF). In particular, the paramilitary operations conducted by the OSS were the precursor to unconventional warfare techniques used by today's SF community and OSS is considered a "legacy unit" of U.S. Army Special Operations Forces (ARSOF; 99).

The extremely specialized and dangerous nature of OSS operations created a high-stress environment for operators (82). Many operators were often unaware and unprepared for the variability and demands of their assignment. Consequently, a high percentage of operatives developed psychiatric problems (75; 82). As described by the OSS Assessment Staff (75):

The organization has been recruiting too many men, civilian or military, who have intelligence and sometimes the necessary mechanical training but who lack common sense, know nothing about working with men or how to look after the welfare and the morale of men under them. We simply must have men who can shoulder responsibility and use initiative with common sense. Simply because a man has intelligence does not qualify him for this type of work. In some instances we also have had men who fall into the class of the high-strung or emotional type. We simply cannot use men of that type in the field when they have to live with Chinese, eat Chinese food, and be under pressure at times. In most cases these men have suffered nervous breakdowns and other nervous ailments. Whether men are recruited in the States or here in the field they must be checked by a doctor and a psychiatrist before being pronounced fit for the field. The check by a psychiatrist is especially desirable (p. 15).

Rising rates of battlefield casualties from psychiatric distress and growing concerns about the quality of candidates led to the creation of the first formal, methodological, psychological screening process in the U.S. military (75; 82). Given the overlap in OSS and ARSOF missions, the three-day A&S paradigm implemented by the OSS staff in 1948 remains the foundation upon which current SOF A&S programs are modeled (15; 78; 79; 82). Today, operational psychologists working in A&S billets are encouraged to read *The Assessment of Men (1948)*, which offers a detailed account of the OSS staff's A&S endeavors (28). **Nearly 70 years after the initial insights offered by the OSS staff, personnel selection and classification remain critically essential to the Army and was the focus of the present research project.**

Post-Vietnam War

In response to a directive by Congress in 1950 to create a joint selection and classification battery, the DoD developed the Armed Forces Qualification Test (AFQT) (62). Modeled after the AGCT, the AFQT allowed for the uniformed psychological screening of recruits, regardless of service branch (35; 62). However, after the Vietnam War, each service was given the option to administer service-specific tests, despite the existence of a joint service psychological test (35; 62). The use of service-specific assessment measures increased the burden on exam stations and, as a result, the services returned to a joint service A&S battery in 1976 – the Armed Service Vocational Aptitude Battery (ASVAB) (35; 62; 85). The AFQT remains in use as a psychological screener and the overall score is derived from two mathematics and two verbal subtests of the ASVAB (35; 62; 85). In addition to the AFQT “line score,” the ASVAB subtests can be combined to create other line scores that are used to determine if a recruit meets the

minimum classification requirements for particular MOS assignments. Line scores provide information about general cognitive aptitude and help inform classification decisions. Line scores also are referred to as “aptitude composites.” The general technical (GT) line score, used in the present study, and other ASVAB line scores are discussed below.

In the 1980s, substantial scoring and norming problems resulted in the selection of hundreds of thousands of personnel from the bottom third of eligible recruits (35; 62; 85). Researchers incorrectly assumed the ASVAB scores had the same meaning as the AFQT scores they replaced (62). When the ASVAB was introduced in 1976, marked differences in test norming conditions (*e.g.*, poor motivation, fatigue, and failing to include personnel with low reference test scores) resulted in the inflation of percentile scores (62). Inflated scaled scores, especially in the below average range, resulted in overestimates of candidate ability and led to the inappropriate selection of many low qualifying personnel (35). The ASVAB’s psychometric shortcomings brought national attention to the military’s selection and classification process (62; 85).

Despite a temporary reduction in troop quality, several positive changes emerged from the ASVAB’s psychometric problems. First, managers in the Office of the Assistant Secretary of Defense for Force Management and Personnel (OASD-FM&P) created a more accurate scale based on data from men and women; the previous scale was restricted to data gathered only from men. Second, a massive research effort was undertaken to validate the predictive ability of the ASVAB. Third, there was increased supervision of the testing process and a steering committee of flag officers and testing psychologists regularly reviewed the testing program (62; 85). Determining the

predictive validity of the ASVAB was no small task and each service engaged in independent research to address Congress' Joint-Service Job Performance Measurement/Enlistment Standards (JPM) mandate (35). Interpreting Congress' JPM mandate as the minimum requirement, the Army conducted extensive research activities, Projects A and B, to examine the entire selection and classification program (35).

Projects A and B

Considered one of the most influential projects in the history of U.S. military selection and classification, Project A evaluated new soldiers from accessions through their second tour of duty (35). Researchers examined concurrent and longitudinal data in an effort to identify the relationships between predictors and performance (35). In a large review of the ASVAB validity studies, Welsh *et al.* (106) affirmed that the ASVAB and AFQT predicted final school grades, self-paced technical school completion times, first-term attrition, and job performance with correlations ranging from 0.55 to 0.60. As Project A was beginning, Project B was initiated. Project B involved the development of a computerized job assignment program to match the most qualified applicants with duty assignments that would become available within the year. Prior to the development of this computerized assignment system, the Army filled job positions based on immediate vacancy and whether the soldier met the minimum qualifications (35).

Although Projects A and B mark major advances in the Army's selection and classification program, their findings are based on the examination of entry-level Army general-purpose force positions. In contrast, candidates in the Army specialty-training program examined by the present study have been assessed and selected to train with the top 3% of the military – they are not entry-level personnel. The elite nature of the

personnel selected for these programs makes them stand out from the general-purpose force. Therefore, findings from Projects A and B have limited utility in A&S models for SOF, special duty, and high-risk assignments. **The current project built on the foundational research in Projects A and B and sought to identify the variables most predictive of graduation and attrition from the MOS phase of an Army specialty organization that assesses, selects, and trains the top-performing enlisted personnel.**

Military Selection, Assessment, and Classification in the 21st century

This section reviews major advancements and current practices in military assessment, selection, and classification in the 21st Century to identify which approaches are relevant to MOS classification in an Army specialty-training program. The first section begins with a brief overview of the current assessment and classification process starting with an applicant's first contact with a recruiter and ending with completion of Military Entrance Processing Station (MEPS). The second section provides a summary of non-cognitive assessment measures and their classification utility. The third section reviews two major Army research initiatives designed to improve noncommissioned officer (NCO) management, particularly in the areas of attrition and "talent management."

Army NCO leadership at the highest levels are calling for improvements to the classification process (60; 103). The current project responds to those calls by specifically examining the classification procedures in an elite Army specialty-training program. The assessment methods (*e.g.*, non-cognitive measures) and predictor variables (*e.g.*, personality characteristics) reviewed next are primarily implemented at MEPS for use in the A&S of conventional (*i.e.*, general purpose) Army forces. However, emerging

research highlights the potential use of these techniques with elite military specialty organizations. After reviewing current practices, the section concludes by focusing on the variables most likely to add to existing SOF A&S batteries and outlines how the current project addressed remaining research gaps.

Selection and Classification from Recruiter to Military Entrance Processing Station (MEPS)

The military's A&S process begins as soon as an applicant is in contact with a service recruiter (*e.g.*, Army, Navy, Air Force, Marine Corps). Service recruiters conduct formal and informal screenings, which may include the administration of a brief aptitude test, review of education status, and an assessment of general fitness (13; 87). If an applicant is a promising fit, then he/she is scheduled for a full formal assessment at MEPS (13; 87). At MEPS, prospective recruits undergo a series of tests including the ASVAB (if not completed previously) and a physical evaluation. Based on the outcomes of testing and the needs of the service branch, a list of possible career fields is generated. Applicants then meet with a service-specific guidance counselor to discuss the range of jobs for which they have qualified (13; 87). If the applicant and counselor agree on an assignment, then the applicant is sworn into the service (*e.g.*, Army) (13; 87). If no agreement can be made, then the applicant is free to leave MEPS without a service obligation (13; 87).

The AFQT and ASVAB remain the Army's primary tools to assess a recruit's aptitude for training. The AFQT score determines whether an applicant has the minimum scores required to enlist and is comprised of subtests assessing arithmetic reasoning, mathematical knowledge, paragraph completion, and word knowledge. The ASVAB is comprised of 10 tests (Table 1) which can be combined to calculate 10 aptitude

composite scores, commonly referred to as “line scores” (Table 2). Line scores provide information about general cognitive aptitude and help inform classification decisions (87). To qualify for a particular MOS, an applicant must meet the minimum scores required for assignment in that field. For example, a candidate interested in assignment as an Army intelligence analyst must obtain a minimum skilled technician (ST) score of 105 (Table 2). Candidates interested in applying for ARSOF must obtain a minimum of 110 on the GT score and 110 on either the combat (CO) or ST scores (depending on the specific SOF MOS). As stated above, some job specialties also require physical fitness tests as part of the assignment process.

The ASVAB is a measure of an applicant’s aptitude across a variety of domains and is used as a proxy for intelligence. In a review of 172 civilian and military studies that considered psychometric properties, the ASVAB emerged as predictive of success in military technical schools, first-term attrition, and performance (106). However, it is important to note that the Army allows applicants to take the ASVAB more than once and offers classes designed to improve an applicant’s score. As a result, an applicant’s ASVAB score (and subsequent line scores) can improve with practice. Depending on the amount of preparation, ASVAB scores may be influenced by practice effects.

Therefore, the current project used candidates’ entrance GT score (*i.e.*, the first ASVAB attempt) and other measures of intelligence as cognitive predictor variables.

The Introduction of Non-cognitive Measures

Current research is examining ways to improve A&S programs by exploring the individual and collective contributions of non-cognitive (*e.g.*, personality) tests (57; 71; 86; 87; 108). Non-cognitive measures examine personality factors and vocational

interests (86). These measures provide additional information about the psychological functioning of recruits and reduce the emphasis on educational achievement (*e.g.*, attainment of a high school diploma) (87; 108). The inclusion of personality measures such as the Tailored Adapted Personality Assessment System (TAPAS) and interest inventories may provide incremental validity to the Army's classification and assignment process.

The Army first implemented TAPAS in 2009 to assess the attitudes and behaviors of civilians who were interested in serving in the Army (104). Initially, the measure was administered with the ASVAB as a screening/assessment tool. Recent research has examined the effectiveness of TAPAS to identify soldiers for special duty assignments (104). Previous attempts to use personality measures to assess for character traits associated with good job performance (*e.g.*, the Assessment of Background and Life Experiences [ABLE] and Assessment of Individual Motivation [AIM]) were unsuccessful because of the measures' high susceptibility to faking (85; 95). The TAPAS was specifically designed to minimize faking and positive image management among applicants (95; 103).

The TAPAS is a computerized multidimensional pairwise preference that examines 15 domains: dominance, attention-seeking, achievement, responsibility, non-delinquency, adjustment, even-temperedness, optimism, intellectual efficiency, tolerance, physical conditioning, adventure seeking, courage, situational awareness, and team orientation (71; 72; 104; 108). The majority of TAPAS domains were derived from the Big Five Model, which examines five dimensions/facets of a person's personality including neuroticism, openness, conscientiousness, extroversion, and agreeableness (72).

The five domains not associated with the Big Five (*e.g.*, physical conditioning, adventure seeking, courage, situational awareness, and team orientation) were added to assess important military-specific traits (71; 72). Research on the efficacy of TAPAS for use with MOS classification is promising and is discussed below (71; 72). It is important to note that logistical requirements preclude the use of the TAPAS in a group administration format, so the Army specialty-training program examined by this study does not use the TAPAS. Instead, the training program uses the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF) to assess personality facets (see Methods section for a full description). As such, further discussion of the TAPAS is included below to emphasize the value of including personality measures in A&S batteries.

Maximizing Noncommissioned Officer (NCO) Performance for the 21st Century and Army Class Effort

High attrition rates (25 to 30%) during first-term enlistments (108) and mandates from leaders in the general purpose force to develop an MOS competency evaluation system have led to a new wave of research (23; 87). Feedback from senior Army leaders (60; 103), especially within the NCO corps, stress a need to focus on “talent management,” which begins with an accurate identification of the skills associated with each MOS. Talent management involves fostering individual soldier characteristics (*e.g.*, skills, knowledge, attributes, and behaviors) to maximize performance and potential (37). In an address delivered at the NCO Solarium at Fort Leavenworth, Sergeant Major of the Army (SMA) Dailey stated that the biggest problems facing the Army are readiness, leadership, and talent management (103). His comments reflect concerns by senior Army NCOs about classification accuracy. As stated by one group of NCOs at the conference, “the service needs to do a better job of using the tools at its disposal to put the right

soldier in the right job” (p. 8) (60). In response, SMA Dailey indicated the Army was engaging in efforts to improve the current classification system: “The Army is in the process of evaluating all of the skills needed in each military occupational specialty, or MOS, and will be matching that to the knowledge, skills and attributes of Soldiers...” (p. 3; 103). These statements underscore the importance of developing selection and classification batteries that are valid, reliable, and highly predictive of success.

Since 2000, the Army has initiated two major research projects designed to improve selection, classification, and evaluation of enlisted personnel (57). The first project, which began in 2000, was organized into three phases: [a] Maximizing Noncommissioned Officer Performance for the 21st Century, NCO21; [b] New Predictors for Selecting and Assigning Future Force Soldiers, Select21; and [c] Performance Measures for 21st Century Soldier Assessment, PerformM21. NCO21 was the first phase of the program, the goal of which was to conduct a needs assessment to identify and validate non-cognitive predictors (*i.e.*, personality traits) of enlisted personnel performance (23; 55; 57). The second phase, Select21, sought to improve selection and classification of new enlisted personnel (57). During this phase, researchers evaluated prototype competency assessments for core and occupation-specific skills (56). The third phase, PerformM21, examined the utility of long-term competency assessments for enlisted personnel (57). Key findings from these research projects are discussed below.

In 2006, the Army Research Institute (ARI) initiated the second major research project entitled *Army Class, Validating Future Force Performance Measures*. Army Class built on the first set of research projects (*i.e.*, NCO21, Select21, and PerformM21)

and was designed to maximize the efficiency and accuracy of the Army's enlisted personnel selection and classification system (57). Army Class was a six-year effort with target goals for each year of study. Across the six years, the project aimed to improve MOS selection and classification, improve MOS reclassification (*i.e.*, the process by which an NCO in one MOS transitions to a different MOS), and conduct a longitudinal validation of selection measures based on the collection of job performance data (57). These research projects were designed to assist the conventional Army with attrition problems and issues related to talent management.

These research initiatives did not specifically involve elite organizations, so the findings may not generalize to SOF and other specialty communities. As previously mentioned, SOF constitutes less than 3% of the military (76) and operators are selected for their superior intellectual and physical functioning. Assessment batteries used by SOF, special duty, and high-risk A&S programs differ from those used by MEPS, so findings from the aforementioned studies may not be directly generalizable to these communities. However, assessment measures examining domains similar to those evaluated by MEPS (*e.g.*, intelligence, aptitude, personality) may prove useful to predict performance in elite organizations.

Attrition

Enlisted soldiers first term of service refers to their initial service obligation, usually four years (58; 108), and is a reflection of their ability to adapt to a military lifestyle (108). Historical attrition research indicates that 83% of the personnel who separate from the military in their first six months of service leave for the following reasons: medical disqualification, failure to meet performance standards,

fraudulent/erroneous entry, and behavioral problems (1). In 1997, it was estimated that a 10% reduction in attrition would save the services approximately \$12 million per year (1). In a 2004 literature review examining first-term attrition, the U.S. Army Center for Health Promotion and Preventive Medicine identified several risk factors which can be generally grouped into the following categories: lower education and intellectual functioning, low social conformity/morals (*e.g.*, following laws, rules, and regulations), mental health history (*e.g.*, depression, trauma, medication), and medical history (*e.g.*, hearing, skin, back, and knee disorders) (54).

In 2014, White *et al.* (108) examined the validity of cognitive (*i.e.*, ASVAB) and non-cognitive (*i.e.*, TAPAS) measures to predict Army attrition among a sample of 480,000 Army applicants. All applicants were tested at MEPS from October 2009 to December 2013. The sample was 70% Caucasian males and 96.8% were categorized as “Tier 1,” meaning they either had high school diplomas, college degrees, or educational credentials (*e.g.*, adult education diploma). Applicants higher in physical conditioning and with higher scores on the even-tempered and achievement facets of the TAPAS were less likely to attrite after 18 months of service. Selflessness and social desirability emerged as modest contributors to the statistical prediction model. Adding the ASVAB, particularly the mathematical knowledge and assembling object subtests, increased the predictive validity of the model. The results from this study indicated that the combined use of cognitive and non-cognitive tests could increase the predictive validity of an attrition model. The authors postulated that their findings may have been stronger had they examined attrition by MOS.

A variety of factors influence attrition rates. Failing to assign soldiers to an MOS that is a “good-fit” is one of those factors. **The present study examined cognitive and personality factors as potential predictors of MOS attrition and graduation.**

However, the Army specialty-training course under examination does not use the TAPAS as a measure of personality. Instead, the organization uses the MMPI-2-RF to assess personality attributes. Although the MMPI has a long history of use in the military (22), no research studies were found about the predictive validity of the MMPI-2-RF in MOS classification decisions. To address this research gap, the present project examined whether the MMPI-2-RF predicted attrition and graduation from a specialty MOS training course. The next section reviews similar research efforts by the conventional Army designed to improve talent management.

Talent Management/Competency

As described previously, ARI initiated a project in 2006 designed to develop competency assessments for enlisted personnel (23). The program was divided into three phases with the second phase (Select21) focused on developing MOS competency assessments (56). The project examined five MOSs: wheeled vehicle mechanic, health care specialist, Abrams armor crewman, Patriot air defense control operator/maintainer, and military police. The outcome of the project was the development of competency assessments examining soldier proficiency after entry into the general force. Scores from these assessments provided peer-to-peer comparisons and examined each soldier’s promotion potential (23). Although the information generated by these competency evaluations provided important information about MOS specific competency, the sample

did not include soldiers assigned to high-risk special operations billets, which limited generalizability to the community under study.

Nye *et al.* (72) examined the effectiveness of TAPAS to classify Army applicants into MOSs. The primary goal of this research was to examine whether the TAPAS could improve MOS classification and predict MOS performance. The project analyzed the predictive validity of the following variables: job knowledge tests, performance evaluations, attitude measures, and attrition data for soldiers in high-density (*i.e.*, largest) fields. The sample included 151,625 soldiers who processed through MEPS between May 2009 and June 2011 into the following MOSs: infantry, combat medics, military police, and motor transport operators. The authors categorized the outcome variables into three categories: can-do performance (*e.g.*, scores from Army and MOS job knowledge tests), will-do performance (*e.g.*, Army and MOS performance ratings, Army physical fitness scores [APFT], disciplinary actions), and attrition (*i.e.*, 6-month attrition). TAPAS composites were significantly related to job knowledge test scores, APFT, disciplinary problems, and attrition. The authors noted that TAPAS composites differentially predicted MOS success. In fact, 40 to 50% of the sample would have performed better in a different MOS, highlighting the value in using TAPAS as a measure to inform classification decisions. As previously mentioned, the TAPAS is not uniformly used in A&S paradigms. Other personality measures, such as the MMPI-2-RF, Revised NEO Personality Inventory (NEO PI-R), and Sixteen Personality Factor Questionnaire (16PF), are frequently used in high-risk A&S programs (28; 106). To address this research gap, the present project examined whether the MMPI-2-RF predicted attrition and graduation from the MOS phase of an Army specialty program.

Of the MOSs examined by Nye *et al.* (72), infantry and combat medic were the only two MOSs that were similar to the MOSs examined in the current research project. Infantrymen were found to have higher mean scores on physical conditioning and adjustment, but lower mean scores on tolerance, selflessness, and order, relative to the overall sample. TAPAS also was able to predict high scorers on the APFT and job knowledge test, but did not provide incremental validity above the ASVAB when the combat aptitude composite was included in the model. With regard to the combat medics, the TAPAS found higher mean scores on intellectual efficiency, even-temperedness, and attention seeking, and lower scores on order. Similar to findings for infantrymen, the TAPAS was a better predictor of will-do performance than can-do performance among combat medics, and the physical conditioning score emerged as a predictor of will-do performance and attrition. Overall, this line of research supports the use of personality measures as a tool for selection, assessment, and classification of MOS courses that train infantry and medical skills.

Although the aforementioned research focused primarily on the conventional Army, efforts are being made by researchers to examine the predictive validity of the TAPAS and other non-cognitive tests in A&S programs for SOF and specialty-training organizations (43; 71). **The current project was designed to contribute to classification research by examining MOS graduation and attrition in a high-risk specialty-training command that uses the MMPI-2-RF to measure personality factors.**

Assessing High-Risk Personnel

Given the highly specialized, technical, and psychologically-demanding requirements of their profession, special operators must be able to function physically and psychologically under substantial stress (78; 79). Because of the operational sensitivity and secrecy that frequently characterize SOF missions, there is often misunderstanding and confusion related to the operations carried out by these organizations. Picano and Roland (78) provide a general description of SOF operational capabilities:

High-risk military personnel typically engage in critical and sensitive national security missions; employ non-routine, nonstandard, or unconventional military tactics; deploy frequently and often for prolonged durations to denied or hostile environments in various cultural settings; operate fairly independently without much logistical or tactical support; and often encounter unknown and uncontrollable situational factors demanding ingenuity, expertise, initiative and a high degree of common sense to avoid mission failure (p.148).

Because of the versatility of missions and operational environments, there is no single “profile” for the perfect SOF operator (82).

Cognitive and Non-cognitive Factors

SOF A&S programs currently use limited psychological and physical assessment data to predict the performance of candidates under real-world stressors in harsh and unpredictable environments (82; 102). To ensure the right individuals are selected for these key roles, all special operators undergo a series of evaluations to determine whether they possess the physical and psychological traits required for success in this arena (82).

Although the A&S process differs among organizations, it generally involves assessments of personality, intelligence, and physical functioning.

A series of studies in the late 1980s and 1990s identified several characteristics predictive of selection from a SOF A&S program (16; 19; 25; 42; 66; 81; 82). These characteristics are summarized by Picano *et al.* (79) and include: psychological hardiness, exceptional stress tolerance, emotional stability, high levels of physical fitness, intrinsic motivation, initiative, competitive drive, exceptional reliability, integrity, and sound judgment and reasoning under stress. Gregariousness and interpersonal skills among operators vary by SOF organization; that is, some organizations are comprised of personnel with more of these traits and some are comprised of personnel with fewer of these traits (79). SOF A&S programs consider these characteristics and other variables during the selection process. Ultimately, SOF A&S programs are designed to rule out candidates who do not meet the minimum criteria, while also ruling in candidates who appear to possess the qualities most predictive of successful performance (25; 79; 82; 102).

As discussed in the previous section, research examining the predictive validity of non-cognitive assessment measures has identified the TAPAS as a useful instrument for selection, assessment, and classification of conventional force recruits. The potential utility of the TAPAS for use with SOF is also being examined. Using correlation and regression analyses, Nye *et al.* (71) examined the predictive validity of the TAPAS to identify which candidates would be selected for participation in an ARSOF training program. After examining 1,216 candidates who attended Special Forces A&S (SFAS) from February to June 2012, the authors examined TAPAS scales and found that 61% of

the candidates who scored in the highest quintile were selected for the Special Forces Qualification Course and only 35% of the individuals in the lowest quintile were selected to advance. Selected ARSOF candidates scored higher than a comparison MEPS sample on the following domains: physical conditioning, achievement, tolerance, and courage. ARSOF scored lower than MEPS personnel on attention-seeking and non-delinquency. Optimism, physical conditioning, and responsibility emerged as the strongest predictors of ARSOF selection; however, the authors note the magnitudes of the associations were relatively weak, possibly because of the small sample size and violations of statistical assumptions (71). Overall, the TAPAS appeared useful in predicting which candidates were selected for advancement in the Army's Special Warfare Center and School (SWCS). The authors observed that the TAPAS also may prove useful to improve MOS assignment in ARSOF and other specialty programs. However, the Army specialty-training program examined by the present study is not using the TAPAS; instead, they are using the MMPI-2-RF to assess candidate personality attributes. **The present study examined the predictive validity of the non-cognitive test (*i.e.*, MMPI-2-RF) currently in use.**

These research studies elucidate the cognitive and non-cognitive characteristics associated with success in SOF A&S programs. However, A&S is considered a preliminary stage in the SOF training pipeline. After selection, candidates must successfully pass several stages in a SOF qualification course (including MOS training) prior to earning entry into the organization. Although the characteristics highlighted above are informative, they are not necessarily predictive of MOS training completion. Rising rates of recycle, reclass, and relief from MOS training courses in an Army

specialty-training program have motivated leaders to assess the current MOS assignment process. **The primary purpose of the present study was to determine whether the variables gathered during A&S predicted graduation and attrition from four different Army MOS courses.**

Present Project: Predicting Attrition in an Army Specialty-Training Program

The goal of psychological A&S is to identify personnel best suited for a particular job by matching applicant aptitude with job requirements (85; 87). At this time, no single tool can perfectly match applicants with jobs for which they are best suited (87).

Variable job demands, unpredictable environments, unique military stressors, and a host of individual and social characteristics make it difficult to create a standardized test battery that fits the needs of every organization. Further, the majority of research on A&S has focused on selecting an applicant/candidate for entry into a training pipeline, not for successful completion of an MOS course. Increasing attrition rates, problems with talent management, and a high reliance on SOF units in the current operational environment underscore the essential nature of accurate job classification for both the conventional Army and elite military organizations.

The current project was designed to contribute to the growing body of research on accurate MOS classification by: [a] assessing the predictive validity of a test battery used by an Army specialty-training program, and [b] attempting to develop an MOS classification algorithm (*i.e.*, decision tree) to improve classification assignment. This project extended previous research and went beyond the identification of predictor variables; its ultimate goal was to identify an MOS assignment decision tree that leaders could use to improve classification accuracy.

The combination of tests in the psychological battery used by an Army specialty-training program is different than the tests administered at MEPS, so findings from Select21, NCO21, and PerformM21 have limited utility. Further, the individual characteristics of candidates in this program are different from the standard Army population. Therefore, existing literature does not fully satisfy the research question: which demographic, military, physical, and psychological variables are associated with the highest attrition and highest graduation rates in the MOS phase of an Army specialty-training program?

The operational tempo has drastically expanded the employment of SOF across the globe and SOCOM is expected to maintain a high operational tempo (3; 4; 63; 84). Identifying personnel who possess the psychological and physical robustness to fulfill the operational requirements our nation has placed on SOF, high-risk, and other specialty duty appointments is an important task. **The present project aimed to enhance MOS assignment in an Army specialty-training organization by identifying the psychological, physical, demographic, and military factors most predictive of graduation and attrition.**

CHAPTER 2: Specific Aims and Hypotheses

OVERVIEW AND SPECIFIC AIMS

Maximizing graduation and minimizing attrition from MOS courses in an Army specialty-training program is a high priority for military leaders and trainers. The purpose of the present project was: [a] to identify the variables most predictive of graduation and attrition (*i.e.*, recycle or relief) from four distinct Army MOS courses, and to develop a classification tree that assists personnel managers with MOS assignment. This study addressed two specific aims.

Specific Aim 1

Identify variables that optimally predict graduation and attrition (*i.e.*, recycle and relief) from MOS courses in an Army specialty-training program.

Hypothesis 1a

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS graduation** better than chance.

Hypothesis 1b

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS recycle** better than chance.

Hypothesis 1c

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS relief** better than chance.

Specific Aim 1 Rationale

Personnel interested in assessing for this Army specialty program must demonstrate higher than average intellectual skills (*i.e.*, minimum GT score of 110) and meet certain fitness standards on Army fitness tests. Applicants who successfully complete the assessment phase and advance into the training pipeline have proven they are top performers; therefore, it is unclear why some candidates recycle or fail the MOS course and some do not. The primary goal of this specific aim was to identify possible aptitude differences between the groups and determine which variables were most predictive of both graduation and attrition from each MOS course.

Specific Aim 1 addressed whether a particular combination of the predictor variables predicted MOS course graduation and MOS attrition for each of the four MOS courses (*i.e.*, Weapons, Engineering, Medical, and Communications). The included predictor variables were selected for two primary reasons: [a] a large body of literature features the effectiveness of these variables in accurately predicting course graduation from a variety of civilian and military courses, and [b] these variables have been collected by the Army A&S program for many years (Table 3).

Specific Aim 2

Develop prediction rules, using the significant predictor variables identified in Specific Aim 1, to identify candidates most likely to graduate or attrite (*i.e.*, recycle or fail) from their assigned MOS course. Specific Aim 2 also evaluated the classification tree against a subsample of the population.

Hypothesis 2

The classification tree model will accurately classify MOS outcome (*i.e.*, graduation, recycle, or relief) better than chance.

Specific Aim 2 Rationale

Specific Aim 2 explored potential interactions between demographic, military, fitness, orienteering, cognitive, and personality variables in determining MOS graduation and attrition. Specific Aim 2 went beyond identifying key variables and sought to create an MOS classification tree that improved classification accuracy and provided MOS assignment personnel a straightforward methodology for classifying candidates into MOS courses.

CHAPTER 3: Methods

OVERVIEW

A retrospective analysis of MOS training outcome data from fiscal years 2014 through 2017 was used to identify the variables most predictive of graduation and attrition from an Army specialty-training school. The data were collected from 1,460 male, active duty, enlisted, Army service members during assessment and selection (A&S).

RESEARCH DESIGN AND PROCEDURES

Data Management

The research was conducted on previously de-identified data, collected by operational psychologists and instructors assigned to an Army specialty-training program. The following variables were selected from the database as potential predictors based on consideration of relevant studies described in Chapter 1 and discussions with operational psychologists who are experts in SOF A&S.

Demographic -- Demographics and Military: age, previous MOS, rank, years of service, and MOS preferences

Performance -- Fitness: physical fitness test, ruck/forced march, and obstacle course

Performance -- Orienteering: land navigation scores

Psychological -- Cognitive: Multidimensional Aptitude Battery II (MAB-II), entrance ASVAB GT score, highest level of education, and Test of Adult Basic Education (TABE)

Psychological -- Personality: Minnesota Multiphasic Personality Inventory-2-
Restructured Form (MMPI-2-RF)

The outcome variable was a candidate's final course disposition at the conclusion of MOS training. Final MOS course disposition was examined as a categorical variable with three levels:

- 1) MOS graduation
- 2) MOS recycle
- 3) MOS relief/other non-successful completion

Participants

The present project used de-identified data gathered from 1,460 enlisted, male, Army, active duty service members between 19 and 40 years old who began the MOS course at an east coast Army specialty-training program between fiscal years 2014 and 2017.

Procedures

Demographic, performance, and psychological data were collected during the A&S phase of the training program. Outcome data (*i.e.*, graduation, recycle, or relief) from participants that began the MOS phase of the program were linked to A&S data for the purpose of this examination. A research psychologist at the Army specialty unit de-identified the data prior to transfer to USUHS. Data transfer from the Army specialty unit occurred after USUHS and Army IRB approval (Appendix B, Figures 1 to 3).

Assessment Measures

1. *Demographics and Military Questionnaire*: The demographics questionnaire is a self-report measure that includes: age, race, branch of service, previous MOS (1 = Operations, 2 = Operations Support, 3 = Force Sustainment), rank/grade, active duty time in service, entrance and current GT scores, military training (*e.g.*, Airborne, Air Assault, Pathfinder), combat, gender, marital status, number of children, education level, years of school, highest level of school, security clearance, and native language.
2. *Army Physical Fitness Test (APFT)*: The APFT, outlined in Training Circular 3-22.20 (12), measures a candidate's upper and lower body muscular endurance and cardiorespiratory fitness. The test consisted of three timed events administered in this order: push-ups, sit-ups, and a 2-mile run (12). The three events were executed in the same order each time with no less than 10 minutes and no more than 20 minutes between each event; the total APFT time could not exceed 2 hours. Raw scores on each event (*e.g.*, number of push-ups and sit-ups in two minutes and 2-mile run time) were converted into scores ranging from 0-100 for each event. Converted scores were based on the candidate's age and gender. Event scores were summed to provide an overall APFT score ranging from 0-300 with higher scores indicating more advanced levels of fitness (12).
3. *Obstacle Course*: The obstacle course, described in Army Training Circular 3-22-20, assesses fundamental skills and abilities related to negotiating obstacles. Obstacle negotiation involves crawling, walking, climbing, creeping, running, and jumping while carrying field equipment. The obstacle course provided information about a candidate's physical conditioning, motor skills, and confidence (12). Candidates were scored on their ability to successfully negotiate obstacles and evaluated on the total time it took to

complete the course. Final scores ranged from 0-100 with higher scores representing better obstacle negotiation skills.

4. *Ruck/Forced March*: The Ruck March, a type of foot march (also known as forced march and road march), involves the movement of troops and equipment across varying terrains primarily by foot and with little external supports (10). As described in Field Manual No. 21-18, forced marches are characterized as maximum effort events that require individual soldiers to carry heavy loads and traverse long distances, at high rates of speed, and using high levels of exertion (10). The Ruck March measured each candidate's physical and muscular endurance. Lower scores indicated faster mean times.

5. *Land Navigation*: Land Navigation, outlined in Field Manual 3-25.26, measures a candidate's ability to navigate from one predetermined location on a map to another predetermined location using terrain association, a compass, and map. No other forms of assistance are provided. Land navigation is considered a useful tool to measure a candidate's motivation, perseverance, fitness, and orienteering skills (11; 21). During this event, a variety of factors (*e.g.*, lack of performance feedback, lack of support, night exercises, and fatigue) stressed candidates physically and psychologically. Candidates were timed on their ability to accurately navigate from one point to the next and scored on the total number of correct points located within the allotted time.

6. *Multidimensional Aptitude Battery II (MAB-II)*: The MAB-II, developed by Douglas N. Jackson in 1984, is a valid and reliable assessment of aptitude and intelligence (48). The MAB-II is group administered in a paper-and-pencil format and consists of 334 items arranged into 10 subtests, which each take 7 minutes to complete. Results from the MAB provided information about a respondent's full scale, verbal, and performance IQs

and were reported as standard scores. There were five subtests in each domain, which assessed distinct intellectual areas and were reported as scaled scores. The verbal subtests included: information, comprehension, arithmetic, similarities, and vocabulary. The performance subtests included: digit symbol, picture completion, spatial, picture arrangement, and object assembly. Test-retest reliabilities scores were found to be high: Verbal IQ (0.95), Performance IQ (0.96), and Full Scale IQ (0.97) (47; 48).

In a study comparing the MAB and the Wechsler Adult Intelligence Scores-Revised (WAIS-R), Carless (24) found high correlations between the tests on Verbal ($r = 0.77$), Performance ($r = 0.79$), and Full Scale ($r = 0.81$) scores. The MAB-II has successfully been used with high-risk Air Force personnel. Chappelle *et al.* (27) used the MAB-II to establish population-specific norms for U.S. Air Force AC-130 Sensor Operators (SO). An examination of intelligence and aptitude among SOs from 1st Special Operations Group showed SOs scored significantly higher than the general population on Verbal, Performance, and Full Scale IQs. Also, operators evidenced higher scores on domains assessing visual-performance/aptitude (27). The authors concluded that the MAB-II was helpful in establishing population-specific cognitive norms and in identifying personnel with the desired attributes during A&S.

7. *Test of Adult Basic Education (TABE)*: The TABE was developed by Data Recognition Corporation (DRC)|CTB based on feedback from educational experts about the current academic expectations for adult curricula (30). The TABE was normed using national standards and validated using a sample of 34,000 students (30). The TABE is a valid and reliable multiple choice, pencil-and-paper test that assesses a respondent's basic skills in the areas of reading (50 items), math (90 items), and language (55 items).

Reliability estimates for test version Form 9, level A (Advanced, grade level equivalents 9-12.9) range from 0.90 to 0.92 (67).

Historically, Army Special Forces (SF) has used the TABE as part of the A&S psychological test battery. From 1995 to 1998, 46% of A&S candidates earned scores equivalent to that of the average graduating high school senior and 60% earned scores equivalent to a high school senior (42). Math scores emerged as the most important predictor and lower grade equivalent scores were associated with higher rates of A&S drop out (42).

8. *Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF)*:

The MMPI-2-RF, developed by Yossef Ben-Porath and Auke Tellegen (2008), is a standardized, multi-dimensional assessment used to assist in identifying predominant aspects of personality structure and psychopathology (39). The MMPI-2-RF examined possible areas of concern by comparing patterns of responses provided by an individual to similar patterns of responses with those that have been given psychological diagnoses. In addition, the MMPI-2-RF examined an individual's test-taking attitudes, which can directly impact the validity of the personality inventory including an individual's efforts to exaggerate distress or promote an unusually positive image, either intentionally or unintentionally. The measure is a paper-and-pencil format consisting of 338 items organized into 50 scales (eight Validity Scales; three Higher-Order Scales; nine Restructured Clinical Scales; 23 Specific Problem Scales; two Interest Scales; and five Personality Psychopathology [PSY-5] Scales) (17; 39). Four to 10 items assess each specific problem scale (39).

The MMPI has a long history of use with military populations (22; 32; 42; 74; 98).

Butcher *et al.* (22) examined the appropriateness of using the MMPI-2 re-standardized norms (the predecessor to the MMPI-2-RF) with military personnel. After examining a sample of 1,156 male participants from four branches of the Armed services (Army, Navy, Air Force, and Marines), Butcher *et al.* (22) concluded that military specific norms were unnecessary because there were minimal score differences between the military population and the re-standardization population. MMPI-2-RF scores are presented as *T* scores.

Outcome Variables

MOS training courses ranged from 24 to 57 weeks depending on the specific course (*e.g.*, Medic or Weapons expert). Each MOS course was independently run and instructed; however, all MOS courses included classroom academic instruction and field training exercises.

1. *MOS graduate.* MOS graduate included all candidates who completed the course without recycling or repeating any block(s) of instruction (Table 4).
2. *MOS recycle.* MOS recycle referred to candidates who were required to repeat any portion of the MOS course in order to graduate.
3. *MOS relief/other non-successful completion.* MOS relief or other non-successful completion referred to candidates who began, but did not finish, the MOS course to which they were originally assigned. Non-successful completion included, but was not limited to, failure to meet the minimum course requirements, candidate voluntary withdrawal, and behavioral

problems that warranted removal from the entire training program (*e.g.*, arrest).

DATA ANALYTIC PLAN

Missing Data and Software Analysis

Data Management

Data were evaluated for missing values, outliers, and distribution using SPSS v24 and Microsoft Excel 2013. In instances where greater than 10% of data were missing for a particular candidate or a particular variable, the candidate or variable was dropped from the analyses because the sample size was sufficiently large. Outliers were examined and considered for exclusion (45). Distributions were examined to determine whether transformations were necessary (45).

Sample and Power

Participants were 1,460 enlisted, male, Army, active duty service members who began an MOS course at an east coast Army specialty-training program. The number of candidates needed for each MOS was estimated using an *a priori* power analysis using G*Power (31). Because this study used a multinomial logistic regression, two separate independent logistic regression power analyses were calculated for the probability of recycling versus graduating ($N = 87$), and the probability of failing versus graduating ($N = 116$); the larger sample size was selected. G*Power indicated that a total of 464 candidates (approximately 116 candidates per MOS), was sufficient to detect a medium effect (*i.e.*, 2.0 odds ratio) with 80% power ($\beta = .20$).

The data were randomly divided by SPSS into two groups using a 60:40 split (Table 5). Data from 60% of the candidates (862 total) served as the primary sample for

the multinomial logistic regression and classification tree modeling. Data from the remaining 40% of the candidates (roughly 83 to 193 depending on the MOS) were used to validate the multinomial logistic regression and the classification tree models. The 60:40 split was selected because it was considered a reasonable compromise between the 55:45 data-partitioning recommendation provided by SAS Enterprise Miner (91) and the 70:30 recommendation found in regression and model-validation literature (41; 105).

Specific Aim 1

The purpose of Specific Aim 1 was to identify variables that optimally predicted graduation and attrition from MOS courses in an Army specialty-training program. To address this specific aim, multinomial logistic regression analyses were conducted on categorical (*e.g.*, previous MOS, grade, and MOS preferences) and continuous (*e.g.*, MAB-II, MMPI-2-RF, and APFT) predictor variables to determine their relationship with the categorical outcome variable (*i.e.*, final MOS disposition: graduation, recycle, or relief). The regression model was built using the seven-step model building strategy proposed by Hosmer *et al.* (44):

[a] First, important covariates were identified using a Pearson's Chi-square for categorical variables and a univariate logistic regression model for continuous covariates. Cells that totaled zero were examined and Kruskal-Wallis statistics were reported instead of Pearson's Chi-square. Variables with a p value of 0.25 were originally considered for inclusion in the multivariate analyses; however, in some MOSs a few number of variables met this criteria. To err on the side of inclusiveness, $p \leq 0.28$ was used as the cutoff instead. Examination of these variables (*i.e.*, variables with p value ≤ 0.28) using backwards stepwise regression

resulted in statistical problems (*i.e.*, models that could not converge); therefore, a forward step-wise procedure with all predictor variables was used instead (44).

[b] Steps two through four of Hosmer's model building strategy were automatically conducted by SPSS as part of the forward stepwise procedure.

Entry and removal criteria were set as 0.10 and 0.15, respectively (44).

[c] A *main effects model* was developed after examining the linearity of continuous variables using Box-Tidwell tests (44).

[d] Possible interaction terms were conceptually derived (see Results) and added to the model. At the conclusion of this step, a *preliminary final model* was created (44).

[e] Pearson's Chi-square analyses and external validation using a randomly selected subsample (40% of the candidates) were used to assess model adequacy and fit (44). Specifically, the regression parameters from the proposed model were used to construct predicted values for the 40% validation sample. Predicted and observed outcomes were compared in this sample to assess overall model fit and predictive accuracy for each regression model was examined.

Specific Aim 2

Classification and regression trees (CART) have gained popularity in the field of psychology (52; 53; 68; 95; 97). These statistical methods provide clinicians with a practical tool that can improve decision-making. As outlined by Merkle and Shaffer (68), the primary benefits of CART include: [a] it is non-parametric and does not utilize significance tests, [b] it can handle all types of variables (continuous, categorical, and ordinal), [c] missing data do not require imputation, [d] it is invariant to monotonic

transformation of the predictor variables, and [e] it is robust to outliers. In this study, CART analyses were designed to complement traditional regression analyses and were included to provide the end-user (Army specialty unit) with a decision-making tool for improving MOS assignment accuracy.

The goal of Specific Aim 2 was to develop prediction rules (*i.e.*, a classification tree), using the significant predictor variables identified in Specific Aim 1, to identify candidates most likely to graduate or attrite (*i.e.*, recycle or fail) from their assigned MOS course. However, because a limited number of variables emerged as significant in the regression analyses, all 18 variables were entered into the CART model as predictors for each MOS. Specific Aim 2 also evaluated the classification tree against a subsample of the population (*i.e.*, 40% subsample).

Salford Predictive Modeler (SPM) version 8.0 was used to build CART models for each MOS. Target (*i.e.*, outcome variable) and predictor variables were identified and class names were assigned to categorical outcome variables. The 60% subsample was designated as the “learn” group and the 40% subsample was designated as the “test” sample. Priors refers to the prior frequency distribution; these parameters were set to “learn,” which means that probabilities were matched to the learn sample frequencies (Table 19; approximately 60% [graduates], 30% [recycles], and 10% [reliefs]). The remaining model parameters were set to the CART default entries.

Development of the classification tree began with an examination of predictor variables to identify the optimal binary split, or cutpoint, for each predictor variable (51; 68). For continuous variables, a threshold for the cutpoint was identified (*e.g.*, > 249.50 or ≤ 249.50 on the APFT). For categorical predictors, cutpoints were between all binary

groupings. For example, the MOS preference variable is comprised of the following MOS classifications: B (Weapons), C (Engineering), D (Medical), and E (Communication). Cutpoints include B versus C/D/E; C versus B/D/E; D versus B/C/E; and E versus B/C/D. The “Gini” improvement measure was used to examine the cutpoints of all predictor variables and to identify the split that created the greatest level of homogeneity (59; 77; 97). The purity of each subset was examined and an *entropy* score (*e.g.*, Gini index) was calculated. Subsets with entropy scores of 0 were considered pure subsets (*e.g.*, all APFT scores in the subset are greater than 249.50) (97). An entropy score of 1 indicated both outcomes were equally represented. Examining entropy before and after the cutpoint provided an *information gain* value. In other words, how much more information was gained about the outcome variable by including the cutpoint. These steps were repeated for each cutpoint until no additional cutpoints improved the prediction (59).

To address problems of *overfitting*, or creating a tree that was specific to the current dataset and therefore could not be generalized, post-pruning and validation procedures were implemented. Post-pruning procedures involved removing each node (*i.e.*, branch) and then measuring the performance of the pruned tree on the validation data. These post-pruning procedures were repeated until removal of a node became harmful (51). The final model referred to as the “optimal tree” was the tree that performed best on the test data (*i.e.*, 40% subsample).

Validation procedures involved applying the decision rule generated by CART (based on 60% of candidates) to a separate validation sample (remaining 40% of candidates) and calculating measures of predictive accuracy.

CHAPTER 4: Results

This section provides an overview of data analytic procedures from database development to completion of the statistical analyses.

DATA MANAGEMENT RESULTS

Data provided by the Army specialty-training organization were organized into three separate databases. The first database was presented in SPSS and included demographic data and psychological testing data for 1,722 candidates and 269 variables. This database, referred to as COG (cognitive) database, was exported into Excel for data cleaning purposes. The second database was presented in Excel and included performance and educational data (*e.g.*, TABE [Test of Adult Basic Education] scores and level of education) for 1,839 candidates and 33 variables; this database was named PERF (performance) database. The third database was presented in Excel and contained information about candidate military occupational specialty (MOS) assignment and course disposition. The third database was named OUTCOME database and included information for 2,415 candidates and eight variables. Predictor variables were extracted from each database and merged into a single database titled FINAL. The FINAL database contained a total of 46 variables and 1,460 candidates (see Methods for detailed discussion of the predictor variables).

Data Cleaning

Two thousand and thirty-six cases were retained for database consolidation. A total of 379 cases were eliminated because they did not have corresponding performance or psychological testing data. For example, although candidate one is listed in the

OUTCOME data spreadsheet, he did not have corresponding data in COG or PERF spreadsheets. Because candidate one was missing data for the predictor variables, his information was not merged into the FINAL database. After the creation of a single Excel database, preliminary analyses of each predictor variable were conducted to identify missing data. Preliminary analyses identified one candidate, number 288, as missing the majority of performance data including: Army Physical Fitness Test (APFT), ruck march, and land navigation scores. Candidate 288 was subsequently removed from the analyses prior to the creation of the FINAL database.

After merging the databases into a FINAL database, preliminary analyses using Excel were conducted to identify missing or out-of-range data (*i.e.*, a value was outside the range of possible scores). In many cases redundant data across the three spreadsheets allowed for the substitution of data that were identified as missing or out of range. Of note, data substitution was a process whereby missing or out-of-range data were recalculated using information provided in the original databases. For example, date of birth was used to calculate missing candidate ages. In this case, data substitution does not refer to data imputation, which is a separate and distinct process. Data substitution strategies were implemented for the following variables and are discussed next: entrance GT (General Technical) scores, age, years of school, Multidimensional Aptitude Battery (MAB) data, MOS training assignment (*i.e.*, Weapons, Engineer, Medical, and Communications experts), outcome data, and APFT.

Entrance General Technical (GT) Scores

Separate GT scores were provided across the three independent databases. This study was interested in entrance GT scores so data from the entrance GT column in the

COG database was the primary reference source. A total of 160 candidates had out-of-range (*e.g.*, scores ≤ 11) or missing entrance GT scores. A minimum GT score of 110 is required for consideration into the Army specialty-training unit under examination.

Therefore, it is impossible that GT scores equaling 11 or less are accurate. To address this discrepancy, entrance GT scores for these 160 candidates were replaced using a two-step process. First, missing data were replaced with GT data from the PERF database. Second, in the event that GT data from the PERF database were unavailable, current GT data were used from the COG database. After consolidating GT scores across the three databases, only four candidates were missing GT data.

Age

Two candidates had ages out of range (*e.g.*, 2 years old). Their ages were subsequently calculated by subtracting date of birth from the date of testing, which provided their ages at the time of assessment and selection (A&S).

Years of Education (YRS ED)

Both years of education (YRS ED) and years of school were provided in separate databases. Years of education is considered equivalent to years of school. Missing or out-of-range data (*e.g.*, ≤ 8 years of school) were addressed by consolidating these data into a single variable, YRS ED.

Multidimensional Aptitude Battery (MAB)

MAB data for verbal, performance, and full scale intelligence scales were provided in two separate databases. Two candidates had missing data from at least one of the databases. Missing data were addressed via database consolidation.

Training MOS and Training Outcome

Six candidates were missing MOS training assignment data (*i.e.*, assignment as Weapons, Engineer, Medical, or Communications experts). However, the outcome data spreadsheet provided information about the training MOS course number to which these candidates were assigned. The MOS training courses are directly associated with MOS assignment, thereby providing information for the missing data. Candidates who recycled or reclassified into another training MOS had separate lines of data in the OUTCOME database. Candidates with multiple lines of data were consolidated into a single line and coded as follows: candidates who completed their MOS training course without recycling or reclassifying were given an outcome disposition of graduate; candidates who recycled (one or more times) and eventually graduated their assigned MOS course were classified as recycle; candidates who failed to complete their assigned MOS course (*i.e.*, on first attempt or after recycling) were classified as relief.

Army Physical Fitness Test (APFT)

Eighty-three candidates had APFT run scores outside the possible run range. Because the APFT total score is a composite score comprised of push-ups, sit-ups, and run, the overall APFT score was retained for inclusion in the analyses. All APFT total scores fell within the expected range.

Quality Assurance

After data cleaning and consolidation of the three databases, the FINAL database was validated in two ways. First, Excel formulas were used to identify any discrepancies in data by comparing columns of data from the FINAL spreadsheet to data from the individual spreadsheets. For example, training MOS categories (*i.e.*, Weapons, Engineer,

Medical, and Communications experts) in the consolidated database were compared to the training MOS categories provided in the original OUTCOME database. All discrepancies identified by Excel were the result of input errors and were corrected following careful examination of the original databases. Second, an independent reviewer checked the final database against the three separate databases using the following steps: [a] a reviewer examined data for every 10 cases. If discrepancies were found, then the reviewer manually corrected the data entries; [b] if there were multiple errors in the preceding 10 cases, then the reviewer examined candidate data for every five cases; [c] if multiple errors were identified, then the reviewer examined each case individually.

Missing Data, Skewness, Kurtosis, and Multicollinearity

After creation and validation of the FINAL consolidated database, each variable and case were independently analyzed using IBM SPSS Statistics v24 to determine the percentage of missing data. Two variables, obstacle course (OCourse) and MOS preferences (MOSPref), each had missing data rates averaging 13% (Tables 6 to 9). The OCourse variable is a measurement of physical fitness, which is a characteristic also measured by APFT, Ruck March, and Land Navigation; therefore, the variable was dropped (Table 16). Examination of this variable using univariate analyses during the model building stage (see Specific Aim 1) indicated that OCourse scores were not significantly associated with the outcome variable.

MOSPref captures whether the candidate received their first, second, third, or fourth choice MOS assignment. Approximately 66% of candidates were assigned to their first choice and approximately 16% were assigned to their second choice; approximately

8% or fewer received third or fourth choices. Examination of this variable using univariate analyses during the model building stage (see Specific Aim 1) indicated MOSPref was not significantly associated with the outcome variable. Given that approximately 82% of candidates received their first or second choice, there was greater than 10% missing data, and the non-significant findings from the univariate analyses, the variable was dropped. Consultation with a statistical expert also resulted in the recommendation to drop both of these variables from the analyses. Only eight candidates were missing one data point, which is less than 10%, so all candidates were retained for analyses.

ANOVA assumptions were sometimes violated and several variables had significant values for skewness, kurtosis, and/or normality (Tables 10 to 12). Visual examination of the data (*i.e.*, histograms) revealed that the data were appropriately distributed for this population. According to Field (34), large datasets are likely to produce significant skewness, kurtosis, and normality findings even though the data are appropriately distributed.

Linear regression analyses were conducted to examine multicollinearity among variables. All variables had variable inflation factors (VIF) less than 4.0. Although VIQ and PIQ had large variance proportions, bivariate analyses found only moderate correlations between the two variables (*i.e.*, $r = 0.39$). Based on recommendations by Field (34), and consultation with statistical experts, these variables were retained and statistical assumptions for ANOVA and multinomial logistic regression were determined adequately met.

Variable Selection

Intelligence and Education Variables

The FINAL database began with 46 predictor variables. However, several variables measured similar constructs. For example, the MAB is comprised of 10 subtests (*i.e.*, five verbal and five performance) that collectively provide information about each candidate's Full Scale IQ (FSIQ), Verbal IQ (VIQ), and Performance IQ (PIQ). Entrance GT scores also provided information on intellectual functioning; in particular, each candidate's general cognitive aptitude. Given that the MAB and GT scores were characteristically similar, it was necessary to examine the relationship between these variables. Similarly, the Test of Adult Basic Education (TABE) provided information about candidates' math, reading, and language skills and was expressed by educational level. For example, the highest scoring candidates earned scores consistent with the 12.9 grade reading, math, and language levels. These scores may prove similar to candidates' level of education. To minimize the inclusion of redundant, or highly correlated variables, correlational analyses were conducted to examine the degree of association between variables (Tables 13 to 17).

Correlational analyses were conducted to examine the relationships between the five verbal subtests, five performance subtests, and the VIQ, PIQ, and FSIQ (Table 14). A Pearson's product-moment correlation coefficient ≥ 0.5 accounts for 25% of the variance and was considered a large effect size (34). Overall, the VIQ and PIQ scores were correlated highly with FSIQ ($r > 0.80$), but only moderately correlated with each other ($r = 0.39$). Additionally, FSIQ was highly correlated with Entrance GT ($r > 0.63$, Table 13), another measure of intelligence. Entrance GT also correlated highly with VIQ ($r > 0.71$, Table 13). Therefore, FSIQ and Entrance GT were removed as predictor

variables and VIQ and PIQ were retained. The five verbal subtests correlated highly with VIQ ($r \geq 0.60$) and minimally with PIQ ($0.24 \leq r \leq 0.37$). The five performance subtests correlated highly with PIQ ($r \geq 0.60$) but less so with VIQ ($0.21 \leq r \leq 0.30$). Given the high correlations between the subtests and their corresponding index scores, the VIQ and PIQ index scores were retained for analyses and the 10 subtest and FSIQ scores were eliminated as predictor variables.

Correlational analyses of TABE scores and years of education (YRS ED; Table 15) showed small effect sizes ($0.05 \leq r \leq 0.18$). These small effects sizes were unusual given that years of education is conceptually related to grade-level academic performance. Closer examination of the data indicated these findings were likely the result of a TABE ceiling effect. TABE scores provided information about grade level performance in reading, math, and language; candidate grades ranged from 1.9 to 12.9. YRS ED was self-reported data that referred to the total years of school each candidate completed. YRS ED ranged from 10 to 20 years. Given the restricted range of the TABE, it is likely that candidates were performing at grade levels above 12.9, which the TABE is unable to measure. Consequently, the relationship between the TABE and YRS ED produced a small effect size. Based on the possible ceiling effect created by the TABE, YRS ED was the variable selected for retention in the analyses. The three TABE subtests were excluded as predictor variables.

Personality Variables

The associations between MMPI-2-RF Higher Order (3 scales), Restructured Clinical (9 scales), and PSY-5 (5 scales) scales were examined using correlational analyses (Table 17). Pearson's product-moment correlation coefficients > 0.5 were

considered large. Eight scales had Pearson's correlation coefficient scores < 0.5 and were retained in the analyses: Emotional/Internalizing Dysfunction (EID), Thought Dysfunction (THD), Behavioral/Externalizing Dysfunction (BXD), Somatic Complaints (RC1), Cynicism (RC3), Ideas of Persecution (RC6), Aggressiveness Revised (AGGR-r), and Introversion/Low Positive Emotionality-Revised (INTR-r). Nine scales exceeded the cutoff score ($r > 0.5$) and were eliminated: Demoralization (RCd), Low Positive Emotions (RC2), Antisocial Behavior (RC4), Dysfunctional Negative Emotions (RC7), Aberrant Experiences (RC8), Hypomanic Activation (RC9), Psychoticism-Revised (PSYC-r), Disconstraint-Revised (DISC-r), and Negative Emotionality/Neuroticism-Revised (NEGE-r). Overall, eight MMPI-2-RF scales were retained for the final analyses.

The FINAL database began with 46 variables. Correlational analyses allowed for the removal of 24 cognitive/intelligence and psychological variables. A single APFT total score was included in lieu of the three separate event scores (*i.e.*, push-ups, sit-ups and run) and missing data analyses resulted in the removal of two variables (*i.e.*, OCourse and MOSPref). Ultimately, 18 variables were utilized to develop the logistic regression model (*i.e.*, five demographic, three performance, two cognitive/intellectual, and eight personality; Table 18).

SPECIFIC AIM 1

Specific Aim 1 was designed to identify the demographic, performance, and psychological variables most predictive of graduation and attrition from four distinct MOS courses in an Army specialty-training program (see Methods section). The intent of this Specific Aim was to identify the most parsimonious model for each MOS using

the 18 variables identified in the previous section. The primary objective of these analyses was to provide unit leaders from an Army specialty-training program with a useful tool for improving MOS assignment. Characteristics of each MOS sample and results from the multinomial logistic regressions are discussed next.

Descriptive Data Analyses

Descriptive statistics analyses for each MOS showed comparable rates of graduation, recycle, relief, and missing data, respectively, for candidates assigned as Weapons (57.1%, 30.0%, 8.1%, and 4.7%), Engineer (53.1%, 35.3%, 5.8%, 5.8%), and Communications experts (63.1%, 22.8%, 7.2%, and 7.0%; Table 19). However, data for the Medical experts MOS was unusual and showed zero reliefs and a high rate of missing data, approximately two to three times more than the other MOSs (Table 19) (38). These findings were discussed with the research psychologist from the Army specialty-training unit who indicated that graduation is the only consistently recorded outcome for the Medical experts (MAJ E. Deagle, personal communication, June 7, 2017). MAJ Deagle speculated that the missing data for the Medical experts consisted of recycle, reliefs, and candidates still participating in the training course. After consultation with statistical experts, it was decided that data from the Medical MOS would be organized into two outcomes: graduates and Others (Table 20). The category “Others” included all data coded as missing or recycle.

Sample Characteristics

Overall, the sample of 1,460 candidates that began the MOS phase of an Army specialty-training unit had the following mean (M) scores on the predictor variables of interest (Tables 21 to 25; Appendix C): **Age** ($M = 24.89$, $SD = 3.32$), **PrevMOS** (Mode =

1), **Grade** (Mode= 4), **MOSPref** (Mode = 1), **ADTIS** ($M = 3.03$, $SD = 2.93$), **YRS ED** ($M = 13.97$, $SD = 1.76$), **APFT** score ($M = 280.38$, $SD = 14.61$), **OCourse** ($M = 90.92$, $SD = 8.09$), **Ruck March** ($M = 108.81$, $SD = 9.59$), **Land NAV** ($M = 3.19$, $SD = 0.67$), **VIQ** ($M = 109.26$, $SD = 8.75$), **PIQ** ($M = 114.44$, $SD = 8.39$), **EID** ($M = 34.79$, $SD = 4.77$), **THD** ($M = 43.22$, $SD = 5.95$), **BXD** ($M = 46.35$, $SD = 7.22$), **RC1** ($M = 41.22$, $SD = 5.85$), **RC3** ($M = 43.40$, $SD = 7.76$), **RC6** ($M = 45.49$, $SD = 5.57$), **AGGRr** ($M = 52.22$, $SD = 7.54$), and **INTRr** ($M = 44.74$, $SD = 7.00$).

Univariate analyses were conducted on all continuous variables. Levene's homogeneity of variance tests were significant for the following independent variables: ADTIS, YRS ED, Land NAV, VIQ, THD, RC3, and RC6 (Tables 27, 31, 34, 35, 38, 41, and 42). For these variables, the non-parametric Kruskal-Wallis statistic is reported.

There were three categorical variables in this study: Grade, PrevMOS, and MOSPref. Grade is the military equivalent for rank and considered an ordinal variable; therefore, grade was examined using the Kruskal-Wallis test (Table 28). The remaining categorical variables, PrevMOS and MOSPref, had cell counts with less than five observations thereby violating the Chi-square assumptions. To account for these violations, Fisher's exact tests were calculated and results are reported in Tables 29 and 30. Sample characteristics and between-group differences on the outcome variable are discussed for each MOS next and can be found in Appendix A (Tables 26 to 48). In addition to univariate and Chi-square analyses, between MOS differences (*e.g.*, Weapons versus Engineer experts) also were analyzed and can be found in Appendix A (Tables 26 to 44).

Weapons Experts

Weapons experts had an average age of 24.87 ($SD = 3.45$) years old, had 3.34 ($SD = 3.06$) years active duty time in service, were predominantly Corporals (Grade = E-4), were previously assigned to an *Operations functional area* MOS, and reported 13.66 ($SD = 1.67$) years of education. Comparisons between Weapons experts who graduated, recycled, or were relieved showed **no significant differences** for the following variables: Age, ADTIS, previous MOS, YRS ED, APFT, Ruck March, Land NAV, VIQ, EID, THD, BXD, RC1, RC3, RC6, AGGRr, and INTRr (Tables 26 to 44).

Univariate analyses and *post hoc* tests examining **PIQ** showed a significant difference between graduates ($M = 114.42$, $SD = 8.20$) and reliefs ($M = 110.56$, $SD = 7.63$) with graduates evidencing higher PIQ scores; however, the effect size was small $F(2, 419) = 4.07$, $p = 0.02$, $\eta^2 = 0.02$ (Table 36).

Engineer Experts

Engineer experts had an average age of 24.88 ($SD = 3.50$) years old, had 3.06 years ($SD = 2.97$) active duty time in service, were predominantly Corporals (Grade = E-4), were previously assigned to an *Operations functional area* MOS, and reported 13.81 ($SD = 1.74$) years of education. Comparisons between Engineer experts who graduated, recycled, or were relieved showed **no significant differences** for the following variables: PrevMOS, YRS ED, Ruck March, Land NAV, PIQ, EID, THD, BXD, RC1, RC6, AGGRr, and INTRr (Tables 26 to 44).

Univariate analyses showed **significant differences** on the outcome variable for the following predictors: **Age**, $F(2, 352) = 4.54$, $p = 0.01$, $\eta^2 = 0.03$; **ADTIS**, $H(2) = 7.54$, $p = 0.02$; **Grade**, $H(2) = 8.01$, $p = 0.02$, **APFT**, $F(2, 352) = 8.05$, $p < 0.01$, $\eta^2 =$

0.04; **VIQ**, $F(2, 352) = 4.18, p = 0.02, \eta^2 = 0.02$, and **RC3**, $F(2, 352) = 4.34, p = 0.01, \eta^2 = 0.02$ (Tables 26, 27, 28, 32, 35, 41). *Post hoc* examinations indicated that Engineers graduates ($M = 2.74, SD = 2.74$) had less time in service than recycled candidates ($M = 3.68, SD = 3.29$). Graduates also had higher APFT scores ($M = 282.39, SD = 14.79$) and were more likely to be Corporals (46%) than candidates who recycled (28.6%) or were relieved (40.9%). Relieved candidates were younger ($M = 23.18, SD = 2.94$), had lower VIQ ($M = 104.41, SD = 8.00$) scores than graduates ($M = 109.13, SD = 7.89$), and had higher RC3 (cynicism) scores ($M = 47.95, SD = 8.75$) than recycled ($M = 43.86, SD = 7.54$) and graduated candidates ($M = 43.12, SD = 7.05$; Tables 26, 27, 28, 32, 35, 41).

Medical Experts

Medical experts had an average age of 24.91 ($SD = 3.07$) years old, had 2.56 ($SD = 2.64$) years active duty time in service, were predominantly Corporals (Grade = E-4), were previously assigned as an *Operations functional area* MOS, and reported 14.52 ($SD = 1.74$) years of education. Comparisons between Medical experts who graduated and those categorized as Other showed **no significant between group differences** for the following variables: Age, ADTIS, Grade, PrevMOS, APFT, Ruck March, VIQ, PIQ, EID, THD, BXD, RC1, RC3, RC6, and AGGRr (Tables 45 to 48).

Independent-samples *t*-tests showed **graduates had significantly higher YRS ED**, ($M = 14.66, SD = 1.75$), $t(221) = 2.12, p = 0.04$, than Others, ($M = 14.06, SD = 1.66$); **lower Land NAV** scores ($M = 3.26, SD = 0.63$), $t(221) = -2.00, p = 0.05$, than Others ($M = 3.46, SD = 0.63$); and **higher INTRr** scores ($M = 46.03, SD = 7.60$), $t(221) = 4.13, p < 0.01$, than Others ($M = 41.86, SD = 5.82$; Tables 45, 46, 48).

Communications Experts

Communications experts had an average age of 24.85 ($SD = 3.17$) years old, had 2.88 ($SD = 2.87$) years active duty time in service, were predominantly Corporals (Grade = E-4), were previously assigned to an *Operations functional area* MOS, and reported 14.16 ($SD = 1.81$) years of education. Comparisons between Communications experts who graduated, recycled, or were relieved showed **no significant differences** for the following variables: Age, ADTIS, Grade, PrevMOS, YRS ED, APFT, Land NAV, VIQ, PIQ, EID, THD, BXD, RC1, RC3, RC6, AGGRr, and INTRr (Tables 26 to 44).

Univariate analyses showed **significant differences** on the outcome variable for **Ruck March**, $F(2, 385) = 3.72, p < 0.03, \eta^2 = 0.02$. *Post hoc* analyses indicated that graduates ($M = 108.33, SD = 9.29$) had faster Ruck March times than recycles ($M = 111.08, SD = 10.22$; Table 33).

Summary of Descriptive Statistical Analysis Findings

Each MOS had one or more statistically significant predictors of outcomes: Weapons = PIQ; Engineers = Age, ADTIS, Grade, APFT, VIQ, and RC3; Medical = YRS ED, Land NAV, INTRr; and Communications = Ruck March. It is noteworthy that these predictors differed among the MOSs and the significant predictors had small effect sizes (all $\eta^2 < 0.04$).

Model Building

Logistic regression models were developed using a random subsample of 60% of candidates from each MOS (Table 5). First, an examination of continuous and categorical variables using univariate and Pearson's Chi-square analyses were conducted. Variables with a p value ≤ 0.28 were retained for the model. The total number of

predictors varied by MOS and ranged from 5 to 13; however, when these variables were entered into the model using backwards stepwise procedures, numerical problems (*e.g.*, infinite parameter estimates) prevented completion of the analyses. After consultation with a statistical expert, it was decided that the best approach for variable identification at this stage was the inclusion of all 18 variables and the execution of forward stepwise procedures. Steps two through four of Hosmer's (44) model building strategy were automatically conducted by SPSS as part of the forward stepwise procedures. Entry and removal criteria were set as 0.10 and 0.15, respectively.

Next, Box-Tidwell tests were used to examine the nonlinearity of the association between each continuous independent variable and outcome for each MOS (18). Although between one and four variables emerged as significant for each MOS, a visual examination of the plots revealed no clear nonlinear patterns. Therefore, the variables were not transformed.

There were a high number of possible interaction terms for 18 variables. Based on this high number of possible interactions, the sample size, and the likelihood of Type I error, possible interaction terms were conceptually, not statistically, derived. Given the nature of the research question, Age by each personality variable (*i.e.*, MMPI-2-RF scales), were considered the most appropriate interactions to examine. Therefore, eight interactions were added to the model for a total of 26 variables (Table 18).

A total of 26 variables were used for model development on a 60% randomly selected subsample. Forward stepwise procedures were conducted in SPSS resulting in preliminary final models for each MOS (Tables 49 to 52). "Preliminary final" models

were those models that were determined based on the 60% subsample; they became “final” models if they were validated with the remaining 40% subsample.

Weapons Expert

The preliminary final model for the Weapons MOS was **significant**, $\chi^2 (2, 236) = 6.97, p = 0.03$; however model fit was poor for predicting overall training outcome. One variable, **Age** $\times$ **EID**, emerged as significant in the preliminary final model for the recycle group (OR = 1.00; 95% CI: 0.996, 1.000; $p = 0.02$; Table 49). The main effects, Age and EID, did not emerge as significant.

Engineer Expert

The preliminary final model for the Engineers' MOS was **significant**, $\chi^2 (2, 210) = 36.03, p < 0.01$; however model fit was poor for predicting overall training outcome. Two variables, **Age** (OR = 1.10; 95% CI: 1.01, 1.19; $p = 0.03$) and **APFT** (OR = 0.97; 95% CI: 0.95, 0.99; $p = 0.01$), emerged as significant in the preliminary final model, for the recycle group. Three variables emerged as significant for the relief group: **Age** (OR = 0.56; 95% CI: 0.38, 0.83; $p \leq 0.01$), **APFT** (OR = 0.96; 95% CI: 0.91, 1.00; $p = 0.04$), and **RC3** (OR = 1.13; 95% CI: 1.03, 1.23; $p = 0.01$; Table 50).

Medical Expert

The preliminary final model for the Medical MOS was **significant**, $\chi^2 (2, 140) = 10.36, p < 0.01$; however model fit was poor for predicting overall training outcome. Two variables, **RC1** (OR = 1.08; 95% CI: 1.00, 1.17; $p = 0.05$) and **INTRr** (OR = 0.93; 95% CI: 0.87, 0.99; $p = 0.03$), emerged as significant in the final model for the Other group (Table 51).

Communications Expert

The preliminary final model for the communications MOS was **not significant**, $\chi^2 (2,225) = 4.94, p = 0.09$; model fit was poor for predicting overall training outcome. **YRS ED** emerged as significant in the preliminary final model for the relief group (OR = 0.70; 95% CI: 0.49, 1.01; $p = 0.05$; Table 52).

Model Validation

The adequacy and fit of the preliminary final models were assessed by subsequent analyses with the remaining 40% subsample. Regression parameters from the proposed model were used to construct predicted values for the 40% subsample. Model findings and predictive accuracy are discussed for each MOS next.

Weapons Expert

A multinomial logistic regression analysis was conducted for Age, EID, and the interaction variable, **Age $\times$ EID**, using the 40% subsample of Weapons experts. Model fit was **not significant**, $\chi^2 (6, N = 185) = 5.62, p = 0.47$, and the model was poor for predicting overall training outcome. The interaction term **Age $\times$ EID** did not significantly predict outcomes for the recycle group, OR = 0.99; 95% CI: 0.97, 1.01; $p = 0.51$, suggesting the model may have been overfit to the 60% subsample (Table 49). The final model accurately predicted outcomes for 58.9% of the sample, with a 98.2% accuracy rate for predicting graduates, a 6.3% accuracy rate for recycles, and a 0% accuracy rate for predicting reliefs (Table 53).

Engineer Expert

A multinomial logistic regression analysis was conducted for **Age**, **APFT**, and **RC3**, using the 40% subsample of Engineer experts. Model fit was **not significant**, $\chi^2(6, 143) = 11.16, p = 0.08$, and the model was poor for predicting overall training outcome. For recycles, **APFT** emerged as **significant**, OR = 0.97; 95% CI: 0.95, 1.00; $p = 0.02$; however, **Age** (OR = 1.01; 95% CI: 0.90, 1.12; $p = 0.93$) and **RC3** were **not significant** (OR = 1.04; 95% CI: 0.99, 1.09; $p = 0.15$). For reliefs, none of the variables emerged as significant: **Age** (OR = 0.99; 95% CI: 0.81, 1.22; $p = 0.95$), **APFT** (OR = 0.97; 95% CI: 0.93, 1.01; $p = 0.10$), and **RC3** (OR = 1.06; 95% CI: 0.97, 1.15; $p = 0.22$). Regression findings indicated that for every one-point increase in a candidate's overall APFT score, the odds of recycling decreased by 3% (Table 50). The final model accurately predicted outcomes for 57.3% of the sample, with a 79.7% accuracy rate for predicting graduates, a 35.8% accuracy rate for recycles, and a 0% accuracy rate for predicting reliefs (Table 54).

Medical Expert

Multinomial logistic regression analyses were conducted for **RC1** and **INTRr** using the 40% subsample of Medical experts. Model fit was **significant**, $\chi^2(2, 83) = 6.32, p = 0.04$, and the model was poor for predicting overall training outcome. For the Other group, **INTRr** emerged as **significant** (OR = 0.91; 95% CI: 0.84, 0.99; $p = 0.02$). **RC1** (OR = 1.03; 95% CI: 0.91, 1.16; $p = 0.66$) was **not significant**. Regression findings indicated that for every one-point increase in a candidate's **INTRr** score, the odds of being classified as Other decreased by 9% (Table 51). The final model accurately

predicted outcomes for 72.3% of the sample, with a 96.8% accuracy rate for predicting graduates and a 0% accuracy rate for predicting Other (Table 55).

Communications Expert

A multinomial logistic regression analysis was conducted for YRS ED using the 40% subsample of Communication experts. Model fit was **not significant**, $\chi^2(2, 160) = 1.26, p = 0.53$. Model fit was poor for predicting overall training outcome. **YRS ED** was **not significant** in the final model for the relief group, (OR = 0.99; 95% CI: 0.74, 1.33; $p = 0.97$), suggesting the model may have been overfit to the 60% sample (Table 52). The final model accurately predicted outcomes for 70% of the sample, with a 100% accuracy rate for predicting graduates, a 0% accuracy rate for recycles and reliefs (Table 56).

Specific Aim 1 - Multinomial Logistic Regression Summary

Overall, the logistic regression model for the Engineer experts found that APFT scores were able to predict candidates who recycle; for every one point increase in a candidate's overall APFT score, the odds of recycling decreased by 3%. The final model accurately predicted outcomes for 57.3% of the sample and had higher accuracy predicting graduates (79.7%) as compared to recycles (35.8%) and reliefs (0%). The model for Medical experts found that INTRr scores were able to predict candidates categorized as Other; for every one point increase in INTRr scores, the odds of being classified as Other decreased by 9%. The model had an overall accuracy rate of 72.3% with a higher probability of predicting graduates (96.8%) than Other (0%). Logistic regression models for Weapons and Communications experts did not identify variables predictive of recycle and relief (Tables 49 to 56).

Hypothesis 1a

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS graduation** better than chance **was confirmed**. That is, the percentage of candidates who were first-time graduates were correctly predicted by the models > 79% of the time for each MOS (98.2% for Weapons; 79.7% for Engineers; 96.8% for Medical; 100% for Communications experts).

Hypothesis 1b

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS recycle** better than chance **was not confirmed**. The percentage of recycled candidates correctly predicted by the models were low.

Hypothesis 1c

The combination of demographic, military, fitness, orienteering, cognitive, and personality variables included in the dataset will predict **MOS relief** better than chance **was not confirmed**. The percentage of recycled candidates correctly predicted by the models were low.

SPECIFIC AIM 2

The purpose of Specific Aim 2 was to develop a classification and regression tree (CART) to aid psychologists at an Army specialty-training command with their MOS-assignment process. The original goal of this specific aim was to use the significant predictor variables identified in Specific Aim 1 to develop prediction rules that forecasted

graduation and attrition from each MOS. However, as described in Specific Aim 1, results from multinomial logistic regression modeling failed to validate on two (Weapons and Communications experts) of the four MOS 40% subsamples suggesting the models were overfit to the 60% sample from which they were developed. The two models (Medical and Engineer experts) that were validated on the 40% subsample showed poor rates of predicting recycles and reliefs. Consequently, there were an insufficient number of significant variables for CART execution. To address this shortcoming, and based on input from a statistical consultant, all 18 variables were entered into the CART analyses for each MOS.

Similar to Specific Aim 1, the sample was divided into 60% and 40% subsamples (Table 5). The 60% subsample was used to develop prediction rules using CART analyses and the 40% subsample was used to validate the CART model. Results of the CART analyses for each MOS are discussed next.

Classification and Regression Tree Development

Originally, the data analytic plan was to use the significant variables identified by the multinomial logistic regression as the predictor variables for CART model development. However, only two variables, one for the Engineer MOS and one for the Medical MOS, emerged as significant. As a result, all 18 variables (no interactions) were entered into the CART model for each MOS. It was not necessary to enter interactions into the analyses because CART examines the variables in a binary fashion. As such, interaction variables were omitted.

CART is robust to issues of parametricity, can handle all types of variables (continuous, categorical, and ordinal), does not require imputation of missing data, is

invariant to monotonic transformation of the predictor variables, and is robust to outliers. Therefore, no transformations or changes to the data were required prior to analyses. CART analyses began with the identification of the optimal splitting point and was followed by a forward growing and backward pruning process to create the optimal tree (90). Using the same 60% and 40% subsamples as the multinomial logistic regression, all trees or were created using the 60% subsample and tested on the 40% subsample. CART outcomes and predictive accuracy are discussed for each MOS next.

Weapons Experts

The optimal CART tree (*i.e.*, the tree with the highest predictive accuracy) for Weapons experts included two splits: **VIQ** and **YRS ED**. Tree details from the *test group* (*i.e.*, 40% subsample) showed the following starting values: graduated, $N = 110$, 59.5%; recycled, $N = 59$, 31.9%; relieved, $N = 16$, 8.6% (Figure 4).

First, the candidates were divided into two groups based on **VIQ: scores ≤ 100.50** (graduated, $N = 18$, 56.3%; recycled, $N = 12$, 37.5%; relieved, $N = 2$, 6.3%) and **scores > 100.50** (graduated, $N = 92$, 60.1%; recycled, $N = 47$, 30.7%; relieved, $N = 14$, 9.2%). Candidates with VIQ scores ≤ 100.50 were further partitioned by **YRS ED: years ≤ 12.50** (graduated, $N = 9$, 69.2%; recycled, $N = 4$, 30.8%; relieved, $N = 0$, 0%) and **YRS ED > 12.50** (graduated, $N = 9$, 47.4%; recycled, $N = 8$, 42.1%; relieved, $N = 2$, 10.5%). Using these two variables, the CART model correctly predicted 58.9% of cases overall, 91.8% of graduates (misclassified 8.2%), 13.6% of recycles (misclassified 86.4%), and 0% of reliefs (misclassified 100%; Table 57).

Engineer Experts

The optimal CART tree for Engineer experts included four splits, **AGE**, **PrevMOS**, **VIQ**, and **ADTIS**. Tree details from the *test group* showed the following starting values: graduated, $N = 79$, 55.2%; recycled, $N = 53$, 37.1%; relieved, $N = 11$, 7.7% (Figure 5).

First, the candidates were divided into two groups based on **AGE: years ≤ 21.5** (graduated, $N = 13$, 59.1%; recycled, $N = 7$, 31.8%; relieved, $N = 2$, 9.1%) and **years > 21.5** (graduated, $N = 66$, 54.5%; recycled, $N = 46$, 38.0%; relieved, $N = 9$, 7.4%). Candidates older than 21.5 were further partitioned by **PrevMOS**: candidates with previous assignments in **Operation Support/Force Sustainment** functional areas (graduated, $N = 8$, 61.5%; recycled, $N = 5$, 38.5%; relieved, $N = 0$, 0%) and **Operations functional areas** (graduated, $N = 58$, 53.7%; recycled, $N = 41$, 38.0%; relieved, $N = 9$, 8.3%). Candidates previously assigned as Operations were then split by **VIQ: scores ≤ 96.50** (graduated, $N = 3$, 37.5%; recycled, $N = 3$, 37.5%; relieved, $N = 2$, 25.0%) and **scores > 96.50** (graduated, $N = 55$, 55%; recycled, $N = 38$, 38%; relieved, $N = 7$, 7%). Candidates scoring higher than 96.5 were split one final time based on **ADTIS: years ≤ 7.30** (graduated, $N = 53$, 56.4%; recycled, $N = 34$, 36.2%; relieved, $N = 7$, 7.4%) and **years > 7.30** (graduated, $N = 2$, 33.3%; recycled, $N = 4$, 66.7%; relieved, $N = 0$, 0%).

Based on these four variables, the CART model correctly predicted 54.6% of cases overall, 83.5% of graduates (misclassified 16.5%), 22.6% of recycles (misclassified 77.4%), and 0% of reliefs (misclassified 100%; Table 58).

Medical Experts

The optimal CART tree for Medical experts produced a single split for **ADTIS** at 9 years and 5 months. Tree details from the *test group* showed the following starting values: graduated, $N = 62$, 74.7%; Other, $N = 21$, 25.3% (Figure 6).

Candidates were divided into two groups based on **ADTIS: years ≤ 9.05** (graduated, $N = 61$, 76.3%; Other, $N = 19$, 23.8%) and **years > 9.05** (graduated, $N = 1$, 33.3%; Other, $N = 2$, 66.7%). Based on one variable, the CART model correctly predicted 75.9% of cases overall, 98.4% of graduates (misclassified 1.6%) and 9.5% of Other (misclassified 90.5%; Table 59).

Communications Experts

The optimal CART tree for Communications experts produced a single split for **APFT** at **249.50** points. Tree details from the *test group* showed the following starting values: graduated, $N = 112$, 70%; recycled, $N = 32$, 20%, relieved, $N = 16$, 10% (Figure 7).

Candidates were divided into two groups based on **APFT: scores ≤ 249.50** (graduated, $N = 1$, 50%; recycled, $N = 1$, 50%, relieved, $N = 0$, 0%) and **APFT scores > 249.50** (graduated, $N = 111$, 70.3%; recycled, $N = 31$, 19.6%, relieved, $N = 16$, 10.1%). Based on one variable, the CART model correctly predicted 70% of cases overall, 99.1% of graduates (misclassified 0.9%), 3.1% of recycles (misclassified 96.9%), and 0% of reliefs (misclassified 100%; Table 60).

Additional Analyses

The CART analysis for each MOS had high rates ($> 77\%$) of misclassification for recycles and reliefs. Because of CART's poor predictive accuracy, two additional

analytical strategies were conducted: [a] additional CART analyses with adjusted probability parameters for the target class (*i.e.*, reliefs); and [b] Random Forests analyses were performed for each MOS to attempt to generate more diverse and stable tree models.

As described in the Methods section, the current study used “priors” based on the outcome frequencies of the 60% subsample (Table 19; approximately 60% [graduates], 30% [recycles], and 10% [reliefs]). “Priors” is a predetermined frequency distribution that refers to the probability that a certain outcome is presented to the tree. Based on discussion with CART analyses experts, the CART models were reanalyzed using two additional sets of “priors.” First, “priors” were set to an equal distribution (*i.e.*, 1/3, 1/3, 1/3), which equalized the misclassification rates for each outcome. In other words, each outcome had an equal probability of being selected. Second, with the “priors” still set to equal distributions, the target class was identified as reliefs and the trees with the lowest misclassifications and highest correct predictions were selected. Both additional analyses revealed reduced misclassification rates for recycles and reliefs but increased misclassification rates for graduates (Tables 61 to 64). Neither parameter adjustments improved the classification outcomes meaningfully.

Random Forests are another type of ensemble learning method. Random Forests use a single tree algorithm to grow several hundred trees using a bootstrap sample of data (51; 97). Random Forests create diverse trees by randomly sampling from all the predictor variables and providing the mode output of the classes (89). Random Forests analyses based on a *200 tree model* (*i.e.*, 200 separate trees were grown) with *balanced classes*, *four N predictors* (89), and *10 frequency progress reports* also failed to produce

a model with high correct classification and low misclassification rates (Tables 65 to 68). Overall, additional analyses using modified CART parameters (*i.e.*, different “priors”) and Random Forests provided little additional value above the original CART models which were based on 60% subsample “learn” parameters.

Specific Aim 2 – CART Summary

Overall, the classification trees correctly predicted graduates more than 80% of the time. However, the CART models also misclassified recycles and reliefs more than 77% of the time. Given the high number of graduates in the sample ($> 60\%$) compared to recycles ($\leq 35\%$) and reliefs ($< 8\%$), it is not surprising that the models were able to accurately predict graduates.

Hypothesis 2

The classification tree model will accurately classify MOS outcome (*i.e.*, graduation, recycle, or relief) better than chance **was partially confirmed**. CART validation models correctly classified graduates $> 80\%$ of the time, but failed to classify the other outcomes $> 25\%$ of the time.

CHAPTER 5: Discussion

PURPOSE

Optimally matching individuals with the military occupational specialty (MOS) assignments for which they are best suited is an important task. Improving MOS assignment has the potential to maximize individual performance, team effectiveness, and organizational stewardship by ensuring the right individuals are placed in the right jobs. This doctoral research project aimed to assist an Army specialty-training unit improve their MOS assignment process by identifying the variables most predictive of graduation and attrition from four separate MOSs (*i.e.*, Weapons, Engineer, Medical, and Communication experts). It was hypothesized that a combination of demographic, military, fitness, orienteering, cognitive, and personality variables would predict MOS graduation, recycle, and relief better than chance. For the Medical experts, there were two outcome classifications: graduation and Other (see Results section for detailed explanation). A secondary aim was to develop an MOS assignment classification tree to aid leaders with their MOS assignment decisions. It was hypothesized that the classification tree model would accurately classify MOS outcomes better than chance. This section discusses the implications and findings from multinomial logistic regression and classification and regression tree analyses (CART) using a sample of 1,460 candidates who began the MOS phase of an Army specialty-training program. This chapter is organized into the following sections: variables of interest; study findings; possible explanations for findings; study strengths, limitations, and future directions.

VARIABLES OF INTEREST

Many research studies have sought to identify a common set of characteristics among individuals selected into special operations organizations (16; 19; 21; 27; 32; 66; 78; 79; 81; 109). In their comprehensive review of the literature, Picano *et al.* (79), identified the following core attributes that best characterize individuals in high-risk operational assignments: exceptional stress tolerance, emotional stability, high levels of physical fitness, intrinsic motivation, initiative, competitive drive, exceptional reliability and integrity, sound judgment and reasoning under stress, tough minded, independent, interpersonally skilled, and more or less gregarious. These attributes emerged as important for assessment and selection (A&S) programs and many high-risk organizations specifically assess for these characteristics. Although critically important to assess for these characteristics during A&S, Picano *et al.* (79), assert that specific mission profiles may require additional attributes for success. As such, identifying personnel best suited for specialty training assignments may require the examination of additional characteristics. Findings from this study are consistent with Picano *et al.* (73) in that the core attributes important during A&S may have limited utility when seeking to identify the personnel most likely to attrite from four specialized MOS courses.

STUDY FINDINGS

The present study aimed to contribute to existing literature about personnel selection for high-risk occupations/missions by identifying whether demographic (*i.e.*, demographic and military), performance (*i.e.*, fitness and orienteering), and psychological variables (*i.e.*, cognitive and personality) collected during A&S could predict MOS

completion. This study found that the combination of variables was unable to predict attrition during the MOS phase of training.

Multinomial Logistic Regression

Using multinomial logistic regression, the predictive validity of 18 variables and eight interactions were examined for four MOSs (*i.e.*, Weapons, Engineers, Medical, and Communications experts). Overall, only two variables emerged as significantly predictive of MOS outcome. One variable, Army Physical Fitness Test (APFT), emerged as significant for the **Engineer expert MOS**; however, because the overall model was not significant this finding should be interpreted with caution. This finding suggested that for Engineer experts, lower APFT scores were associated with greater odds of recycling. Interestingly, lower APFT scores were not associated with higher odds of course relief. Examination of APFT scores showed only a five-point difference (out of 300 points) between candidates who graduated and recycled; however, the magnitude of this difference was small. Given the small effect size and the homogeneity of the sample with regard to fitness scores, it is unlikely that APFT is a meaningful differentiator between graduates and recycles.

A second variable, Introversion/Low Positive Emotionality (INTRr), emerged as significant for the **Medical expert MOS** and indicated that higher INTRr scores were associated with decreased odds of being classified as Other. Similar to APFT scores for the Engineer MOS, there was only a four-point difference (on a 100 point scale) in scores between candidates classified as graduates and Others. Introversion/Low Positive Emotionality primarily measures affective disposition. Individuals who score in the

clinical range ($T\ scores \geq 65$) are characterized as experiencing social disengagement and anhedonia (39). Both Medical graduates and Others had scores less than 50 (normal range, $T\ scores = 45$ to 57) suggesting neither group is evidencing unusual levels of affective disturbance.

Further, outcome classification problems for the Medical MOS limited interpretation of this finding. The Other category is comprised of candidates who recycled, failed, or were still enrolled in the course. Given that some candidates classified as Other may eventually graduate from the medical course without recycling, the relationship between INTRr and graduation rates must be interpreted cautiously. No other demographic, performance, or psychological variables were associated with MOS outcomes for any of the groups.

Regression models revealed poor predictive accuracy for identifying candidates classified as recycle, relief, or Other. High accuracy rates for predicting graduates were not surprising given the high number of graduates ($> 60\%$) in the sample. Additionally, the sample was highly homogenous. In cases where significant between-group differences were found, the effect sizes were small, which indicated that between-group differences were likely trivial. These findings were not surprising given that A&S programs are designed to identify candidates with a common set of core attributes who perform at the highest levels (78; 79). Given that candidates in this sample were selected based on their similarities (*e.g.*, demographic, physical, and psychological) with the personnel who comprise the organization, it was difficult to identify meaningful differences using variables collected during A&S.

Classification and Regression Tree (CART)

CART analyses were conducted to develop prediction rules for MOS assignment. The purpose of this aim was to provide leaders in an Army Specialty Unit with a tool that enabled rapid data-driven decision-making. However, no MOS classification tree produced the same predictor variables as the corresponding MOS regression model.

CART modeling for **Weapons experts** identified Verbal Intelligence Quotient (VIQ) and Years of Education (YRS ED) as optimal classification decision points. This finding differed from the regression model for Weapons experts which identified the interaction of Age $\times$ Emotional and Internalizing Dysfunction (EID) as a potential predictor during the model building stage (this interaction term was not significant in the final regression model). Similar to the Weapons expert regression model, the CART evidenced high misclassification rates for the recycles and reliefs.

CART modeling for **Engineer experts** identified Age, Previous MOS, VIQ, and Active Duty Time in Service (ADTIS) as optimal splitting variables. Age was the only variable that appeared in both the regression and CART models for the Engineer experts. Although both models identified Age as an important predictor variable, neither model had strong predictive accuracy for recycles and reliefs. The CART had high misclassification rates for recycles (77.4%) and reliefs (100%) suggesting age was not able to accurately identify the individuals least likely to graduate.

CART modeling for **Medical experts** identified a single splitting variable, ADTIS, for the optimal tree. This finding differed from the regression model for Medical experts, which initially considered Cynicism (RC3) and INTRr as predictor variables.

Although ADTIS was considered the best variable for a single-leaf tree, the model misclassified 90.5% of candidates in the Other category.

CART modeling for **Communications experts** also produced a single split, APFT, which differed from the variable YRS ED, which emerged during regression model building. Again, CART misclassified a high rate of recycles (96.9%) and reliefs (100%).

CART analyses were conducted with the intention of producing a decision tree model for MOS assignment. Although each MOS analysis produced a tree with at least one split, the high rates of misclassification for target outcomes (recycle and relief) indicated the models were unstable. Overall, the CART analyses revealed: [a] predictor variables were not consistent across the four MOSs; [b] the CART analyses provided different findings than the regression models; and [c] neither regression nor CART models meaningfully predicted or distinguished candidates most likely to recycle or fail. Further, additional analyses using modified CART model parameters (*i.e.*, changing the probability of recycles and reliefs) and Random Forests affirmed the conclusion that accurate MOS classification models cannot be developed using the 18 predictor variables.

Summary of Study Findings

The predictor variables examined in this study were selected because of their historical success at identifying the candidates most likely to complete high-risk A&S courses. However, in the present study, only two variables were able to predict MOS course attrition, APFT for Engineers and INTRr for Medics. These findings highlighted potential limitations of using data collected during A&S to predict outcomes at later

phases of the training program. For example, measures of cognitive/intellectual functioning are useful for identifying candidates most likely to be selected into high-risk training programs (27; 81; 109). However, verbal and performance intelligence scores, did not emerge as predictive of MOS outcome in any of the regression models. Statistically significant differences emerged between outcomes (*e.g.*, graduate, recycle, relief) within some MOSs (*i.e.*, Weapons and Engineer experts); however, the effect sizes were small. For instance, there was a statistically significant difference in VIQ scores between Engineers who graduated and Engineers who were relieved; graduates had higher VIQ scores. Similarly, Weapons candidates who graduated had higher PIQ scores than candidates who were relieved. Although statistically significant, the effect sizes were small.

Of note, candidates assigned to the Medical MOS had the highest PIQ and VIQ scores, but problems with the recycle and relief data prevented direct comparison of the four MOSs at each level of the outcome variable. Based on these findings, it is unlikely that VIQ and PIQ are useful measures for determining MOS assignment for the Weapons, Engineer, and Communications experts. However, based on the intellectual demands of the Medical course, individuals with lower IQs are not likely to be successful in this course. The average VIQ and PIQ scores for individuals classified in the Other category were approximately 5 points and 2 points higher (respectively) than the highest VIQ and PIQ scores in the recycle and relief groups for any of the other three MOSs. Although IQ is used as an example here, several variables produced similar results whereby statistically significant differences were accompanied by small effect sizes.

There are three possible explanations for the findings that only two variables emerged as predictive of two MOS outcomes: [a] **current instruments** do not sufficiently differentiate candidates on the characteristics of interest, [b] other, **unassessed**, performance and psychological **attributes** are associated with MOS attrition, and/or [c] **current assignment procedures are as effective as possible** at assigning candidates to MOSs and no additional improvements are needed. Each of these points is discussed next.

POSSIBLE EXPLANATION OF FINDINGS

Current Instruments

The predictor variables in this study have a history of successfully identifying candidates with the core attributes of high-risk operators. As such, several of these measurements should remain part of psychological testing batteries. In particular, measures of intelligence, personality, and motivation should be retained.

Intelligence Quotient (IQ) and Personality

The Multidimensional Aptitude Battery (MAB-II) provides information about a candidate's IQ, which is an important factor for passing selection. Similarly, the MMPI-2-RF is a useful instrument for identifying candidates with, or at risk of developing, psychopathology. As such, these measures are important screening tools for "selecting-out" vulnerable candidates (69; 78). However, with the exception of the Medics, the MMPI-2-RF was not useful in differentiating outcomes for the Weapons, Engineers, and Communications MOSs, highlighting the instrument's limited utility for informing other phases of the training course.

It is possible that other personality measures such as the *Revised NEO Personality Inventory* (NEO PI-R), the *16 Personality Factors* (16PF), and *Tailored Adaptive Personality Assessment System* (TAPAS), can provide important information about personality dimensions associated with success in more specialized segments of the training course. The aforementioned measures all assess, to some degree, the Big Five factors of personality (*i.e.*, openness, conscientiousness, extraversion, agreeableness, neuroticism). Big Five factors such as conscientiousness, openness, extraversion, and low neuroticism (*i.e.*, emotional stability) are already identified as important traits for high-risk personnel (20; 81).

In addition to providing information about the five major personality dimensions, these measures already have been used successfully by military organizations. For example, the TAPAS was effective at classifying conventional force Army personnel into MOSs. A study by Nye *et al.* (72) showed the TAPAS differentially predicted MOS in an Army sample and noted that 40% to 50% of the sample would have performed better in another job classification. Inclusion of personality measures, such as those mentioned above, that more directly assess the personality attributes that have been identified as strengths among high-risk personnel, may prove useful at identifying what personality characteristics, if any, predict success for the different MOS courses and at different stages of the training program.

Motivation

The present study examined MOS preferences as a proxy for motivation. It was assumed that individuals assigned to their first or second choice job assignments would

be more motivated than individuals assigned to their third or fourth choices. Examination of this variable showed a high rate of missing data (average of 13%), which precluded the use of this variable in the final analyses. No other measures of motivation were available. Inclusion of validated measures designed to assess various aspects of motivation (*e.g.*, intrinsic motivation, work motivation, and achievement orientation) might provide important information about a candidate's drive or willingness to succeed (50).

Kilcullen *et al.* (50) examined the attributes associated with on-the-job performance of Special Forces (SF) personnel. Demographic, cognitive, motivation, and fitness variables were included in the analyses. The SF personnel sampled were fully-fledged members of a Special Forces Operational Detachment. Motivational factors in this study included a variety of attributes including social intelligence, achievement orientation, work motivation, cognitive flexibility, and several others. The authors reported that motivation, not cognitive or fitness attributes, significantly predicted performance, highlighting the value in examining this construct. Of note, the authors commented about the possibility of ceiling effects for the cognitive and fitness variables.

In addition to improving upon existing measures, it also is important to consider that different attributes might influence success at later stages of training, where tasks are more specialized. The next section discusses attributes that may influence success during the MOS phase of the training program.

Unassessed Attributes

As stated by Picano *et al.* (79), the attributes common to high-risk personnel are important to assess when considering whether a candidate is a good fit for the organization (*e.g.*, during A&S). However, as the training course or mission requirements become more specialized, success may be contingent on other qualities. The predictor variables measured in this study were selected based on their proven history of identifying candidates most likely to graduate from high-risk A&S programs. It was hypothesized that similar characteristics (*e.g.*, intelligence, fitness, etc.) would predict MOS completion. However, those hypotheses were not fully substantiated. Based on the findings of this study, it is possible that characteristics assessed during A&S become less important than other attributes as candidates' progress through the training course. Factors such as emotional/social intelligence and team orientation, which were not assessed in this study, might be more predictive of MOS outcome than the variables analyzed.

Emotional/Social Intelligence

The military recognizes emotional intelligence (EI) as an important factor in its comprehensive fitness program (93). EI refers to a person's relational skills, ability to understand themselves and others, and the ability to adapt to the environment successfully (26; 93). Mayer *et al.* (65) identified four branches of emotional intelligence: [1] perceiving emotions, [2] facilitating thought with emotion, [3] understanding emotions, and [4] managing emotions. In their study of Army SF personnel, Kilcullen *et al.* (50), examined a similar attribute, social intelligence, which

they considered a subcomponent of motivation. Social intelligence was defined as the ability to assess or intuit the motives, feelings, and intentions of others (50).

Adaptability and the ability to navigate “human terrain” are important characteristics for some high-risk personnel who operate autonomously, without external supports, and in environments that require interactions with, or reliance on, other people, agencies, and countries (110). For example, Army SF Operational Detachment Alphas frequently engage in activities such as Foreign Internal Defense (FID), which involve high levels of social interaction and collaboration for mission accomplishment (88). This quality may be especially important among members of small teams. Including a measure that assesses emotional or social intelligence such as the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT; 65) could provide valuable information about emotional attributes that are related to success within certain job specialties.

Peer Evaluations

Peer evaluations are commonly used in high-risk training courses (*e.g.*, Ranger School and Special Forces A&S). The Army specialty-training unit regularly conducts peer evaluations because they provide unique information about each candidate’s typical level of performance and because they offer instructors additional insights into interpersonal characteristics and motivation (110).

Zazanis *et al.* (110) investigated the association between interpersonal relationships, motivation, and peer evaluations. Interestingly, the authors found that peer evaluations during A&S ($r = 0.29, p < 0.01$) were better than staff ratings ($r = 0.00, p = ns$) at predicting final training outcomes. The study also examined the relationships

between staff ratings at two time points: [a] during A&S and [b] during the training course, across five factors (*i.e.*, motivation, interpersonal performance, task performance, overall current performance, and future on-the-job performance). Peer ratings were examined at the same time points for the same five factors. Findings indicated that peer ratings during A&S were significantly correlated with peer ratings from training with medium effect sizes; conversely, staff ratings during A&S were not significantly correlated with staff ratings during training, and effect sizes were small. These results highlighted the reliability of peer evaluations and supported the assertion that they provide additional information about personality and behavioral factors, which may be unobserved by instructors at different phases of training. Of note, researchers found higher correlations among peer evaluations, as well as between peer and instructor evaluations, on assessments of task-performance; lower correlations were found on assessments of motivation and interpersonal facets (110). These findings reinforce the notion that independent assessments of social/emotional intelligence and motivation would be worthwhile additions to psychological testing batteries.

Peer evaluations indirectly capture emotional/social intelligence and motivational attributes, which are not directly examined in the current assessment paradigm.

Examination of the association between peer evaluations and MOS training outcomes may provide information about characteristics differentiating these groups. Although somewhat limited compared to peer evaluations, instructor evaluations also might provide important, and different, information about candidates most likely to recycle or be relieved from MOS training.

Instructor Evaluations

Instructor evaluations offer another source of information about candidate attributes, especially evaluations from MOS course instructors. Although instructors follow a formal grading rubric constructed specifically for each MOS, personal factors unique to each instructor (*e.g.*, team orientation, social cohesiveness, extraversion, etc.) likely influence their evaluation and characterization of student performance, at least to a small degree. Rater bias is a common phenomenon that refers to performance-rating errors as a result of evaluator attitudes, beliefs, and experiences (46; 64). In addition to whether a candidate is able to execute the technical components of a task, instructors also are implicitly assessing each candidate's "fit" as a member of a small team or within the larger organization. Instructors are assessing, albeit indirectly, factors such as emotional/social intelligence, team-orientation, and social cohesion. Unconscious biases might be influencing instructor ratings and these factors may account for some of the variance in MOS outcome. Examining instructor beliefs about why candidates' graduate, recycle, or fail might provide useful information about important attributes worthy of formal assessment. Identifying unassessed attributes via instructor interviews and focus groups may highlight additional domains for psychological examination.

Current Assignment Procedures are as Effective as Possible

The third, and perhaps simplest explanation for the poor fit of the regression and CART models in the present study, is that current assignment procedures are optimally categorizing candidates into MOS assignments. The sample examined in this study was highly homogeneous and few variables differentiated MOS outcomes. Based on the sample of candidates assessed, the recycle and relief rates may be as low as can be

accurately predicted. Therefore, it may not be possible to develop a “formula” or “algorithm” to better identify and predict which candidates will attrite during MOS training.

Summary

Although it is possible that current assignment procedures are as effective as possible, until the additional research is conducted, this explanation cannot be affirmed. Attributes important during A&S may remain important during MOS; however, it is possible that different instruments are needed to assess the personnel who have progressed to this more advanced phase of training. Additionally, skills and attributes not assessed by the A&S psychological battery might be accounting for the variance in MOS outcomes. Identifying other sources of data could prove useful to differentiate candidates. Broader assessment batteries that include measures of healthy personality functioning, motivation, emotional/social intelligence, peer evaluations, and instructor evaluations might reveal important attributes uniquely important for success in each MOS course.

STUDY STRENGTHS, LIMITATIONS AND FUTURE DIRECTIONS

Strengths

Rising attrition rates in a high-risk military organization prompted the need for a critical examination of the unit’s MOS assignment procedures. The majority of research examining high-risk occupations is restricted to the identification of characteristics that predict completion of an A&S program. There were limited published studies examining the factors associated with success at different phases of training and the factors

associated with success when assigned to operational teams. Picano *et al.* (79) noted this dearth of literature and highlighted the need for additional analyses of candidates at later stages of training. The present study was specifically designed to address this research gap and to answer the organization's questions about MOS attrition, which are major strengths.

Additionally, the present study examined each MOS separately (*i.e.*, Weapons, Engineer, Medical, and Communications experts) using data already collected by the unit. The intent was to provide the organization with ways to improve their current procedures by validating the current assessment battery. Specifically, the utility of current measures was examined and recommendations for alternative measures were provided.

Classification and regression tree (CART) analyses are gaining popularity in the field of psychology because of their robustness (*e.g.*, lack of statistical assumptions) and their ability to produce simple decision tools for researchers, health care providers, and leaders (52; 53; 68; 96; 97). In the literature review that preceded these analyses, no studies were found that examined MOS classification using CART models. The present study is the first of its kind to attempt CART analyses with the intention of providing military leaders with a data-derived decision-making tool.

Limitations and Future Directions

There were limitations of the current study. These limitations are categorized into: sample size, family-wise error, social desirability response bias, recycle and relief

performance and exit data, and database recording/accuracy. In addition to exploring these limitations, this section also proposes recommendations for future directions.

A priori power analyses indicated that 464 candidates (116 candidates per MOS), were sufficient to detect a medium effect (*i.e.*, 2.0 odds ratio) with 80% power ($\beta = .2$). The final sample exceeded the *a priori* sample size estimate with the smallest group containing 223 candidates. Despite acquiring data for more than the minimum estimated sample size, some secondary analyses indicated the study was underpowered to detect an effect. Additionally, Salford Systems (90) recommends a minimum of 100 records for CART analyses. There were only 83 cases in the Medical MOS validation group (*i.e.*, 40% subsample), which may have impacted decision tree findings. Future research should consider increasing the sample size to ensure adequate power and cases for all statistical analyses. Should a large sample size be used in future studies, Lin *et al.* (61) recommends reporting effect sizes, confidence intervals, and various charts and plots in addition to *p* values.

Another possible limitation could be the increased probability of family-wise error due to the high number of statistical analyses. To examine between-group differences using a categorical outcome variable, the predictor and outcome variables were reverse entered into SPSS for all univariate analyses. This entry pattern required that predictor variables be examined independently, which may have increased the likelihood of spurious findings when examining differences between MOSs (*e.g.*, Weapons compared to Engineers) and outcomes (*e.g.*, graduates compared to recycles). If future studies are concerned about multiple comparisons, a more stringent alpha level (34; 49) or alternative family-wise error control methods (49) should be considered.

Spurious findings during regression and CART analyses were considered unlikely because of the partitioning of data into subsamples. By randomly dividing the data into 60% and 40% subsamples, any spurious findings that emerged during model building (*i.e.*, 60% subsample) were tested during model validation (*i.e.*, 40% subsample).

Trying out for a high-risk training organization is a stressful and challenging undertaking. During A&S candidate performance is continuously monitored and evaluated. Candidates want to be perceived in the most positive, socially desirable light, which may lead to the under-reporting of symptoms. The MMPI-2-RF personality test consists of scales that assess the degree to which candidates are endorsing uncommon virtues and high levels of psychological adjustment. When these scales are elevated, it is difficult to determine whether the test-taker is genuinely functioning well or if they are under-reporting psychopathological problems. Although elevated scores on these scales are common in A&S settings (17), unusually high elevations can make it difficult to determine whether candidate MMPI-2-RF profiles are valid representations of their psychological functioning. Future studies should consider examining the influence of the validity scales on regression model outcomes.

Understanding why candidates are recycled and/or relieved is an important preliminary step when attempting to predict outcomes. The Army specialty-training unit has attempted to collect this form of data; however, information was only available for 32% of the sample. Candidates who recycle or fail are categorized based on the following: academic, administrative, character, compassionate, deceased, investigation, medical, voluntary, and overfill. In addition to ensuring these data are collected for every candidate who attrites, it also would be important to expand the current categories to

further specify outcomes. For example, of the 32% of candidates categorized to an attrition category, 22% were for academic reasons. Further specifying the academic problem (*e.g.*, failed field problem, failed hands-on training, or failed written test) might highlight trends for each MOS. Additionally, conducting exit interviews with relieved candidates might provide additional information about why the candidate attrited from the course.

A major limitation of the present study was the lack of consistently recorded outcome data for the Medical experts. The Medical MOS has historically produced one of the highest attrition rates of any MOS (38), making examination of this group especially important. However, due to inconsistent documentation of candidate data, particularly candidates who were recycled and relieved, the Medical MOS could not be fully analyzed and findings have limited interpretive value. Emphasizing standardized procedures for data recording is critical for future analyses of the Medical MOS. In addition to data recording problems for the Medics, there also were issues with the Army specialty-training unit's overall data management process.

For the past several decades, the Army specialty-training unit for which this study was conducted has collected and archived large amounts of candidate data. However, the absence of a clear data management strategy has resulted in disjointed data collection efforts and large gaps in performance data. The Army specialty-training unit is organized into separate companies, each of which is responsible for a specific phase of training.

Without an official data management Standard Operating Procedure (SOP), each company collects and records performance and testing data differently. The lack of a standardized, centralized, and integrated data collection system jeopardizes the

organization's overall effectiveness by limiting commanders' abilities to make data-driven decisions (29). For example, disparate data recording practices can limit commanders' ability to examine candidates' performance longitudinally, assess training effectiveness, and provide feedback to recruiters about the quality of the candidates assessing for the organization (29). Given the amount of time (approximately 3 months) it took to prepare the database for this project, it is understandable why psychologists and researchers are limited in their ability to rapidly analyze data and provide recommendations to unit leaders.

Efforts are already underway to develop a single comprehensive training database comprised of potentially relevant information (*e.g.*, performance, psychological, demographic, and health data). As this database is constructed, it is imperative that data recording procedures are clearly outlined and that a second, independent reviewer completes checks on the data entered. These quality assurance steps will help ensure the database is accurate and that data can be analyzed in a timely and efficient manner. It also is important that database developers and managers populate the database with pertinent and meaningful information. The database should be comprised of validated measures to support data-driven decision-making (79), as highlighted earlier in this section.

CHAPTER 6: Summary

The present study attempted to identify the characteristics most predictive of graduation, recycle, and relief from four distinct Military Occupational Specialty (MOS) courses for an Army specialty-training unit. Based on existing literature that identified the core characteristics of high-risk personnel, the present research used assessment measures from assessment and selection (A&S) to create prediction (*i.e.*, multinomial logistic regression analyses) and decision tree models (*i.e.*, classification and regression tree [CART]) for each MOS. Predictors included demographic, fitness/performance, and psychological variables for 1,460 male, active duty, enlisted Army candidates that attended training between fiscal years 2014 to 2017.

Logistic regression and CART modeling identified important variables for classifying candidates. However, both modeling strategies: [a] revealed different variables for each MOS; [b] the two data analytic strategies revealed different findings from each other; and [c] all statistical models had high rates of misclassification, indicating they were poor at identifying the candidates most likely to recycle or be relieved from their respective training courses. Overall, the A&S variables selected for inclusion in this study were not successful at predicting MOS training attrition. There are several possible explanations for these findings including: [a] more specific and sensitive assessment instruments are needed, [b] attributes other than those examined (*e.g.*, emotional intelligence) are associated with MOS training outcome, and/or [c] the unit's current MOS classification procedures are as effective as possible. After considering each of these explanations, it is recommended that the Army specialty-training unit

consider expanding their current assessment battery to include measures of healthy personality functioning, motivation, emotional/social intelligence, peer evaluations, and instructor evaluations. It also is recommended that the unit provide clear and detailed exit information (*e.g.*, why the candidate was relieved) and conduct exit surveys to gather information about the candidate's perception of his performance during MOS training.

In addition, the present study identified the need for thorough and careful data management strategies. An enormous amount of data are collected by the Army specialty-training unit throughout the entire training pipeline. The value of these data and meaningfulness of data analyses depend upon accurate data entry and careful quality assurance procedures. The present study found examples of data that needed to be cleaned for correctness by comparing responses to related questions (*e.g.*, birth year versus age). In addition, consolidation and comparison of data sets allowed for the identification and correction of many errors. Dedicated database managers, quality assurance procedures, and appropriate statistical analyses are vital for meaningful interpretation of the data.

CHAPTER 7: Conclusion

Testing batteries administered during assessment and selection (A&S) at an Army specialty-training unit produce a large number of variables, which historically have identified the candidates most likely to be selected at the completion of A&S. Although appropriate for A&S, these variables were not able to successfully predict candidate outcomes during the military occupational specialty (MOS) phase of training. Validating current A&S measures, replacing ineffective assessment tools with improved instruments, examining emotional/social intelligence and motivational domains, and improving database management may result in the identification of attributes associated with MOS training course attrition. Psychologists should work in collaboration with unit leaders to determine what information is relevant to psychological constructs (behaviors, cognitions, and motivations) of interest; how to collect meaningful information relevant to those constructs; and to ensure that the collected information is accurately recorded, analyzed, and interpreted.

APPENDIX A: Tables

Table 1. The Armed Services Vocational Aptitude Battery (ASVAB) test

Test	Description	Domain
General Science (GS)	Knowledge of physical and biological sciences	Science/ Technical
Arithmetic Reasoning (AR)	Ability to solve arithmetic word problems	Math
Work Knowledge (WK)	Ability to select the correct meaning of a word presented in context and to identify best synonym for a given word	Verbal
Paragraph Comprehension (PC)	Ability to obtain information from written passages	Verbal
Mathematics Knowledge (MK)	Knowledge of high school mathematics principles	Math
Electronics Information (EI)	Knowledge of electricity and electronics	Science/ Technical
Auto Information (AI)*	Knowledge of automobile technology	Science/ Technical
Shop Information (SI)*	Knowledge of tools and shop terminology and practices	Science/ Technical
Mechanical Comprehension (MC)	Knowledge of mechanical and physical principles	Science/ Technical
Assembling Objects (AO)	Ability to determine how an object will look when its parts are put together	Spatial

Adapted from http://official-asvab.com/docs/asvab_fact_sheet.pdf

*AI and SI are administered as separate tests in the computerized (CAT)-ASVAB, but combined into one single score (labeled AS). AI and SI are combined into one test (AS) in the paper and pencil (P&P)-ASVAB. Scores on the combined test (AS) are reported for both the CAT-ASVAB and P-P-ASVAB.

Table 2. Army ASVAB Line Scores (aptitude composites)

Test	Description
Armed Forces Qualification Test (AFQT)	Paragraph Comprehension, Word Knowledge, Mathematics Knowledge, and Arithmetic Reasoning
Clerical (CL)	Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Mathematics Knowledge
Combat (CO)	Word Knowledge, Paragraph Comprehension, Auto & Shop, and Mechanical Comprehension
Electronics (EL)	General Science, Arithmetic Reasoning, Mathematics Knowledge, and Electronic Information
Field Artillery (FA)	Arithmetic Reasoning, Mathematics Knowledge, and Mechanical Comprehension
General Maintenance (GM)	General Science, Auto & Shop, Mathematics Knowledge, and Electronic Information
General Technical (GT)	Word Knowledge, Paragraph Comprehension, and Arithmetic Reasoning
Mechanical Maintenance (MM)	Auto & Shop, Mechanical Comprehension, and Electronic Information
Operators and Food (OF)	Word Knowledge, Paragraph Comprehension, Auto & Shop, and Mechanical Comprehension
Surveillance and Communications (SC)	Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, Auto & Shop, and Mechanical Comprehension
Skilled Technical (ST)	Word Knowledge, Paragraph Comprehension, General Science, Mechanical Comprehension, and Mathematics Knowledge

Adapted from: <http://www.goarmy.com/learn/understanding-the-asvab.html>

Table 3. Predictor Variables

Demographic and Military
Age
Previous Military Occupational Specialty (PrevMOS)
Grade/Rank
Active duty time in service (ADTIS)
MOS Preferences (MOSPref)
Years of Education (YRS ED)
Fitness
Army Physical Fitness Test (APFT)
Ruck/Forced March
Obstacle Course (OCourse)
Orienteering
Land Navigation (Land NAV)
Cognitive/Intellectual
Multidimensional Aptitude Battery II (MAB-II)
Entrance General Technical (GT)
Armed Services Vocational Battery (ASVAB) score
Test of Adult Basic Education (TABE)
Personality
Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF)
<i>See Glossary for full list of terms</i>

Table 4. Outcome Variables

MOS (Weapons, Engineers, and Communications)
Graduate
Recycle
Relief
MOS (Medical)
Graduate
Other (Recycle/Relief/Missing data)

Table 5. 60% and 40% Subsample Counts

	60% random sample (expected count)	40% random sample (expected count)	Total
Weapons	250 (266)	193 (177)	443
Engineer	227 (226)	150 (151)	377
Medical	140 (134)	83 (89)	223
Communications	245 (250)	172 (167)	417

Table 6. Missing Data by Variable: Weapons Experts

Variable	Cases			
	Valid		Missing	
	N	%	N	%
Age	443	100.00	0	0.00
ADTIS	443	100.00	0	0.00
Grade	443	100.00	0	0.00
YRSED	443	100.00	0	0.00
ENT GT	441	99.55	2	0.45
APFT	443	100.00	0	0.00
PrevMOS	443	100.00	0	0.00
MOSPref	388	87.58	55	12.40
OCourse	388	87.58	55	12.42
RuckAVG	443	100.00	0	0.00
Land NAV	442	99.77	1	0.23
MAB VIQ	443	100.00	0	0.00
MAB PIQ	443	100.00	0	0.00
EID	443	100.00	0	0.00
THD	443	100.00	0	0.00
BXD	443	100.00	0	0.00
RC1	443	100.00	0	0.00
RC3	443	100.00	0	0.00
RC6	443	100.00	0	0.00
AGGRr	443	100.00	0	0.00
INTRr	443	100.00	0	0.00

Table 7. Missing Data by Variable: Engineer Experts

Variable	Cases			
	Valid		Missing	
	N	%	N	%
Age	377	100.00	0	0.00
ADTIS	376	99.73	1	0.27
Grade	377	100.00	0	0.00
YRSED	377	100.00	0	0.00
ENT GT	376	99.73	1	0.27
APFT	377	100.00	0	0.00
PrevMOS	377	100.00	0	0.00
MOSPref	339	89.90	38	10.10
OCourse	334	88.60	43	11.40
RuckAVG	377	100.00	0	0.00
Land NAV	375	99.47	2	0.53
MAB VIQ	377	100.00	0	0.00
MAB PIQ	377	100.00	0	0.00
EID	377	100.00	0	0.00
THD	377	100.00	0	0.00
BXD	377	100.00	0	0.00
RC1	377	100.00	0	0.00
RC3	377	100.00	0	0.00
RC6	377	100.00	0	0.00
AGGRr	377	100.00	0	0.00
INTRr	377	100.00	0	0.00

Table 8. Missing Data by Variable: Medical Experts

Variable	Cases			
	Valid		Missing	
	N	%	N	%
Age	223	100.00	0	0.00
ADTIS	223	100.00	0	0.00
Grade	223	100.00	0	0.00
YRSED	223	100.00	0	0.00
ENT GT	223	100.00	0	0.00
APFT	223	100.00	0	0.00
PrevMOS	223	100.00	0	0.00
MOSPref	195	87.40	28	12.60
OCourse	199	89.24	24	10.76
RuckAVG	223	100.00	0	0.00
Land NAV	223	100.00	0	0.00
MAB VIQ	223	100.00	0	0.00
MAB PIQ	223	100.00	0	0.00
EID	223	100.00	0	0.00
THD	223	100.00	0	0.00
BXD	223	100.00	0	0.00
RC1	223	100.00	0	0.00
RC3	223	100.00	0	0.00
RC6	223	100.00	0	0.00
AGGRr	223	100.00	0	0.00
INTRr	223	100.00	0	0.00

Table 9. Missing data by Variable: Communications Experts

Variable	Cases			
	Valid		Missing	
	N	%	N	%
Age	417	100.00	0	0.00
ADTIS	417	100.00	0	0.00
Grade	417	100.00	0	0.00
YRSED	417	100.00	0	0.00
ENT GT	416	99.76	1	0.24
APFT	417	100.00	0	0.00
PrevMOS	417	100.00	0	0.00
MOSPref	373	89.40	44	10.60
OCourse	368	88.25	49	11.75
RuckAVG	417	100.00	0	0.00
Land NAV	413	99.04	4	0.96
MAB VIQ	417	100.00	0	0.00
MAB PIQ	417	100.00	0	0.00
EID	417	100.00	0	0.00
THD	417	100.00	0	0.00
BXD	417	100.00	0	0.00
RC1	417	100.00	0	0.00
RC3	417	100.00	0	0.00
RC6	417	100.00	0	0.00
AGGRr	417	100.00	0	0.00
INTRr	417	100.00	0	0.00

Table 10. Skewness and Kurtosis: Demographic and Performance Variables

	Age	ADTIS	Grade	YRS ED	APFT	Ruck	Land Nav
Mean	24.89	3.03	4.18	13.97	280.38	108.81	3.19
SD	3.32	2.93	1.10	1.76	14.61	9.59	0.67
Skewness	0.80	1.23	0.16	0.43	-0.50	0.10	-0.28
SE of Skewness	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Kurtosis	0.52	0.88	-0.05	-0.83	-0.57	-0.09	-0.93
SE of Kurtosis	0.13	0.13	0.13	0.13	0.13	0.13	0.13

ADTIS: Active Duty Time in Service

YRS ED: Years Education

APFT: Army Physical Fitness Test

Land NAV: Land Navigation

Table 11. Skewness and Kurtosis: Psychological Variables

	VIQ	PIQ	EID	THD	BXD
Mean	109.26	114.44	34.79	43.22	46.35
SD	8.75	8.39	4.77	5.95	7.22
Skewness	0.18	0.02	1.38	1.11	0.48
SE of Skewness	0.06	0.06	0.06	0.06	0.06
Kurtosis	-0.39	-0.34	2.83	0.28	0.52
SE of Kurtosis	0.13	0.13	0.13	0.13	0.13

VIQ: Verbal IQ

PIQ: Performance IQ

EID: Emotional/Internalizing Dysfunction

THD: Thought Dysfunction

BXD: Behavioral/Externalizing Dysfunction

Table 12. Skewness and Kurtosis: Psychological Variables

	RC1	RC3	RC6	AGGRr	INTRr
Mean	41.22	43.40	45.49	52.22	44.74
SD	5.85	7.76	5.57	7.54	7.00
Skewness	0.95	0.96	1.97	0.86	0.66
SE of Skewness	0.06	0.06	0.06	0.06	0.06
Kurtosis	0.32	1.52	2.53	0.96	0.89
SE of Kurtosis	0.13	0.13	0.13	0.13	0.13

RC1: Somatic Complaints

RC3: Cynicism

RC6: Ideas of Persecution

AGGRr: Aggressiveness-Revised

INTRr: Introversion/Low Positive Emotionality-Revised

Table 13. Correlational Analyses: Cognitive/Intelligence Tests

	YRS ED	ENT GT	VIQ	PIQ	FSIQ
YRS ED	1	**0.43	**0.33	**0.11	**0.28
ENT GT		1	**0.71	**0.31	**0.63
VIQ			1	**0.39	**0.86
PIQ				1	**0.80
FSIQ					1

***Correlation is significant at the 0.01 level (2-tailed)*

YRS ED: Years Education

ENT GT: Entrance General Technical Score from Armed Services Vocational Battery (ASVAB)

VIQ: Verbal IQ

PIQ: Performance IQ

FSIQ: Full Scale IQ

Table 14. Correlational Analyses: Multidimensional Aptitude Battery (MAB-II)

	FSIQ	VIQ	PIQ	IN	CO	AR	SI	VO	DS	PC	SP	PA	OA
FSIQ	1	**0.86	**0.80	**0.76	**0.62	**0.58	**0.66	**0.66	**0.52	**0.53	**0.55	**0.47	**0.54
VIQ		1	**0.39	**0.86	**0.75	**0.62	**0.77	**0.82	**0.27	**0.30	**0.27	**0.21	**0.25
PIQ			1	**0.37	**0.24	**0.34	**0.29	**0.24	**0.62	**0.61	**0.69	**0.60	**0.68
IN				1	**0.60	**0.36	**0.59	**0.67	**0.24	**0.32	**0.24	**0.20	**0.21
CO					1	**0.29	**0.51	**0.58	**0.15	**0.22	**0.15	**0.13	**0.16
AR						1	**0.37	**0.34	**0.28	**0.12	**0.28	**0.16	**0.25
SI							1	**0.59	**0.21	**0.24	**0.20	**0.16	**0.19
VO								1	**0.13	**0.24	**0.14	**0.15	**0.15
DS									1	**0.17	**0.36	**0.20	**0.26
PC										1	**0.26	**0.24	**0.35
SP											1	**0.25	**0.35
PA												1	**0.27
OA													1

**Correlation is significant at the 0.01 level (2-tailed)

FSIQ: Full Scale IQ; VIQ: Verbal IQ; PIQ: Performance IQ

I: Information

CO: Comprehension

AR: Arithmetic

SI: Similarities

VO: Vocabulary

DS: Digit Span

PC: Picture Completion

SP: Spatial

PA: Picture Arrangement

OA: Object Assembly

Table 15. Correlational Analyses: Level of Education

	YRS ED	TABE READ	TABE MATH	TABE LANG
YRS ED	1	*0.05	**0.18	**0.18
TABE READ		1.0	0.03	**0.08
TABE MATH			1	**0.32
TABE LANG				1

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

Sample characteristics: Years of Education ranged from 10 to 20; TABE Reading Range: 8.3 to 12.9; TABE Math Range: 7.1 to 12.9; TABE Language Range: 1.9 to 12.9

Table 16. Correlational Analyses: Performance Variables

	APFT	OCourse	Forced March	Land NAV
APFT	1	**0.17	** -0.21	<-0.01
OCourse		1	-0.05	**0.13
Forced March			1	** -0.17
Land NAV				1

**Correlation is significant at the 0.01 level (2-tailed)

Table 17. Correlational Analyses: Personality Variables

	EID	THD	BXD	RCd	RC1	RC2	RC3	RC4	RC6	RC7	RC8	RC9	AGGRr	PSYCr	DISCr	NEGEr	INTRr
EID	1	**0.25	**0.27	**0.63	**0.30	**0.59	**0.35	**0.33	**0.25	**0.57	**0.27	**0.28	0.03	**0.25	**0.20	**0.62	**0.31
THD		1	**0.31	**0.25	**0.21	<0.01	**0.33	**0.27	**0.47	**0.36	**0.85	**0.41	**0.24	**0.90	**0.27	**0.35	-0.12
BXD			1	**0.24	**0.17	-0.02	**0.44	**0.78	**0.28	**0.44	**0.36	**0.72	**0.50	**0.29	**0.93	**0.46	**-.024
RCd				1	**0.34	**0.23	**0.31	**0.26	**0.26	**0.47	**0.27	**0.29	**0.08	**0.26	**0.19	**0.44	**0.07
RC1					1	**0.06	**0.22	**0.20	**0.19	**0.28	**0.21	**0.20	**0.09	**0.21	**0.13	**0.30	-0.04
RC2						1	0.05	**0.12	0.04	**0.15	<-0.01	**-.014	**-.031	0.02	-0.05	**0.13	**0.57
RC3							1	**0.37	**0.32	**0.45	**0.35	**0.55	**0.43	**0.34	**0.39	**0.45	-0.05
RC4								1	**0.26	**0.42	**0.33	**0.45	**0.33	**0.26	**0.74	**0.44	-0.04
RC6									1	**0.34	**0.30	**0.34	**0.23	**0.39	**0.23	**0.33	*-.005
RC7										1	**0.42	**0.52	**0.29	**0.36	**0.37	**0.73	-0.04
RC8											1	**0.45	**0.24	**0.85	**0.33	**0.40	**-.015
RC9												1	**0.65	**0.39	**0.65	**0.54	**-.050
AGGRr													1	**0.21	**0.41	**0.32	**-.027
PSYCr														1	**0.26	**0.34	**-.011
DISCr															1	**0.39	**-.025
NEGEr																1	**-.009
INTRr																	1

**Correlation is significant at the 0.01 level (2-tailed)

EID: Emotional/Internalizing Dysfunction; THD: Thought Dysfunction; BXD: Behavioral/Externalizing Dysfunction; RCd: Demoralization; RC1: Somatic Complaints; RC2: Low Positive Emotions; RC3: Cynicism; RC4: Antisocial Behavior; RC6: Ideas of Persecution; RC7: Dysfunctional Negative Emotions; RC8: Aberrant Experiences; RC9: Hypomanic Activation; AGGRr: Aggressiveness-Revised; PSYCr: Psychoticism-Revised; DISCr: Disconstraint-Revised; NEGEr: Negative Emotionality/Neuroticism-Revised; INTRr: Introversion/Low Positive Emotionality-Revised

Table 18. Final Predictor Variables

Demographic and Military
Age
Grade/Rank
Active duty time in service (ADTIS)
Previous Military Occupational Specialty (PrevMOS)
Years of education (YRS ED)
Performance
Army Physical Fitness Test (APFT)
Ruck/Forced March
Land Navigation (Land NAV)
Cognitive/Intellectual (<i>Multidimensional Aptitude Battery II [MAB-II]</i>)
Verbal Intelligence Quotient (VIQ)
Performance Intelligence Quotient (PIQ)
Personality (<i>Minnesota Multiphasic Personality Inventory-2 Restructured Form [MMPI-2-RF]</i>)
Emotional/Internalizing Dysfunction (EID)
Thought Dysfunction (THD)
Behavioral/Externalizing Dysfunction (BXD)
Somatic Complaints (RC1)
Cynicism (RC3)
Ideas of Persecution (RC6)
Aggressiveness-Revised (AGGRr)
Introversion/Low Positive Emotionality-Revised (INTRr)
Interaction terms for Regression Model
Age x Emotional/Internalizing Dysfunction (Age x EID)
Age x Thought Dysfunction (Age x THD)
Age x Behavioral/Externalizing Dysfunction (Age x BXD)
Age x Somatic Complaints (Age x RC1)
Age x Cynicism (Age x RC3)
Age x Ideas of Persecution (Age x RC6)
Age x Aggressiveness-Revised (Age x AGGRr)
Age x Introversion/Low Positive Emotionality-Revised (Age x INTRr)
<i>See Glossary for full list of terms</i>

Table 19. Outcomes: All Military Occupational Specialties (MOS)

	Weapons	Engineer	Medical	Communications	Total
Graduate	253 (57.10)	200 (53.10)	174 (78.00)	263 (63.10)	890 (61.00)
Recycle	133 (30.00)	133 (35.30)	7 (3.10)	95 (22.80)	368 (25.20)
Relief	36 (8.10)	22 (5.80)		30 (7.20)	88 (6.00)
Missing	21 (4.70)	22 (5.80)	42 (18.80)	29 (7.00)	114 (8.00)
Overall	443 (100.00)	377 (100.00)	223 (100.00)	417 (100.00)	1460 (100.00)

Graduate: Graduated first attempt

Recycle: One or more recycles; candidate eventually graduates

Relief: Reclass or relief

Table 20. Outcomes: Medical Experts

Graduate	174 (78.00)
Other	49 (22.00)
Total	223 (100.00)

Graduate: Graduated first attempt

*Other: Recycled, reclassified,
relieved, or incomplete*

Table 21. Demographic Continuous Variables (All MOS)

	Age	ADTIS	YRS ED	ENT GT
Mean	24.89	3.03	13.97	118.48
SD	3.32	2.93	1.76	11.06

Table 22. Demographic Categorical Variables (All MOS)

	PrevMOS	Grade	MOSPref
Mode	1	4	1

PrevMOS = 1: Operations functional area
Grade = 4: Equivalent to the rank of E4, Corporal

Table 23. Performance Variables (All MOS)

	APFT	OCourse	Ruck	Land NAV
Mean	280.38	90.92	108.81	3.19
SD	14.61	8.09	9.59	0.67

Table 24. Psychological Variables

	VIQ	PIQ	EID	THD	BXD
Mean	109.26	114.44	34.79	43.22	46.35
SD	8.75	8.39	4.77	5.95	7.22

Table 25. Psychological Variables

	RC1	RC3	RC6	AGGRr	INTRr
Mean	41.22	43.40	45.49	52.22	44.74
SD	5.85	7.76	5.57	7.54	7.00

Table 26. Age (*Mean, SD*)

	Weapons	Engineer	Communications	Overall	F statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	24.84 (3.49)	24.71 (3.49)	24.83 (3.21)	24.84 (3.35)	0.10 (0.90)	0.07 (<0.01)
Recycle	24.86 (3.34)	25.43 (3.51)	24.89 (3.10)	25.06 (3.33)	1.15 (0.32)	0.25 (<0.01)
Relief	25.06 (3.65)	23.18 (2.94)	24.87 (3.15)	24.52 (3.37)	2.43 (0.10)	0.48 (0.05)
Missing	24.71 (2.28)	25.55 (3.40)	25.07 (3.50)	24.98 (2.99)	0.38 (0.69)	0.11 (0.01)
†Overall	24.87 (3.45)	24.88 (3.50)	24.85 (3.17)	24.89 (3.32)	0.04 (0.96)	0.06 (<0.01)
F statistic (p value)	0.06 (0.94)	4.54 (*0.01)	0.01 (0.99)	1.31 (0.27)		
Observed Power (Partial Eta Sq)	0.06 (<0.00)	0.77 (0.03)	0.05 (< 0.01)	0.29 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences between Engineer recycles and reliefs; relieved Engineers were youngest

Table 27. Active Duty Time in Service (ADTIS; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	3.50 (3.07)	2.74 (2.74)	2.67 (2.72)	2.87 (2.81)	6.42 (<*0.01)	0.90 (0.02)
Recycle	3.05 (2.93)	3.68 (3.29)	3.21 (3.13)	3.30 (3.14)	1.43 (0.24)	0.31 (<0.01)
Relief	3.27 (3.45)	2.22 (2.42)	2.92 (2.57)	2.89 (2.93)	0.88 (0.42)	0.20 (0.02)
Missing	2.54 (2.34)	4.68 (3.47)	3.68 (3.42)	3.44 (3.15)	2.40 (0.10)	0.47 (0.07)
†Overall	3.34 (3.06)	3.06 (2.97)	2.88 (2.87)	3.03 (2.93)	2.14 (0.12)	0.44 (<0.01)
Levene's F (p value)	1.34 (0.26)	6.42 (<*0.01)	2.09 (0.13)	2.25 (0.11)		
Kruskal-Wallis [KW] (p value)	3.23 (0.20)	7.54 (*0.02)	2.75 (0.25)	2.46 (0.29)		

*Statistically significant

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences: 1) Weapons graduates had more ADTIS than Engineers and communications graduates, and 2) within the Engineer MOS, graduates had less ADTIS than recycles

Table 28. Grade/Rank (*Mode, Valid%*)

	Weapons	Engineer	Communications	Overall	KW (<i>p</i> value)
Graduate	4 (36.80)	4 (46.00)	4 (47.90)	4 (43.40)	4.17 (0.12)
Recycle	4 (34.60)	4 (28.60)	4 (43.20)	4 (34.60)	2.74 (0.26)
Relief	†3 (25.00)	3 (40.90)	4 (56.70)	4 (35.20)	1.73 (0.42)
Missing	4 (42.90)	4 (45.50)	4 (37.90)	4 (41.70)	2.62 (0.27)
Overall	4 (35.40)	4 (38.50)	4 (46.80)	4 (40.20)	0.24 (0.89)
KW (<i>p</i> value)	2.09 (0.35)	8.01 (*0.02)	0.08 (0.96)	2.51 (0.29)	

*Statistically significant

†Multiple modes, lowest shown

Post hoc analyses showed statistically significant differences: 1) Engineer graduates had higher rank than recycles and failures, and 2) Engineer recycles had higher rank than reliefs

Table 29. Previous MOS (PrevMOS; *Mode, Valid%*)

	Weapons	Engineer	Communications	Overall	Fisher's exact
Graduate	1 (91.30)	1 (93.00)	1 (87.80)	1 (90.50)	0.09
Recycle	1 (90.20)	1 (85.70)	1 (90.50)	1 (88.60)	0.70
Relief	1 (88.90)	1 (95.50)	1 (90.00)	1 (90.90)	1.00
Missing	1 (95.20)	1 (77.30)	1 (86.20)	1 (86.10)	0.10
Overall	1 (91.00)	1 (89.70)	1 (88.50)	1 (89.70)	
Chi Square	0.57 (0.97)	5.85 (0.21)	0.89 (0.93)	1.44 (0.84)	
(p value)					
Fisher's exact	0.89	0.25	0.96		

Table 30. MOS Preferences (MOSPref; *Mode, Valid%*)

	Weapons	Engineer	Communications	Overall	Fisher's exact
Graduate	1 (78.80)	1 (86.20)	1 (49.00)	1 (72.20)	<*0.01
Recycle	1 (76.90)	1 (85.70)	1 (64.40)	1 (76.80)	*0.01
Relief	1 (78.80)	1 (85.00)	1 (53.60)	1 (71.60)	*0.02
Missing	1 (76.20)	1 (89.50)	1 (70.80)	1 (78.10)	0.29
Overall	1 (78.10)	1 (86.10)	1 (58.20)	1 (73.80)	
Chi Square	3.17 (0.79)	0.84 (0.99)	4.10 (0.66)	2.90 (0.82)	
(p value)					
Fisher's exact	0.84	0.97	0.68		

*Statistically significant

Table 31. Years of Education (YRS ED; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	13.56 (1.63)	13.98 (1.77)	14.29 (1.85)	14.08 (1.79)	5.28 (*0.01)	20.40 (<*0.01)
Recycle	13.70 (1.73)	13.77 (1.74)	13.95 (1.79)	13.77 (1.74)	0.13 (0.88)	0.94 (0.63)
Relief	13.72 (1.70)	13.18 (1.18)	13.80 (1.75)	13.61 (1.61)	3.25 (*0.04)	1.36 (0.51)
Missing	14.52 (1.69)	13.18 (1.65)	14.03 (1.45)	14.04 (1.67)	1.00 (0.37)	8.61 (*0.01)
†Overall	13.66 (1.67)	3.81 (1.74)	14.16 (1.81)	13.97 (1.76)	3.10 (*0.05)	16.87 (<*0.01)
Levene's F (p value)	0.57 (0.57)	4.87 (*<0.01)	0.37 (0.69)	1.60 (0.20)		
KW (p value)	0.75 (0.69)	3.64 (0.16)	3.34 (0.19)	3.06 (0.22)		

*Statistically significant

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences between Weapons, Engineers, and Communications graduates; Weapons graduates had fewer years of education

Table 32. Army Physical Fitness Test score (APFT; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	280.09 (15.89)	282.39 (14.79)	281.22 (13.97)	281.20 (14.56)	1.32 (0.27)	0.29 (<0.01)
Recycle	280.14 (14.46)	276.65 (14.46)	278.79 (15.85)	278.49 (14.78)	1.86 (0.16)	0.39 (0.01)
Relief	278.75 (15.25)	273.73 (14.98)	277.90 (14.89)	277.20 (15.03)	0.81 (0.45)	0.18 (0.02)
Missing	284.10 (13.00)	285.18 (12.85)	281.14 (13.21)	282.47 (13.23)	0.67 (0.52)	0.16 (0.02)
*Overall	280.19 (15.27)	280.02 (14.89)	280.42 (14.44)	280.38 (14.61)	†††0.07 (0.93)	0.06 (<0.01)
F statistic (p value)	0.13 (0.88)	8.05 (<*0.01)	1.44 (0.24)	††5.50 (<*0.01)		
Observed Power (Partial Eta Sq)	0.07 (<0.01)	0.96 (0.04)	0.31 (<0.01)	0.85 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

††F statistic comparing outcomes

†††F statistic comparing MOSs

Post hoc analyses showed statistically significant differences: 1) Engineer MOS graduates had higher APFT scores than recycles and reliefs, and 2) across MOSs, graduates had higher APFT scores than recycles and reliefs

Table 33. Ruck/Forced March (*Mean, SD*)

	Weapons	Engineer	Communications	Overall	F statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	108.57 (10.19)	108.18 (9.56)	108.33 (9.29)	108.29 (9.59)	0.10 (0.91)	0.07 (<0.01)
Recycle	108.67 (9.47)	109.22 (9.88)	111.08 (10.22)	109.52 (9.86)	1.76 (0.17)	0.37 (0.01)
Relief	110.39 (8.87)	109.30 (9.26)	111.22 (7.62)	110.40 (8.50)	0.32 (0.73)	0.10 (<0.01)
Missing	106.91 (8.62)	110.25 (9.91)	111.65 (10.00)	109.43 (9.24)	1.52 (0.23)	0.31 (0.04)
†Overall	108.67 (9.80)	108.73 (9.66)	109.40 (9.52)	108.81 (9.59)	0.72 (0.49)	0.17 (<0.01)
F statistic (p value)	0.54 (0.58)	0.52 (0.59)	3.72 (*0.03)	2.85 (0.06)		
Observed Power (Partial Eta Sq)	0.14 (<0.01)	0.14 (<0.01)	0.68 (0.02)	0.56 (<0.01)		

†Means and SDs calculated without missing data

Post hoc analyses showed MOS graduates had significantly different scores than recycles; graduates had faster ruck march times

Post hoc analyses showed statistically significant differences between Communications graduates and recycles; graduates had faster ruck march times

Table 34. Land Navigation (Land NAV; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	3.21 (0.62)	3.28 (0.63)	3.09 (0.69)	3.20 (0.65)	1.28 (0.28)	9.40 (<*0.01)
Recycle	3.21 (0.64)	3.16 (0.74)	3.17 (0.74)	3.19 (0.65)	2.28 (0.10)	0.34 (0.84)
Relief	2.96 (0.66)	3.05 (0.67)	3.10 (0.69)	3.03 (0.67)	0.27 (0.77)	0.77 (0.68)
Missing	3.10 (0.60)	3.09 (0.80)	3.20 (0.84)	3.26 (0.73)	2.06 (0.14)	0.69 (0.71)
†Overall	3.19 (0.63)	3.21 (0.67)	3.12 (0.71)	3.19 (0.67)	3.97 (*0.02)	3.94 (0.14)
F statistic (p value)	2.70 (0.07)	2.02 (0.13)	0.46 (0.64)	2.26 (0.11)		
Observed Power (Partial Eta Sq)	0.53 (0.01)	0.42 (0.01)	0.12 (<0.01)	0.46 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences between Engineer graduates and Communications and Weapons graduates; Engineers had higher land navigation scores

Table 35. Verbal IQ (VIQ; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	107.94 (7.88)	109.13 (7.89)	108.81 (9.16)	109.93 (8.69)	2.96 (*0.05)	2.81 (0.25)
Recycle	106.77 (8.38)	107.51 (8.53)	109.65 (9.79)	107.90 (8.88)	1.54 (0.22)	4.78 (0.09)
Relief	106.22 (6.65)	104.41 (8.00)	108.20 (9.47)	106.44 (8.08)	1.56 (0.22)	2.10 (0.35)
Missing	105.24 (5.27)	106.14 (7.85)	110.66 (8.60)	110.56 (8.32)	3.23 (*0.05)	5.52 (0.06)
†Overall	107.32 (7.86)	108.11 (8.20)	109.09 (9.28)	109.26 (8.75)	4.95 (<*0.01)	7.03 (*0.03)
F statistic (p value)	1.39 (0.25)	4.18 (*0.02)	0.39 (0.68)	††3.03 (*0.05)		
Observed Power (Partial Eta Sq)	0.30 (<0.01)	0.73 (0.02)	0.11 (<0.01)	0.59 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

††No significant post hoc tests

Post hoc analyses showed statistically significant differences between Weapons experts and Communications experts; Weapons experts had lower VIQ scores

Post hoc analyses showed statistically significant differences between Engineers who graduated and were relieved; Engineer graduates had higher VIQ scores

Table 36. Performance IQ (PIQ; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	114.42 (8.20)	115.08 (7.98)	113.12 (8.58)	114.75 (8.32)	3.42 (*0.03)	0.64 (<0.01)
Recycle	113.23 (7.58)	114.59 (8.48)	114.28 (8.67)	114.06 (8.19)	0.98 (0.38)	0.22 (<0.01)
Relief	110.56 (7.63)	110.77 (5.66)	112.90 (8.93)	111.41 (7.68)	0.86 (0.43)	0.19 (0.02)
Missing	114.19 (8.86)	114.05 (9.51)	114.69 (10.33)	115.60 (9.54)	0.03 (0.97)	0.06 (<0.01)
†Overall	113.74 (8.05)	114.59 (8.17)	113.48 (8.75)	114.44 (8.39)	1.92 (0.15)	0.40 (<0.01)
F statistic (p value)	4.07 (*0.02)	2.83 (0.06)	0.69 (0.50)	4.32 (*0.01)		
Observed Power (Partial Eta Sq)	0.72 (0.02)	0.55 (0.02)	0.17 (<0.01)	0.75 (<0.01)		

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences between: 1) Weapons graduates and reliefs; graduates had higher PIQ scores, 2) Engineers and Communications graduates; Engineers had higher PIQ scores, and 3) reliefs (overall) and graduates and recycles; reliefs had lower PIQ scores

Table 37. Emotional/Internalizing Dysfunction (EID; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	34.47 (4.82)	34.74 (4.56)	35.08 (4.83)	34.78 (4.77)	1.04 (0.35)	0.23 (<0.01)
Recycle	33.95 (4.28)	35.26 (4.99)	35.26 (4.99)	34.77 (4.77)	††3.22 (*0.04)	0.61 (0.02)
Relief	35.36 (6.38)	34.50 (4.23)	36.57 (4.39)	35.56 (5.27)	1.02 (0.37)	0.22 (0.02)
Missing	35.38 (4.79)	34.32 (3.58)	33.97 (4.61)	34.32 (4.33)	0.66 (0.52)	0.16 (0.02)
†Overall	34.43 (4.82)	34.88 (4.64)	35.15 (4.83)	34.79 (4.77)	2.49 (0.08)	0.50 (<0.01)
F statistic (p value)	1.33 (0.27)	0.58 (0.56)	1.28 (0.28)	1.09 (0.34)		
Observed Power (Partial Eta Sq)	0.29 (<0.01)	0.15 (<0.01)	0.28 (<0.01)	0.24 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

††Post hocs do not show significant differences between groups on recycle

Table 38. Thought Dysfunction (THD; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	43.27 (6.04)	42.83 (5.42)	43.57 (6.11)	43.04 (5.79)	3.53 (*0.03)	1.27 (0.53)
Recycle	43.11 (6.16)	44.03 (6.21)	43.46 (5.94)	43.53 (6.09)	0.40 (0.67)	2.03 (0.36)
Relief	42.56 (5.67)	46.00 (8.52)	45.63 (7.33)	44.47 (7.14)	3.85 (*0.03)	4.23 (0.12)
Missing	42.38 (6.00)	42.50 (6.24)	42.31 (5.20)	42.66 (5.64)	0.19 (0.83)	0.02 (0.99)
†Overall	43.12 (6.03)	43.42 (6.01)	43.61 (6.12)	43.44 (6.07)	1.02 (0.36)	1.84 (0.40)
Levene's F (p value)	0.76 (0.47)	7.83 (<*0.01)	2.89 (0.06)	4.46 (*0.01)		
KW (p value)	0.49 (0.78)	5.01 (0.08)	2.55 (0.28)	1.99 (0.37)		

*Statistically significant

Table 39. Behavioral/Externalizing Dysfunction (BXD; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	46.94 (7.92)	46.55 (8.03)	45.60 (6.87)	46.26 (7.40)	2.13 (0.12)	0.44 (<0.01)
Recycle	46.41 (6.66)	46.68 (7.81)	45.64 (6.51)	46.28 (7.04)	0.62 (0.54)	0.15 (<0.01)
Relief	46.72 (6.18)	48.18 (7.75)	48.40 (8.56)	47.66 (7.41)	0.49 (0.62)	0.13 (0.01)
Missing	46.52 (6.10)	45.41 (6.65)	45.45 (5.72)	46.29 (6.05)	0.24 (0.79)	0.09 (<0.01)
†Overall	46.74 (7.33)	46.63 (7.85)	45.80 (6.86)	46.35 (7.22)	2.05 (0.13)	0.42 (<0.01)
F statistic (p value)	0.23 (0.79)	0.42 (0.66)	2.25 (0.11)	1.31 (0.27)		
Observed Power (Partial Eta Sq)	0.09 (<0.01)	0.12 (<0.01)	0.46 (0.01)	0.28 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

Table 40. Somatic Complaints (RC1; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	41.39 (6.16)	41.14 (5.73)	41.44 (6.18)	40.99 (5.85)	0.16 (0.85)	0.08 (<0.01)
Recycle	41.47 (5.77)	41.96 (6.39)	41.38 (5.56)	41.60 (5.91)	0.34 (0.71)	0.10 (<0.01)
Relief	40.67 (5.68)	42.95 (5.19)	42.17 (5.83)	41.75 (5.63)	1.26 (0.29)	0.27 (0.03)
Missing	42.33 (5.99)	40.45 (5.29)	41.59 (6.02)	41.38 (5.72)	0.58 (0.57)	0.14 (0.02)
†Overall	41.40 (5.99)	41.49 (5.92)	41.49 (5.99)	41.22 (5.85)	0.03 (0.97)	0.06 (<0.01)
F statistic (p value)	0.27 (0.77)	1.42 (0.24)	0.21 (0.81)	0.40 (0.67)		
Observed Power (Partial Eta Sq)	0.09 (<0.01)	0.30 (<0.01)	0.08 (<0.01)	0.11 (<0.01)		

†Means and SDs calculated without missing data

Table 41. Cynicism (RC3; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	43.87 (8.56)	43.12 (7.05)	43.38 (8.10)	43.26 (7.76)	3.07 (*0.05)	0.36 (0.84)
Recycle	43.49 (8.00)	43.86 (7.54)	42.95 (7.23)	43.52 (7.65)	0.61 (0.55)	1.23 (0.54)
Relief	43.33 (8.44)	47.95 (8.75)	45.77 (8.83)	45.32 (8.75)	0.04 (0.96)	4.59 (0.10)
Missing	42.14 (7.57)	42.77 (7.55)	41.48 (6.57)	42.58 (7.15)	0.16 (0.85)	0.22 (0.89)
†Overall	43.63 (8.33)	43.64 (7.42)	43.32 (7.89)	43.40 (7.76)	1.01 (0.37)	1.10 (0.58)
F statistic (p value)	0.13 (0.88)	4.34 (*0.01)	1.47 (0.23)	2.18 (0.11)		
Observed Power (Partial Eta Sq)	0.07 (<0.01)	0.75 (0.02)	0.31 (<0.01)	0.45 (<0.01)		

*Statistically significant

†Means and SDs calculated without missing data

Post hoc analyses showed statistically significant differences between Engineer reliefs and Engineer graduates and recycles; reliefs had higher cynicism scores

Table 42. Ideas of Persecution (RC6; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	Levene's F (p value)	KW (p value)
Graduate	45.45 (5.49)	45.48 (5.46)	45.54 (5.73)	45.42 (5.49)	0.09 (0.92)	0.01 (1.00)
Recycle	45.51 (5.63)	46.05 (6.26)	45.07 (5.08)	45.55 (5.69)	3.48 (*0.03)	1.33 (0.51)
Relief	44.44 (4.14)	46.64 (7.14)	47.83 (7.20)	46.15 (6.22)	11.66 (<*0.01)	4.93 (0.09)
Missing	46.57 (6.66)	45.23 (5.81)	44.79 (4.56)	45.34 (5.33)	2.47 (0.09)	1.16 (0.56)
†Overall	45.44 (5.49)	45.73 (5.87)	45.54 (5.65)	45.57 (5.67)	1.18 (0.31)	0.40 (0.82)
Levene's F (p value)	2.83 (0.06)	2.81 (0.06)	6.79 (<*0.01)	1.92 (0.15)		
KW (p value)	1.03 (0.60)	0.82 (0.66)	5.48 (0.07)	0.95 (0.62)		

*Statistically significant

†Means and SDs calculated without missing data

Table 43. Aggressiveness - revised (AGGRr; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	53.27 (8.28)	52.21 (7.33)	51.96 (7.54)	52.30 (7.78)	2.04 (0.13)	0.42 (<0.01)
Recycle	52.47 (7.41)	51.82 (7.05)	51.49 (7.02)	52.04 (7.14)	0.56 (0.58)	0.14 (<0.01)
Relief	52.14 (6.72)	51.41 (6.89)	53.80 (8.60)	52.52 (7.43)	0.73 (0.48)	0.17 (0.02)
Missing	52.48 (7.72)	51.45 (8.51)	50.93 (6.55)	51.96 (7.05)	0.26 (0.77)	0.09 (<0.01)
†Overall	52.90 (7.87)	51.98 (7.26)	51.91 (7.44)	52.22 (7.54)	2.29 (0.10)	0.47 (<0.01)
F statistic (p value)	0.65 (0.52)	0.20 (0.82)	1.08 (0.34)	0.60 (0.55)		
Observed Power (Partial Eta Sq)	0.16 (<0.01)	0.08 (<0.01)	0.24 (<0.01)	0.15 (<0.01)		

†Means and SDs calculated without missing data

Table 44. Introversion/Low Positive Emotionality - revised (INTRr; *Mean, SD*)

	Weapons	Engineer	Communications	Overall	F Statistic (p value)	Observed Power (Partial Eta Sq)
Graduate	44.55 (6.71)	44.54 (6.42)	44.95 (7.43)	44.96 (7.06)	0.28 (0.76)	0.09 (<0.01)
Recycle	44.29 (6.77)	44.86 (6.90)	45.71 (6.77)	44.80 (6.80)	1.19 (0.31)	0.26 (<0.01)
Relief	45.33 (7.56)	42.09 (4.90)	46.13 (8.91)	44.80 (7.60)	1.99 (0.14)	0.40 (0.05)
Missing	44.71 (7.89)	42.00 (5.39)	43.41 (6.60)	42.86 (6.46)	0.89 (0.42)	0.20 (0.03)
†Overall	44.55 (6.84)	44.36 (6.50)	45.10 (7.35)	44.74 (7.00)	1.24 (0.29)	0.27 (<0.01)
F statistic (p value)	0.33 (0.72)	1.71 (0.18)	0.61 (0.54)	0.08 (0.92)		
Observed Power (Partial Eta Sq)	0.10 (<0.01)	0.36 (0.01)	0.15 (<0.01)	0.06 (<0.01)		

†Means and SDs calculated without missing data

Table 45. Demographic Variables (Medical Experts)

	Age (Mean, SD)	ADTIS (Mean, SD)	Grade (Mode, %)	Previous MOS (Mode, %)	MOSPref (Mode, %)	Yrs Ed (Mean, SD)
Graduate	24.98 (3.17)	2.43 (2.47)	4 (47.7%)	1 (82.2%)	1 (87.4%)	14.66 (1.75)
Other	24.67 (2.72)	3.00 (3.18)	4 (53.1%)	1 (79.6%)	1 (86.4%)	14.06 (1.66)
Overall	24.91 (3.07)	2.56 (2.64)	4 (48.9%)	1 (81.6%)	1 (87.2%)	14.52 (1.74)
T Score	0.61 (0.54)	-1.33 (0.18)				2.12 (*0.04)
(p value)						
Chi-Square			††2.78 (0.60)	†0.84	†0.81	
(p value)						

*Statistically significant

†Fisher's exact test reported

††Kruskal-Wallis statistic reported (ordinal variable)

Table 46. Performance Variables (Medical Experts; Mean, SD)

	APFT	Ruck	Land Nav
Graduate	281.43 (13.10)	107.93 (9.24)	3.26 (0.63)
Other	280.71 (13.15)	108.90 (8.95)	3.46 (0.63)
Overall	281.27 (13.08)	108.14 (9.17)	3.30 (0.63)
T Score (p value)	0.34 (0.74)	-0.66 (0.51)	-2.00 (*0.05)

*Statistically significant

Table 47. Cognitive/Intellectual (Medical Experts; *Mean, SD*)

	VIQ	PIQ
Graduate	115.43 (7.77)	117.31 (7.89)
Other	115.08 (6.90)	117.61 (8.92)
Overall	115.35 (7.57)	117.38 (8.11)
T Score (<i>p</i> value)	0.29 (0.78)	-0.23 (0.82)

Table 48. Personality Variables (Medical Experts, *Mean, SD*)

	EID	THD	BXD	RC1	RC3	RC6	AGGRr	INTRr
Graduate	34.84 (4.85)	42.16 (5.26)	45.90 (6.57)	39.56 (4.76)	42.36 (6.68)	45.14 (5.20)	51.50 (7.81)	46.03 (7.60)
Other	34.08 (4.37)	43.08 (5.45)	46.88 (5.95)	41.08 (5.50)	43.76 (7.49)	44.86 (4.60)	53.08 (6.14)	41.86 (5.82)
Overall	34.67 (4.75)	42.36 (5.30)	46.12 (6.44)	39.90 (4.96)	42.66 (6.87)	45.08 (5.07)	51.85 (7.49)	45.11 (7.44)
T Score (<i>p</i> value)	0.99 (0.33)	-1.08 (0.28)	-0.94 (0.35)	-1.91 (0.06)	-1.26 (0.21)	0.35 (0.73)	-1.31 (0.19)	†4.13 (<*0.01)

*Statistically significant

†Equal variances not assumed

Table 49. Multinomial Logistic Regression Models: Weapons Experts

					Model		Validation				
					Chi-Square (<i>p</i> value)		6.97 (*0.03)		5.62 (0.47)		
					Cox and Snell Pseudo R-Square		0.03		0.03		
					Nagelkerke Pseudo R-Square		0.04		0.04		
		Model Building				Validation					
		Odds Ratio		95% Confidence		Odds Ratio		95% Confidence			
Variable		Exp(B)	Interval		<i>p</i> value	Exp(B)	Interval		<i>p</i> value		
Recycle	Age					1.37	0.65	2.88	0.41		
	EID					1.22	0.72	2.07	0.46		
	Age x EID	1.00	0.996	1.000	*0.02	0.99	0.97	1.01	0.51		
*Statistically significant											
Reference category: Graduate											

Table 50. Multinomial Logistic Regression Models: Engineer Experts

						Model	Validation		
						Chi-Square (<i>p</i> value)	36.03 (<*0.01)		11.16 (0.08)
						Cox and Snell Pseudo R-Square	0.16		0.08
						Nagelkerke Pseudo R-Square	0.19		0.09

		Model Building				Validation			
		Odds Ratio	95% Confidence		<i>p</i> value	Odds Ratio	95% Confidence		<i>p</i> value
	Variable	Exp(<i>B</i>)	Interval			Exp(<i>B</i>)	Interval		
Recycle	Age	1.10	1.01	1.19	*0.03	1.01	0.90	1.12	0.93
	APFT	0.97	0.95	0.99	*0.01	0.97	0.95	1.00	*0.02
	RC3	0.99	0.95	1.03	0.64	1.04	0.99	1.09	0.15
Relief	Age	0.56	0.38	0.83	<*0.01	0.99	0.81	1.22	0.95
	APFT	0.96	0.91	1.00	*0.04	0.97	0.93	1.01	0.10
	RC3	1.13	1.03	1.23	*0.01	1.06	0.97	1.15	0.22

**Statistically significant*
Reference category: Graduate

Table 51. Multinomial Logistic Regression Models: Medical Experts

					Model	Validation			
					Chi-Square (<i>p</i> value)	10.36 (<*0.01)	6.32 (*0.04)		
					Cox and Snell Pseudo R-Square	0.07	0.07		
					Nagelkerke Pseudo R-Square	0.11	0.11		

					Model Building		Validation		
					Odds Ratio	95% Confidence	Odds Ratio	95% Confidence	
					Exp(<i>B</i>)	Interval	Exp(<i>B</i>)	Interval	<i>p</i> value
Other	Variable				RC1	1.08 1.00 1.17	1.03	0.91 1.16	0.66
					INTRr	0.93 0.87 0.99	0.91	0.84 0.99	*0.02

**Statistically significant*
Reference category: Graduate

Table 52. Multinomial Logistic Regression Models: Communications Experts

					Model		Validation	
					Chi-Square (<i>p</i> value)	4.94 (<i>p</i> = 0.09)	1.26 (0.53)	
					Cox and Snell Pseudo R-Square	0.02	0.01	
					Nagelkerke Pseudo R-Square	0.03	0.01	

					Model Building		Validation		
					Odds Ratio	95% Confidence			
					Exp(<i>B</i>)	Interval	Odds Ratio	95% Confidence	
							Exp(<i>B</i>)	Interval	<i>p</i> value
Relief	Variable								
	YRS ED	0.70	0.49	1.01	0.05*		0.99	0.74 1.33	0.97

**Statistically significant*
Reference category: Graduate

Table 53. Validation Model Classification Accuracy: Weapons Experts

	Observed		Predicted		% Correct
	N	Graduate	Recycle	Relief	
Graduate	110	108	2	0	98.20
Recycle	59	59	0	0	0.00
Relief	16	15	0	1	6.30
% Overall		59.30	100.00	0.00	58.90

Bold values reflect accurate predictions

Table 54. Validation Model Classification Accuracy: Engineer Experts

	Observed		Predicted		% Correct
	N	Graduate	Recycle	Relief	
Graduate	79	63	16	0	79.70
Recycle	53	34	19	0	35.80
Relief	11	5	6	0	0.00
% Overall		61.80	39.02	0.00	57.30

Bold values reflect accurate predictions

Table 55. Validation Model Classification Accuracy: Medical Experts

	Observed		Predicted		% Correct
	N	Graduate	Other		
Graduate	62	60	2		96.80
Other	21	21	0		0.00
% Overall		74.07	100.00		72.30

Bold values reflect accurate predictions

Table 56. Validation Model Classification Accuracy: Communications Experts

	Observed		Predicted		% Correct
	N	Graduate	Recycle	Relief	
Graduate	112	112	0	0	100.00
Recycle	32	32	0	0	0.00
Relief	16	16	0	0	0.00
% Overall		70.00	0.00	0.00	70.00

Bold values reflect accurate predictions

Table 57. CART Test Model: Weapons Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	110	101	9	0	91.82	8.18
Recycle	59	51	8	0	13.56	86.44
Relief	16	14	2	0	0.00	100.00
Total	185					
Average					35.13	
Overall % Correct					58.92	

Bold values reflect accurate predictions

Table 58. CART Test Model: Engineer Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	79	66	13	0	83.54	16.46
Recycle	53	41	12	0	22.64	77.36
Relief	11	9	2	0	0.00	100.00
Total	143					
Average					35.40	
Overall % Correct					54.55	

Bold values reflect accurate predictions

Table 59. CART Test Model: Medical Experts

	Observed			Predicted	
	Total Class	Graduate	Other	% Correct	% Misclassified
Graduate	62	61	1	98.39	1.61
Other	21	19	2	9.52	90.48
Total	83				
Average				53.96	
Overall % Correct				75.90	

Bold values reflect accurate predictions

Table 60. CART Test Model: Communications Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	112	111	1	0	99.11	0.89
Recycle	32	31	1	0	3.13	96.88
Relief	16	16	0	0	0.00	100.00
Total	160					
Average					34.08	
Overall % Correct					70.00	

Bold values reflect accurate predictions

Table 61. CART Modified Priors: Weapons Experts

	Equal Priors					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	110	18	56	36	16.36	83.64
Recycle	59	12	27	20	45.76	54.24
Relief	16	3	4	9	56.25	43.75
Total	185					
Average					39.46	
Overall % Correct					29.19	

Bold values reflect accurate predictions

	Equal Priors Target Class Reliefs					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	110	25	56	29	22.73	77.27
Recycle	59	18	27	14	45.76	54.24
Relief	16	4	4	8	50.00	50.00
Total	185					
Average					39.50	
Overall % Correct					32.43	

Bold values reflect accurate predictions

Table 62. CART Modified Priors: Engineer Experts

	Equal Priors					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	79	43	29	7	54.43	45.57
Recycle	53	14	31	8	58.49	41.51
Relief	11	4	3	4	36.36	63.64
Total	143					
Average					49.76	
Overall % Correct					54.55	

Bold values reflect accurate predictions

	Equal Priors Target Class Reliefs					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	79	28	21	30	35.44	64.56
Recycle	53	7	25	21	47.17	52.83
Relief	11	2	1	8	72.73	27.27
Total	143					
Average					51.78	
Overall % Correct					42.66	

Bold values reflect accurate predictions

Table 63. CART Modified Priors: Medical Experts

	Equal Priors				
	Observed		Predicted		
	Total Class	Graduate	Other	% Correct	% Misclassified
Graduate	62	19	43	30.65	69.35
Other	21	2	19	90.48	9.52
Total	83				
Average				60.56	
Overall % Correct				45.78	

Bold values reflect accurate predictions

	Equal Priors Target Class Others				
	Observed		Predicted		
	Total Class	Graduate	Other	% Correct	% Misclassified
Graduate	62	19	43	30.65	69.35
Other	21	2	19	90.48	9.52
Total	83				
Average				60.56	
Overall % Correct				45.78	

Bold values reflect accurate predictions

Table 64. CART Modified Priors: Communications Experts

	Equal Priors					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	112	39	35	38	34.82	65.18
Recycle	32	2	15	15	46.88	53.13
Relief	16	2	7	7	43.75	56.25
Total	160					
Average					41.82	
Overall % Correct					38.13	

Bold values reflect accurate predictions

	Equal Priors Target Class Reliefs					
	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	112	39	35	38	34.82	65.18
Recycle	32	2	15	15	46.88	53.13
Relief	16	2	7	7	43.75	56.25
Total	160					
Average					41.82	
Overall % Correct					38.13	

Bold values reflect accurate predictions

Table 65. Random Forests: Weapons Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	110	28	35	47	25.45	74.55
Recycle	59	19	11	29	18.64	81.36
Relief	16	4	3	9	56.25	43.75
Total	185					
Average					33.45	
Overall % Correct					25.95	

Bold values reflect accurate predictions

Table 66. Random Forests: Engineer Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	79	37	26	16	46.84	53.16
Recycle	53	17	20	16	37.74	62.26
Relief	11	1	6	4	36.36	63.64
Total	143					
Average					40.31	
Overall % Correct					42.66	

Bold values reflect accurate predictions

Table 67. Random Forests: Medical Experts

	Observed			Predicted	
	Total Class	Graduate	Other	% Correct	% Misclassified
Graduate	62	4	58	6.45	93.55
Other	21	0	21	100.00	0.00
Total	83				
Average				53.23	
Overall % Correct				30.12	

Bold values reflect accurate predictions

Table 68. Random Forests: Communications Experts

	Observed			Predicted		
	Total Class	Graduate	Recycle	Relief	% Correct	% Misclassified
Graduate	112	27	65	20	24.11	75.89
Recycle	32	6	17	9	53.13	46.88
Relief	16	5	9	2	12.50	87.50
Total	160					
Average					29.91	
Overall % Correct					28.75	

Bold values reflect accurate predictions

APPENDIX B: Figures



UNIFORMED SERVICES UNIVERSITY OF THE HEALTH SCIENCES

4301 JONES BRIDGE ROAD
BETHESDA, MARYLAND 20814-4799
www.usuhs.mil



January 26, 2017

MEMORANDUM FOR AMANDA RICKLING WEBB, M.A and M.S, P and R – UNIFORMED SERVICES UNIVERSITY of the HEALTH SCIENCES

SUBJECT: USUHS IRB #1 (FWA 00001628; DoD Assurance P60001) Approval of Protocol TO-72-8756 for Human Subjects Participation

Congratulations! The Initial Review for your human subjects research protocol entitled "Predicting Attrition in the MOS Phase of an Army Specialty Training Program", was reviewed and approved for execution on January 26, by Dr. Leodayan Bojanowski as an EXEMPT protocol under the provision of Category 4: Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded in a de-identified manner. This approval will be reported to the USUHS IRB #1 scheduled to meet on February 09, 2017.

This project will contribute to the growing body of research on accurate MOS classification by: (a) assessing the predictive validity of a test battery used by an Army Specialty training program, and (b) attempting to develop an MOS classification algorithm (i.e., decision tree) to improve classification assignment. This project extends previous research and goes beyond the identification of predictor variables; its ultimate goal is to identify an MOS assignment decision tree that leaders can use to improve classification accuracy.

You are required to submit amendments to this protocol, changes to the informed consent document (if applicable), adverse event reports, and other information pertinent to human research for this project. No changes to this protocol may be implemented prior to IRB approval. If you have questions regarding this IRB action or questions of a more general nature concerning human participation in research, please contact Christopher Murphy at 301-319-0444 or christopher.murphy.ctr@usuhs.edu at the IRB office.

Leodayan Bojanowski
Exemption Determination Official

Figure 1. Institutional Review Board (IRB) approval letter



DEPARTMENT OF THE ARMY
WOMACK ARMY MEDICAL CENTER
2817 REILLY ROAD
FORT BRAGG, NORTH CAROLINA 28310-7301

MCXC-DME-RES

9 March 2017

MEMORANDUM FOR Mr. Ed Deagle, Director of Research, 1st Special Warfare Training Group (A), U.S. Army John F Kennedy Special Warfare Center and School (USAJFKSWCS)

SUBJECT: USAJFKSWCS Support for Protocol, "Predicting Attrition in the MOS Phase of an Army Specialty Training Program," to be Conducted by CPT Amanda Webb, MS, Graduate Student, Department of Psychology, Uniformed Services University of the Health Sciences (USUHS), Bethesda, MD, WAMC Protocol Number 170202-1.

1. This doctoral research project is designed to address the need to improve MOS assignment by identifying variables that best predict successful graduation versus attrition in an Army specialty training program. Additionally, this project aims to develop a MOS assignment algorithm (i.e., a classification tree) that aids Army specialty unit leaders with classification decisions.
2. It appears that CPT Webb is a USAJFKSWCS asset. Because she is completing this project as part of a USUHS graduate program course requirement, USUHS is the institution engaged in research and has regulatory oversight. The USAJFKSWCS is supporting this research by providing de-identified data to CPT Webb for analysis. Womack Army Medical Center (WAMC) provides regulatory review for research to be supported by USAJFKSWCS.
3. A Command letter of support was provided and a Data Use Agreement is in place. Mr. Deagle is serving as an "honest broker" and providing a de-identified data set as described in the DUA and protocol.
4. The undersigned has determined that this research protocol does not involve human subjects as defined by 32 CFR 219.102 (f). Mr. Deagle is not an Investigator involved in the conduct of the research; he is simply providing a de-identified data set. His parent organization is a consumer of the results, and a DUA agreement is in place. The research may proceed as described with no further requirement for review at WAMC.
5. USUHS reviewed the protocol and determined that it is human subjects research exempt from IRB review IAW 32 CFR 219.101(b)(4). The Principal Investigator must adhere to the reporting requirements provided by USUHS in the approval letter.
6. Point of contact for this action is the undersigned at 910-907-6307.

Caryn L. Duchesneau, CIP
Human Protections Administrator

CF: CPT Amanda Webb, USUHS

Figure 2. Womack Army Medical Center IRB exemption letter

DATA USE AGREEMENT

20 Jan 2017

This Data Use Agreement forms a basis of understanding between The Department of Psychology, Uniformed School of Health Sciences (hereafter referred to as "University") and the U.S. Army John F. Kennedy Special Warfare Center and School (hereafter referred to as "Agency"), Fort Bragg, North Carolina, for intern supported research efforts.

Purpose of Agreement: This Data Use Agreement is designed to provide university students with an opportunity to conduct applied psychological research as part of their graduate studies. USAJFKSWCS will benefit through the completion of additional applied research that enhances the military selection and training programs.

Intern: CPT Amanda Webb

Study Title: Predicting SF.

Study Duration: September 2016 - December 2017

1. All personal identifying information (PII) will be removed from any datasets used in the project by USAJFKSWCS prior to analyses to ensure soldier confidentiality is fully maintained.
2. Use of USAJFKSWCS data is only authorized for the abovementioned study and intern. Additional research/analyses outside the current study and student will require new USAJFKSWCS approval.
3. All data remains the property of USAJFKSWCS.
4. Any future publication of the results from this study will be subject to prior command review and approval for release.
5. DoD regulations do not consider archived anonymous data research with human subjects so DoD Directive 3216.02 does not apply.

Approved By:

Neil E. Grunburg, Ph.D.
Professor, Military and Emergency Medicine
Uniformed Services University of the Health Sciences

Signature

24 Jan 2017

Date

Ed Deagle
Director of Research
1SWTG(A)
USAJFKSWCS

Signature

24 Jan 17

Date

Figure 3. Data sharing agreement

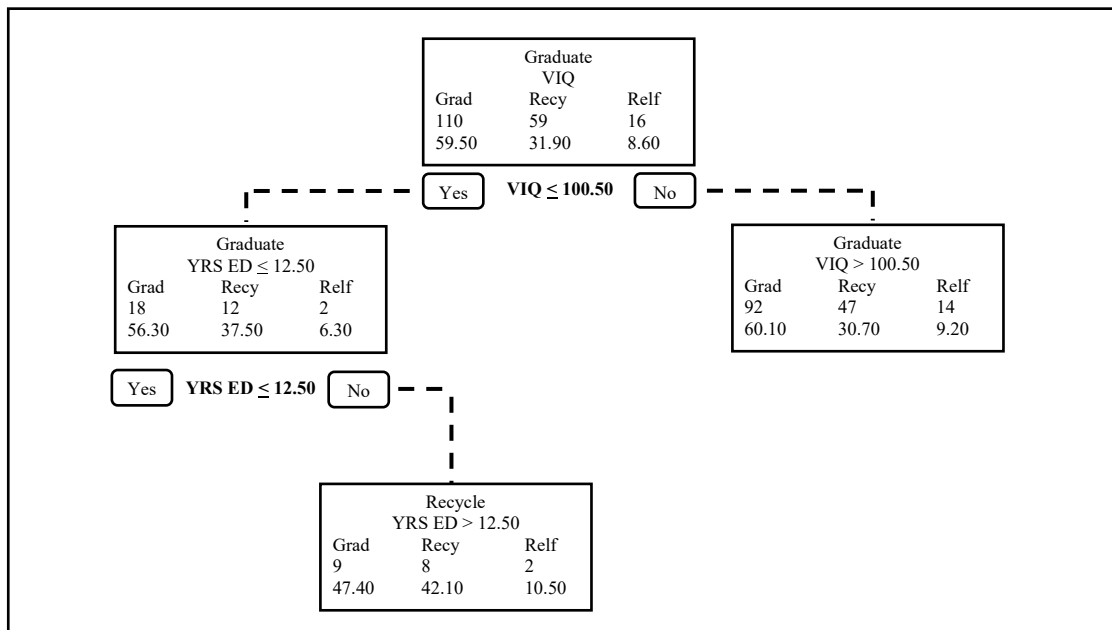


Figure 4. Classification and Regression Tree (CART): Weapons experts
 Grad: graduate; Recy: recycle; Relf: relief; First row = N within leaf;
 Second row = percentage within leaf

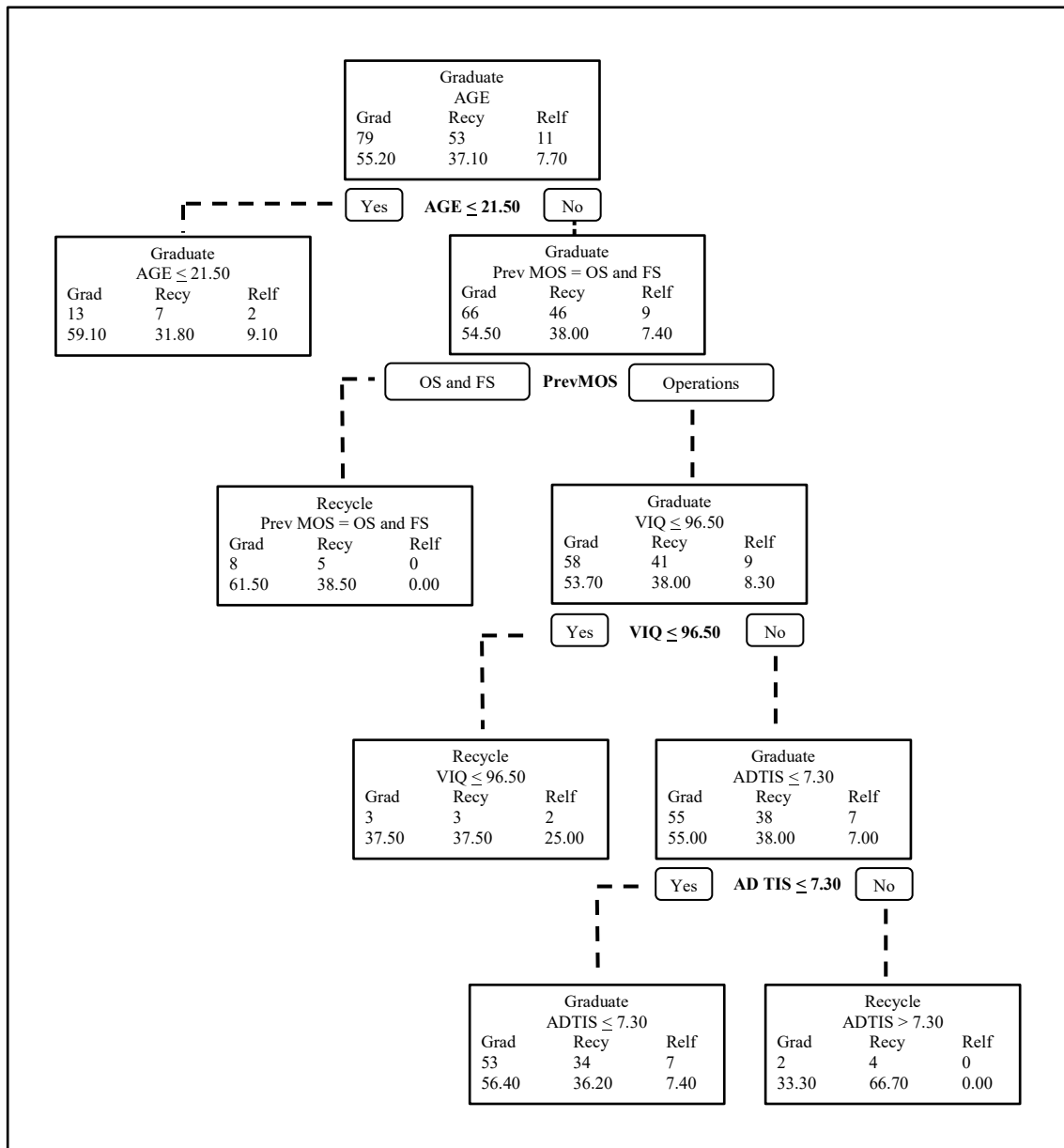


Figure 5. CART: Engineer expert

Grad: graduate; Recy: recycle; Relf: relief; PrevMOS = OS and FS: Previous MOS = Operational Support and Force Sustainment; First row = N within leaf; Second row = percentage within leaf

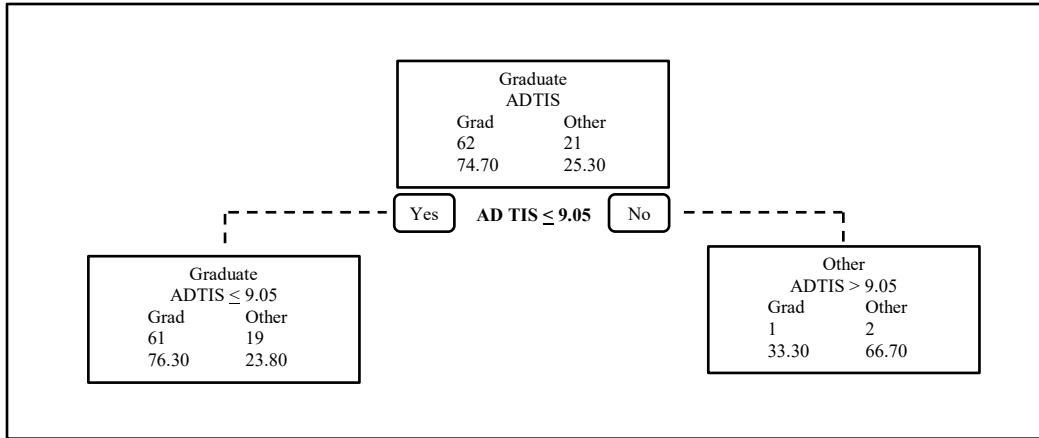


Figure 6. CART: Medical experts

Grad: graduate; Other: recycle, reclass, relief, or incomplete; First row = N within leaf; Second row = percentage within leaf

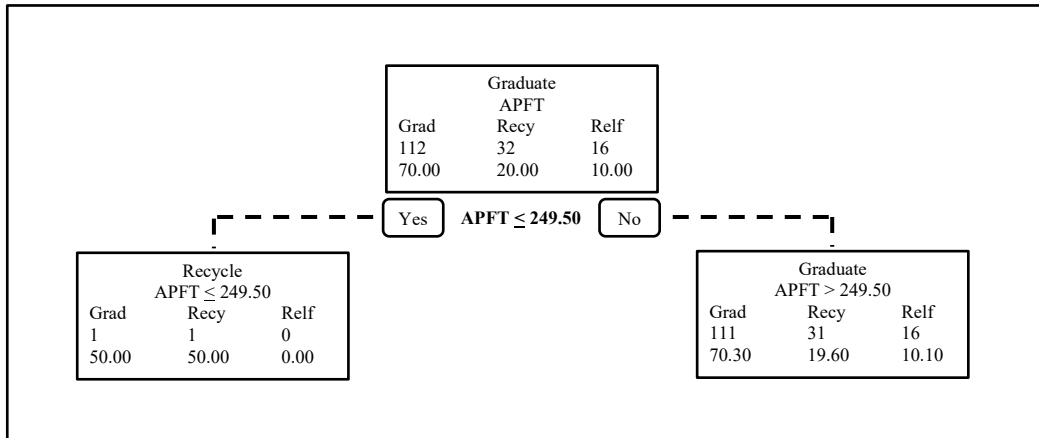


Figure 7. CART: Communications experts

Grad: graduate; Recy: recycle; Relf: relief; First row = N within leaf; Second row = percentage within leaf

APPENDIX C: Glossary

16PF: Sixteen Personality Factor Questionnaire

A&S: Assessment and Selection

AAS: Athletic Activities Score

ABLE: Assessment of Background and Life Experiences

ACB: Army Classification Battery

ACT: American College Testing

AD: Active Duty

ADTIS: Active Duty Time in Service

AFQT: Armed Forces Qualification Test

AGCT: Army General Classification Test

AGGR-r: Aggressiveness-Revised

AIM: Assessment of Individual Motivation

APFT: Army Physical Fitness Test

ARI: Army Research Institute

ARSOF: Army Special Operations Forces

ASVAB: Armed Services Vocational Aptitude Battery

BUD/S: Basic Underwater Demolitions/SEAL

BXD: Behavioral/Externalizing Dysfunction

CART: Classification and Regression Tree

CEER: College Entrance Exam Rank

CFA: Candidate Fitness Assessment

CIA: Central Intelligence Agency

CO: Combat

COI: Coordinator of Information

COG: Cognitive Database

DISC-r: Disconstraint-Revised

DoD: Department of Defense

EAS: Extracurricular Activities Score

EI: Emotional Intelligence

EID: Emotional/Internalizing Dysfunction

ENT GT: Entrance General Technical Score

FAS: Faculty Appraisal Score

FID: Foreign Internal Defense

FINAL: Final Database

FSIQ: Full Scale Intelligence Quotient

FY: Fiscal Year

GT: General Technical

Grad: Graduate

HSGPA: High School Grade Point Average

HSRCS: High School Rank Convert Score

INTRO-r: Introversion/Low Positive Emotionality-revised

IQ: Intelligence Quotient

IRB: Institutional Review Board

JPM: Joint-Service Job Performance Measurement

Land NAV: Land Navigation

MAB-II: Multidimensional Aptitude Battery, 2nd edition

MEPS: Military Entrance Processing Station

MMPI-2-RF: Minnesota Multiphasic Personality Inventory-2 Restructured Form

MOS: Military Occupational Specialty

MOSPref: Military Occupational Specialty Preferences

MSCEIT: Mayer-Salovey-Caruso Emotional Intelligence Test

NCO: Non-commissioned Officer

NEGE-r: Negative Emotionality/Neuroticism-Revised

NEO PI-R: NEO Personality Inventory - Revised

OASD-FM&P: Office of the Assistant Secretary of Defense for Force Management and Personnel

OCourse: Obstacle Course

ODA: Operational Detachment Alpha

OSS: Office of Strategic Services

OUTCOME: Outcome Database

PERF: Performance Database

PIQ: Performance Intelligence Quotient

PrevMOS: Previous Military Occupational Specialty

PSY-5: Personality Psychopathology Scales

PSYC-r: Psychoticism-Revised

RCd: Demoralization

RC1: Somatic Complaints

RC2: Low Positive Emotions

RC3: Cynicism

RC4: Antisocial Behavior

RC6: Ideas of Persecution

RC7: Dysfunctional Negative Emotions

RC8: Aberrant Experiences

RC9: Hypomanic Activation

Recy: Recycle

Relf: Relief

SAT: Scholastic Aptitude Test

SD: Standard Deviation

SES: Socioeconomic Status

SF: Special Forces

SFAS: Special Forces Assessment and Selection

SMA: Sergeant Major of the Army

SO: Sensor Operators

SOCOM: U.S. Special Operations Command

SOF: Special Operations Forces

SPM: Salford Predictive Modeler

ST: Skilled Technician

SWCS: Special Warfare Center and School

TABE: Test of Adult Basic Learning

TAPAS: Tailored Adapted Personality Assessment System

THD: Thought Dysfunction

USASOC: United States Army Special Operations Command

USSOCOM: United States Special Operations Command

VIF: Variable inflation factors

VIQ: Verbal Intelligence Quotient

WAIS-R: Wechsler Adult Intelligence Scores-Revised

WCS: Whole Candidate Score

WWI: World War I WWII: World War II

YRS ED: Years of Education

REFERENCES

1. 1997. Military attrition: Better screening of enlisted personnel could save millions of dollars. In *Subcommittee on Military Personnel, Committee on National Security, House of Representatives: United States General Accounting Office GAO/T-NSIAD-97-120*
2. 2014. *Congressional guide 2014: Admissions to the United States Service Academies*.
[http://www.usma.edu/admissions/shared%20documents/congressional%20guide%20\(2014\).pdf](http://www.usma.edu/admissions/shared%20documents/congressional%20guide%20(2014).pdf)
3. 2016. Statement of General Joseph L. Votel, U.S. Army, Commander United States Special Operations Command. In *House Armed Services Committee, Subcommittee on Emerging Threats and Capabilities*.
<http://docs.house.gov/meetings/AS/AS26/20160301/104552/HHRG-114-AS26-Wstate-VotelJ-20160301.pdf>
4. 2017. *Statement of General Raymond A. Thomas, III U.S. Army Commander United States Special Operations Command*. In *House Armed Services Committee, Subcommittee on Emerging Threats and Capabilities*.
<http://www.socom.mil/Pages/posture-statement-hasc.aspx>
5. Academy: United States Air Force. 2016. *Advice to Applicants*.
<http://www.academyadmissions.com/admissions/advice-to-applicants/all-applicants/>
6. Academy: United States Coast Guard. 2016. *Academy at a glance*. <http://www.uscga.edu/about2.aspx?id=36>
7. Academy: United States Merchant Marine. 2016. *Class profile*. <https://www.usmma.edu/class-profile>
8. Academy: United States Navy. 2016. *United States Naval Academy 2019 class portrait*.
http://www.usna.edu/Admissions/_files/documents/ClassPortrait.pdf
9. Admissions West Point. 2016. *West Point class profile*.
<http://www.usma.edu/admissions/sitepages/class%20profiles.aspx>
10. Army. 1990. Foot marches field manual no. 21-18. ed. Department of the Army: Headquarters Training and Doctrine Command
11. Army. 2005. Map reading and land navigation field manual 3-25.26. Washington, DC: Department of the Army

12. Army. 2010. Army physical readiness training circular no. TC 3-22.20. Washington, DC: United States Army Physical Fitness School
13. Armyenlist.com. 2016. *7 steps to join*.
<http://www.armyenlist.com/national-guard/7-steps-to-join/>
14. Atkinson RC, Geiser P. 2009. Reflections on a century of college admissions tests. *Educational Researcher* 38:665-76
15. Banks LM. 2006. The history of special operations psychological selection. In *Psychology in the service of national security*:83-95. Number of 83-95
16. Beal SA. 2010. The roles of perseverance, cognitive ability, and physical fitness in U.S. Army Special Forces assessment and selection *Rep. RR 1927*, Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA
17. Ben-Porath Y. 2008. An introduction to the MMPI-2-RF (Restructured Form). PsychCorp: Pearson
18. Box GEP, Tidwell PW. 1962. Transformation of the independent variables. *Technometrics* 4:531-50
19. Braun DE, Prusaczyk WK, Goforth Jr. HW, Pratt NC. 1994. Personality profiles of U.S. Navy Sea-Air-Land (SEAL) personnel, Naval Health Research Center, San Diego, CA
20. Bruan P, Wiegand D. 1991. The assessment of complex skills and of personality characteristics in military services. In *Handbook of military psychology*, ed. R Gal, D Mangelsdorff:37-61. New York: Wiley.
21. Busciglio HH, Teplitzky ML, Welborn C. 1991. Project A spatial tests and military orienteering performance in the special forces assessment and selection program, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA
22. Butcher JN, Jeffrey T, Cayton TG, Colligan S, DeVore JR, Minegawa R. 1990. A study of active duty military personnel with the MMPI-2. *Military Psychology* 2:47-61
23. Campbell RC, Knapp DC, Heffner TS. 2005. Planning for the future: Progress toward an NCO competency assessment program, U.S. Army Research Institute for the Behavioral and Social Sciences
24. Carless SA. 2000. The validity of scores on the Multidimensional Aptitude Battery. *Educational and Psychological Measurement* 60:592-603

25. Carlin TM, Sanders M. 1996. Soldier of the future: Assessment and selection of force XXI. *Special Warfare* 9:16-21
26. Caruso DR, Mayer JD, Salovey P. 2002. Relation of an ability measure of emotional intelligence to personality. *Journal of Personality Assessment* 79:306-20
27. Chappelle W, McDonald K, Thompson W, McMillan K, Marley M. 2010. Multiple Aptitude Battery-II normative intelligence test data that distinguish U.S. Air Force AC-130 Gunship sensor operators, Air Force Research Laboratory, Brooks City-Base, TX
28. Christian JR, Picano JJ, Roland RR, Williams TJ. 2010. Guiding principles for assessing and selecting high-risk operational personnel. In *Enhancing human performance in security operations: International and law enforcement perspectives*, ed. PT Bartone, BH Johnsen, J Eid, JM Violanti, JC Laberg:121-42. Springfield, IL: Charles C Thomas.
29. Davidson S. 2014. Army Special Operations Forces (ARSOF) human capital decision support tool. *Rep. AOJK-GPA*, Fort Bragg, NC
30. DRC, CTB. 2016. Discover TABE 9&10. DRC|CTB
31. Faul F, Erdfelder E, Buchner A, Lang A-G. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* 41:1149-60
32. Feeley SP. 1998. *Special Forces assessment and selection*. Naval Postgraduate School, Monterey, CA
33. Feickert A. 2014. U.S. Special Operations Forces (SOF): Background and issues for Congress. Congressional Research Service, Washington, DC
34. Field A. 2013. *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: SAGE Publications Ltd
35. Gade PA, Dudley NM. 2015. Sixty years of the U.S. Army selection and classification test development.
36. Geiser S, Santelices MV. 2007. Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. Berkeley: Center for Studies in Higher Education: University of California, Berkley
37. General Odierno RT. 2015. Leader development and talent management the Army competitive advantage. *Military Review* July/August:9-15

38. Graham SE. 1994. Analysis of special forces medic (18D) attrition, Army Research Institute, Alexandria, VA
39. Greene RL. 2011. *The MMPI-2/MMPI-2-RF: An interpretive manual*. Boston, MA: Allyn & Bacon
40. Hanser LM, Oguz M. 2015. United States service academy admissions: Selecting for success at the Military Academy/West Point and as an officer, RAND, Santa Monica, CA
41. Harrel FE, Lee KL, Mark DB. 1996. Multivariate prognostic models: Issues in developing models, evaluating assumptions and accuracy, and measuring and reducing errors. *Statistics in Medicine* 15:361-87
42. Hazlett GA, Sanders M. 1999. Cognitive and personality assessment in Special Forces assessment and selection. *Special Warfare* 12:14-20
43. Horgen KE, Nye CD, White LA, LaPort KA, Hoffman III RR, et al. 2013. Validation of the noncommissioned officer special assignment battery. *Rep. TR 1328*, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA
44. Hosmer DW, Lemeshow S, Sturdivant RX. 2013. *Applied logistic regression*. John Wiley & Sons
45. Howell DC. 2012. *Statistical methods for psychology*. Belmont, CA: Wadsworth, Cengage Learning
46. Hoyt WT. 2000. Rater bias in psychological research: When is it a problem and what can we do about it. *Psychological Methods* 5:64-86
47. Jackson DN. 2003. *Multidimensional Aptitude Battery - II manual*. Port Huron, MI: SIGMA Assessment Systems, Inc
48. Jackson DN. 2016. *MAB-II*. <http://www.sigmaassessmentsystems.com/assessments/multidimensional-aptitude-battery-ii/#heading-what-is-the-mab-ii>
49. Keselman HJ, Miller CW, Holland B. 2011. Many tests of significance: New methods for controlling type I errors. *Psychological Methods* 16:1-12
50. Kilcullen RN, Mael FA, Goodwin GF, Zazanis MM. 1999. Predicting U.S. Army Special Forces field performance. *Human Performance in Extreme Environments* 4:53-63
51. King MW, Resick PA. 2014. Data mining in psychological treatment research: A primer on classification and regression trees. *Journal of Counseling and Clinical Psychology* 82:895-905

52. Kitsantas P, Hollander M, Lei L. 2006. Using classification trees to assess low birth weight outcomes. *Artificial Intelligence in Medicine* 38:275-89
53. Kitsantas P, Moore TW, Sly DF. 2007. Using classification trees to profile adolescent smoking behaviors. *Addictive Behaviors* 32:9-23
54. Knapik JJ, Jones BH, Hauret K, Darakjy S, Eugene P. 2004. A review of the literature on attrition from the military services: Risk factors for attrition and strategies to reduce attrition, U.S. Army Center for Health Promotion and Preventive Medicine, Aberdeen Proving Ground, MD
55. Knapp DJ, Campbell RC. 2004. Army enlisted personnel competency assessment program phase I (volume I): Needs analysis. *Rep. Technical Report 1151*, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA
56. Knapp DJ, Campbell RC. 2006. Army enlisted personnel competency assessment program: Phase II report. *Rep. Technical Report 1174*, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA
57. Knapp DJ, Heffner T. 2009. Validating future force performance measures (Army Class): End of training longitudinal validation. *Rep. Technical Report 1257*, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA
58. Kubisiak UC, Lentz E, Horgen KE, Bryant RH, Connell PW, et al. 2009. Review of interventions for reducing enlisted attrition in the U.S. military: An update. *Rep. Technical Report #636*, PDRI, Tampa, FL
59. Lewis RJ. 2000. An introduction to classification and regression tree (CART) analysis. In *Annual Meeting of the Society for Academic Emergency Medicine*. San Francisco, CA
60. Lilley K. 2015. 7 ways enlisted leaders want to fix the force. *Army Times*, May 10
61. Lin M, Lucas HC, Shmueli G. 2013. Research commentary - Too big to fail: Large samples and the p-value problem. *Information Systems Research* 24:906-17
62. Maier MH. 1993. Military aptitude testing: The past fifty years, Defense Manpower Data Center, Monterey, CA
63. Malvesti ML. 2010. To serve the nation: U.S. Special Operations Forces in an era of persistent conflict. Center for a New American Security, Washington, DC

64. Mason BA, Gunersel AB, Ney EA. 2014. Cultural and ethnic bias in teacher ratings of behavior: A criterion-focused review. *Psychology in the Schools* 51:1017-30
65. Mayer JD, Salovey P, Caruso DR, Sitarenios G. 2001. Emotional intelligence as a standard intelligence. *Emotion* 1:232-42
66. McDonald DG, Norton JP, Hodgdon JA. 1988. Determinants and effects of training success in U.S. Navy Special Forces (NHRC Report #88-34), Naval Health Research Center, Washington, DC
67. Mellard DF, Anderson G. 2007. Challenges in assessing for postsecondary readiness. *Rep. National Commission on Adult Literacy (NJ1)*
68. Merkle EC, Shaffer VA. 2011. Binary recursive partitioning: Background, methods, and application to psychology. *British Journal of Mathematical and Statistical Psychology* 64:161-81
69. Monahan J, Keener JM. 2012. Fitness-for-duty evaluations. In *Military psychology: Clinical and operational applications, Second Edition*, ed. CH Kennedy, EA Zillmer. New York, NY: The Guilford Press
70. National Association for College and Admission Counseling. 2016. Use of predictive validity studies to inform admission practices, National Association for College and Admission Counseling (NACAC)
71. Nye CD, Beal SA, Drasgow F, Dressel JD, White LA, Stark S. 2014. Assessing the tailored adaptive personality assessment system for army special operations forces personnel. *Rep. RR 1971*, U.S. Army Research Institute for the Behavioral and Social Sciences, Arlington, VA
72. Nye CD, Drasgow F, Chernyshenko OS, Stark S. 2012. Assessing the tailored adaptive personality assessment system (TAPAS) as an MOS qualification instrument, U.S. Army Research Institute, Fort Belvoir, VA
73. Official Site of the ASVAB. *History of military testing*
http://official-asvab.com/history_coun.htm
74. Ogle AD, Barron LG, Fedotova AV. 2016. Job analysis of United States Air Force military training instructor duty: Identification of screening criteria for instructor candidate suitability. *Military Psychology* 28:50-63
75. OSS Assessment Staff. 1948. *Assessment of men: Selection of personnel for the Office of Strategic Services*. New York: Rinehart & Co

76. Pendleton J. 2015. Special operations forces: Opportunities exist to improve transparency of funding and assess potential to lessen some deployments. *Rep. GAO-15-571*
77. Phelps MC, Merkle EC. 2008. Classification and regression trees as alternatives to regression. In *Proceedings of the 4th Annual GRASP Symposium*. Wichita State University
78. Picano JJ, Roland RR. 2012. Assessing psychological suitability for high-risk military jobs. In *The Oxford handbook of military psychology*, ed. JH Laurence, MD Matthews. New York: Oxford University Press
79. Picano JJ, Williams TJ, Roland RR. 2012. Assessment and selection of high-risk operational personnel: Identifying essential psychological attributes. In *Military psychology: Clinical and operational applications*, ed. CH Kennedy, EA Zillmer:50-72. New York: Guilford Press
80. Pincus W. 2014. In planning for the future, the Pentagon sees fewer troops and more technology. In *The Washington Post*, March 3
81. Pleban RJ, Allentoff HL, Thompson TJ. 1989. Preliminary assessment of selected predictors of special forces qualification course success, U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA
82. Prusaczyk WK, Goldberg GM. 2002. Organizational, psychological, and training aspects of special operations forces. In *Medical aspects of harsh environments*, ed. KB Pandolf., RE Burr. Washington, DC: Office of the Surgeon General, Borden Institute
83. Robertson IT, Smith M. 2001. Personnel selection. *Journal of Occupational and Organizational Psychology* 74:441-72
84. Robinson L. 2013. The future of U.S. special operations forces. *Rep. Council Special Report No. 66*, Council on Foreign Relations, New York, NY
85. Rumsey MG. 2012. Military selection and classification in the United States. In *The Oxford handbook of military psychology*:129-47. New York: Oxford University Press
86. Rumsey MG, Arabian JM. 2014. Introduction to the special issue on selected new developments in military enlistment testing. *Military Psychology* 26:131-7
87. Rumsey MG, Arabian JM. 2014. Military enlistment selection and classification: Moving forward. *Military Psychology* 26:221-51
88. Russell TL, Crafts JL, Tagliareni FA, McCloy RA, Barkley P. 1996. Job analysis of Special Forces jobs, U.S. Army Research Institute for the

Behavioral and Social Sciences, Alexandria, VA

89. Salford Systems. 2017. *Random Forests*.
<https://www.salford-systems.com/products/randomforests>
90. Salford Systems. 2017. *Salford Systems data mining and predictive analytics software*. <https://www.salford-systems.com/resources>
91. SAS Enterprise Miner. 2016. *Getting started with SAS enterprise miner: Exploring input data and replacing missing values*.
https://www.youtube.com/watch?v=EBFtrPKQ0zg&list=PLVBcK_IpFVi-xzvJiOlf33UvVbRoLRu0z&index=2
92. Schmitt E, Mazzetti M, Shanker T. 2012. Admiral seeks freer hand in deployment of elite forces. In *The New York Times*, Feb 12
93. Sewell GF. 2011. How emotional intelligence can make a difference. *Military Review*, March-April 201
94. Spinath FM, Spinath B, Harlaar N, Plomin R. 2006. Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence* 34:363-74
95. Stark S, Chernyshenko OS, Drasgow F, Nye CD, White LA, et al. 2014. From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology* 26:153-64
96. Steadman HJ, Silver E, Monahan J, Appelbaum PS, Clark Robbins P, et al. 2000. A classification tree approach to the development of actuarial violence risk assessment tools. *Law and Human Behavior* 24:83-100
97. Strobl C, Malley J, Tutz G. 2009. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* 14:322-48
98. Task Group HFM-171. 2012. Psychological and physiological selection of military special operations forces personnel, North Atlantic Treaty Organization
99. United States Army Special Operations Command (USASOC). *Office of Strategic Services*. <http://www.soc.mil/OSS/index.html>
100. United States General Accounting Office. 1991. DOD Service Academies: Improved cost and performance monitoring needed. Ed. USGA Office. Washington, D.C.
101. USASOC. *Office of Strategic Services*. <http://www.soc.mil/OSS/index.html>

102. Velky JL. 1990. Special Forces assessment and selection. *Special Warfare* 3:12-5
103. Vergun D. 2015. *Dailey: Non-deployable soldiers no.1 problem*. https://www.army.mil/article/158897/Dailey_Non_deployable_Soldiers_No_1_problem/
104. Vergun D. 2015. Personality test helps ensure civilians are compatible for Army life. *Army News, May 15*. https://www.army.mil/article/148691/Personality_test_helps_ensure_civilians_are_compatible_for_Army_life
105. Vittinghoff E, Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. 2011. *Regression methods in biostatistics: Linear, logistic, survival, and repeated measures models*. New York, NY: Springer Science & Business Media
106. Welsh JR, Kucinkas SK, Curran LT. 1990. Armed services vocational battery (ASVAB): Integrative review of validity studies. *Rep. AFHRL-TR-90-22*, Operational Technologies Corporation, San Antonio, TX
107. Westrick PA, Le H, Robbins SB, Radunzel JMR, Schmidt FL. 2015. College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment* 20:23-45
108. White LA, Rumsey MG, Mullins HM, Nye CD, LaPort KA. 2014. Toward a new attrition screening paradigm: Latest Army advances. *Military Psychology* 26:138-52
109. Zazanis MM, Hazlett GA, Kilcullen RN, Sanders MG. 1999. Prescreening methods for Special Forces assessment and selection, Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA
110. Zazanis MM, Zaccaro SJ, Kilcullen RN. 2001. Identifying motivation and interpersonal performance using peer evaluations. *Military Psychology* 13:73-88