Predicting Attrition in a Military Special Program Training Command

by

Brendan J. Finton, M.S.

Doctoral Dissertation submitted to the faculty of the

Department of Medical and Clinical Psychology

Graduate Program of the Uniformed Services University

of the Health Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

2016

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
MEDICAL & CLINICAL PSYCHOLOGY DEPARTMENT

Title of Dissertation: "Predicting Attrition in a Military Special Program Training Command"

Name of Candidate: Brendan J. Finton
Doctor of Philosophy Degree
January 19, 2016

DISSERTATION AND ABSTRACT APPROVED:

DATE:

19 Jan 2016

Dr. Tracy Sbrocco
DEPARTMENT OF MEDICAL & CLINICAL PSYCHOLOGY
Committee Chairperson

19 Jan '16

Dr. Neil E. Grunberg
DEPARTMENT OF MEDICAL & CLINICAL PSYCHOLOGY
Dissertation Advisor

2016-01-19

Dr. Paul E. Rapp
DEPARTMENT OF MEDICAL & CLINICAL PSYCHOLOGY
Committee Member

1/19/2016

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

19 JAN 16

Dr. Carrie H. Kennedy
NAVAL BRANCH HEALTH CLINIC BAHRAIN
Committee Member

ACKNOWLEDGMENTS

I am indebted to more people than I could possibly recognize on this page and likely am unable to see all of the ways that others have contributed to my development. I am especially thankful for the support of my Doctoral Committee: Drs. Tracy Sbrocco, Paul Rapp, Cara Olsen, and Carrie Kennedy. They guided me with professionalism, support, and some good humor. I also am grateful to the faculty and staff of the Medical and Clinical Psychology program for their support and patience through the years.

I would not be here if not for the support of the U. S. Navy, and especially the the Navy Psychology community. I have found many mentors in this community that I know will be there as I continue to mature and grow professionally. I hope to make them proud and give back to the community in kind.

I cannot overstate the importance of the Grunberg lab team who helped in direct and indirect ways throughout my graduate school experience. Ms. Erin Barry was a friend and collaborator who has done more to support me than I am probably even aware. I cannott thank Matt Moosey and Ang Yarnell enough for their support and friendship.

I owe a special thanks to my friends and family. Your support and patience made this project and degree possible. I learn from you all every day, and I am deeply fortunate to have so many loved ones in my life.

Finally, I will forever be grateful to my major advisor and mentor, Dr. Grunberg. I am in awe of his commitment, in word and deed, to students, research, and this country. I know that every action he has taken throughout my graduate career has been with my best interests at heart. I am a better scholar, clinician, and person because of his commitment to my development.

COPYRIGHT STATEMENT

The author hereby certifies that the use of any copyrighted material in the dissertation manuscript entitled: "Predicting Attrition in a Military Special Program Training Command" is appropriately acknowledged and, beyond brief excerpts, is with the permission of the copyright owner.

Brendan J. Finton

May 20, 2016

DISCLAIMER

The views presented here are those of the author and are not to be construed as official or reflecting the views of the Uniformed Services University of the Health Sciences, the Department of Defense or the U.S. Government.

ABSTRACT

Predicting Attrition in a Military Special Program Training Command

Brendan J. Finton, M.S., 2016

Thesis directed by: Neil E. Grunberg, Professor, MPS, MEM

Screening for special assignments within the U.S. military is a top priority of the Department of Defense. Developments that increase the likelihood of selecting and training only the most highly qualified candidates for special assignments contribute to the national security and defense, and increase fiscal responsibility within the military. One method to improve applicant screening is to retrospectively analyze training performance and outcome data to determine the variables that predict success or failure (*i.e.*, attrition) in a training program. Some variables that may be predictive of attrition include: intelligence, physical fitness, age, rank, and relevant psychological constructs (*e.g.*, posttraumatic stress symptoms).

This research analyzed training performance and outcome data from an East Coast U.S. military training command to identify variables that predicted success and failure in this command. It was hypothesized that logistic regression and multiple regression would identify the relative contribution of demographic and psychological factors that enhance prediction of attrition above and beyond current assessment and selection methods. Latent profile analysis also was used to characterize subgroups within the sample population for whom training outcome risk was greater or lesser.

The most efficient model derived via logistic regression correctly identified 98.1% of the successful trainee outcomes but only correctly identified 17.8% of the training failures. Among program completers, the most efficient multiple regression models resulted in adjusted $R^2 = 0.297$ for program GPA and adjusted $R^2 = 0.270$ for instructor ratings. Latent profile analysis revealed a best-fitting, 7-subgroup solution in which the characterized subgroups differed in attrition risk ranging from a 46.3% pass rate to 89.5%. Characterizing trainees into these subgroups may be the most effective approach to predict training program outcomes. These findings resulted in actionable recommendations for the command from which the data were collected. The methods used to analyze the training data may serve as a template for future attrition evaluation for military training programs.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Background

PURPOSE

While the number of traditional ground forces in Iraq and Afghanistan has declined dramatically in the past years, the utilization of non-traditional military forces across the globe has increased. Special operations forces are increasingly called upon to handle sensitive missions (*e.g.*, hostage rescue), the Navy's submarine forces undergo demanding undersea patrols irrespective of the surface wars, and smaller, more mobile military forces are called upon for regional stability and security missions around the globe (*e.g.*, Africa). Given the sensitivity of the missions and unpredictability of the operating environments for these non-traditional forces, assessment and selection (A&S) for special military programs is a national priority.

The purpose of this study was to develop statistical models to predict candidate attrition by identifying key variables that predict training success or failure. Further, this project also sought to characterize subgroups within this special military training command to improve A&S for the command. Special military assignments often come with increased responsibility for the service member and/or the potential for increased national and international visibility in the media for actions taken while serving in these assignments. With these considerations in mind, as well as the potential increase in resource efficiency (*e.g.*, time, money), any developments from this research that increase the likelihood of selecting and training only the most highly qualified candidates contribute to the military mission.

Historically, military efforts at improving A&S have been limited by organizational and practical considerations (50). Some limitations that have been identified as impeding the sharing of A&S developments within military special programs include: the need to maintain assessment validity/security by minimizing outside knowledge of processes; classification considerations that prevent publication and disclosure of the methods and tools used in A&S; and scientists involved in development of and conducting A&S rarely have the time to analyze and publish the available data (50).

This document summarizes the history of A&S within the U.S. military, highlights the adverse impact of military training attrition, and describes some current A&S programs within the military, as well as identifies benefits of improved A&S. Subsequently, demographic, psychological, and military variables that likely contribute to attrition are explored. The final section of Chapter 1 introduces the rationale for this research project and ties these subjects together. Chapter 2 explicates the aims of the study, and Chapter 3 describes the methodology. Chapter 4 presents the findings of the analyses, and Chapter 5 discusses the implications and future directions of the project.

ASSESSMENT AND SELECTION (A&S)

Accurate military assessment and selection (A&S) is critical to identify the right candidates for a special assignment role (selecting in), as well as keeping the wrong candidates out of such specialized, high stress/visibility positions (selecting out) (58). Because the role requirements for special assignments vary considerably between commands and missions, selecting in is a more specialized process that should be tailored to the mission requirements (50). In contrast, there are many behavioral, motivational, and cognitive factors common to special assignments that indicate poor fit. Efforts aimed at identifying the individual risk factors as well as the constellation of variables that are associated with failure to select for a program may be applicable across a wide range of assignments and may serve as a foundation for research tailored to other assignments.

MILITARY TRAINING ATTRITION

By definition, a candidate who fails a training program before reaching the position for which he or she is training is a poor candidate. Failure rates for U.S. military basic training range from approximately 7-14% (25). Attrition prior to completion of the first enlistment across branches is approximately 30%, inclusive of basic training failure (40; 71). The most common reasons for attrition within the first three years of enlistment include performance problems, medical problems, and misconduct.

Early military efforts aimed at reducing attrition through pre-accession screening have been in place to some degree since World War I. These initial assessment and selection (A&S) screening efforts were primarily intelligence testing, and these efforts were at the forefront of practical application of psychological principles (53). Military accession screening measures, including the Army Alpha and Beta Tests of Intelligence, became the foundation of objective testing and applied psychology in the United States (8).

During World War II, efforts at screening for special military jobs included A&S for the Office of Strategic Services (OSS) (28) and screening for fitness for military service using projective tests (74). The use of a multiple choice Rorschach Test administered to groups was found to be a poor screening tool in these populations (34). Handler (28) reported that there was a shift from purely paper-and-pencil testing to a more comprehensive approach during World War II. This new approach sought to select out individuals with low intelligence and those individuals lacking necessary skills, and to select in the highly qualified candidates who might serve in challenging, dynamic environments. The OSS A&S program included traditional psychological tests as well as staff observations, problem-solving scenarios, and intense interviews (28). The A&S staff developed detailed personality sketches of the candidates based on the comprehensive data before evaluating them for the position to which they would likely be assigned, emphasizing that the comprehensive evaluation was better for A&S than measures in isolation.

A milestone in pre-accession screening for the U.S. military occurred in the late 1950s when a statistically significant inverse relationship between level of education and attrition was found, with high school graduates having a 3.15% discharge rate compared to a 17.99% discharge rate for people with less than 12 years of education (23). The U.S. Army utilized this research and additional replications of it to develop a tiered system of risk based on education credentials for accession purposes, with 75% or greater of Tier 1 (*i.e.*, high priority) accessions completing their first three years of service compared with 62% of Tier 2 (*i.e.*, medium priority) accessions (71). Although this system is statistically valid to reduce attrition risk, there have been criticisms regarding the potential discriminating effect of limiting accession based on education. As such, the U.S. Army has continued to refine and expand the screening tools, eventually leading to newer tools for recruit A&S such as the Tailored Adaptive Personality Assessment System, or TAPAS (71). The TAPAS has been utilized as an adjunct screening tool for Army recruiting to screen in and screen out candidates based on personality characteristics and the relative match to Army expectations. The TAPAS is not administered to every potential recruit; rather, it is used for individuals who perform marginally on the aptitude screening (*i.e.*, Armed Services Vocational Aptitude Battery).

An additional component to reduce attrition at the basic training level is to implement programs that enhance retention after selection to entry training. Kubisiak and colleagues (41) conducted a review of the attrition-reduction interventions utilized across and between the different branches of service. Retention-enhancing interventions have been enacted at the level of the recruit such as fitness programs, additional academic training, mental health and social support strategies, financial incentives, and direct counseling. Other initiatives have emphasized changes in administrative and leadership approaches to increase recruit perception of their value and to decrease the ease with which they can depart training (*e.g.*, appropriating the recruits' civilian clothing on entry), to decrease unoccupied time, and reduce training duration (41).

Despite the promise of reduced attrition based on these training interventions, Kubisiak and colleagues (41) caution that attrition can occur at multiple stages and that attrition, in and of itself, is not inherently good or bad for the military. Decreasing attrition at the entry/testing phase (*i.e.*, prior to accession) may increase attrition subsequent to accession (*e.g.*, during basic training). Intervening to increase retention in basic training may result in increased attrition within the first enlistment contract by keeping in recruits who are not suited for military service. Ideally, A&S would identify the best potential recruits to attend basic training, and retention programs could be directed at these candidates with greater potential. However, the importance of A&S within special training programs, or programs for which A&S beyond entry enlistment is required, is more pronounced because of the challenging environments in which the successful trainees operate with limited resources and support.

Special Training Programs Attrition, Assessment, and Selection

Attrition from a special training program is not synonymous with attrition from basic training or attrition from the military itself. The attrition in special program training ranges from 25% to more than 80%, due in large part to the challenging nature of the training (29; 65; 66). These numbers also indicate that a sizeable contingent of inappropriate candidates have been accepted to the training program. Statistical methods for identifying key variables that predict attrition from special training programs, such as logistic regression used in this study, may decrease the attrition by pre-screening potential trainees prior to their acceptance into the training program.

Different approaches, as discussed below, have been utilized to conduct assessment and selection (A&S) within different special military communities. A non-exhaustive but illustrative list of such programs includes aviation, undersea warfare, special operations forces, and astronaut assignments.

Military Aviation

The direct financial benefits of accurate A&S prior to entering flight school cannot be overstated. Basic flight training for military aviators costs more than $1 million per person, with training costs exceeding $9 million for a fully trained, operationally ready pilot (24).

Historically, A&S for aviation candidates has been a part of military aviation since World War I. This unique A&S has developed over time to meet the needs of the U.S. military as aviation technology has advanced (*e.g.*, the development of Unmanned Aerial Systems) as well as advances in testing and test administration (18). A detailed review on the history of assessment and selection of military aviators and astronauts can be found in Kennedy and Kay's *Aeromedical Psychology* (39). Because the role of aviation within each military branch is unique (*e.g.*, different aircraft, different combat and support roles), the U.S. Army, Air Force, and Navy utilize distinct A&S programs to select aviators, though these A&S programs have some functional overlap in content assessed.

The U.S. Army utilizes the Flight Aptitude Selection Test (FAST) to select applicants who are likely to succeed in Army flight training. The FAST is a 200-question, multiple choice test that is purported to assess motivation, coordination, leadership, and physical health (19). The Air Force utilizes the Pilot Candidate Selection Method (PCSM) for aviation A&S. The PCSM is derived using an algorithm based on knowledge, aptitude, and previous flight experience (67). The U.S. Navy, Marine Corps, and Coast Guard utilize the Naval Aviation Selection Test Battery (ASTB) for A&S of pilot and flight officer candidates. The ASTB is comprised of cognitive, personality, life experience, and psychomotor evaluations that have been found to be relevant to success in Naval aviation training (18; 49).

Cox and colleagues (18) note that any changes to the A&S programs within or between branches requires significant time, effort, and planning. The program evaluators have access to the outcome data (pass/fail; reasons for failure, etc.) as well as the initial assessment data. Ongoing evaluation and revision of these programs is warranted to refine and improve prediction of training success and to adapt to changing needs within military aviation. One approach to enhance A&S in these program could be to implement advanced statistical analyses, such as latent profile analysis, to identify subgroups within the accepted aviation candidates. For aviation, these subgroups may identify candidates who are likely to attrite or who might be better suited to different aviation platforms.

Astronauts

A&S for astronaut roles deals with unique mission requirements and stressors of serving in space. Astronauts are selected from civilian (*e.g.*, researchers) and military (*e.g.*, pilots) candidates, and all potential astronauts must undergo an extensive selection process. The Behavioral Health and Performance Group (BHP) within the National Aeronautics and Space Administration (NASA) is responsible for psychiatric evaluations to determine if a candidate meets the medical qualifications, and the BHP makes recommendations on the psychological suitability relevant to space missions (18). This combination of behavioral health data along with training program performance data yields the best predictions for success in the program. However, unique statistical models derived from this training data must be tailored to each special community. As such, the principles on which astronaut selections are founded parallel the rationale for the proposed project. Specifically, that A&S of candidates should be based on job-related tasks and tests as well as psychologically-grounded assessment measures and participant demographics.

Although current A&S efforts have been successful at selecting out psychopathology, there remain challenges in selecting in personnel who will maintain optimal functioning in the stressful operating environment of space (17). Recent efforts have worked to address these concerns by supplementing the psychological and personality testing and interviews with behavioral and functional exercises to better train and evaluate the candidates' abilities to work as a team (18). This effort at prediction aims not only to select out, but to select in the best candidates. In line with these goals, this project was designed to identify candidates likely to attrite from the training, and to identify candidates who are more successful in training.

Submariners

The environmental and operational considerations for personnel assigned to submarines provide unique A&S challenges. Whanger and colleagues (69) highlight the unique challenges for submariners as: (1) a small, enclosed environment; and (2) sustained, social group isolation. Submariners live, work, and eat in a restricted environment with approximately 100 - 150 other service members in which they may remain underwater for upwards of one month.

Given the importance of identifying resilient and non-pathological candidates for these roles, A&S efforts have focused on selecting in and selecting out candidates at multiple stages. Whanger and colleagues (69) note that potential submariners are screened in and out (offered contracts) of the field based on their Armed Services Vocational Aptitude Battery (ASVAB) scores. Candidates then are administered the SUBSCREEN assessment that flags sailors for a subsequent clinical interview at which they may or may not be retained (69).

The SUBSCREEN is a revision of previous screening assessments designed to reduce attrition by identifying at risk individuals at the outset of training (41). The SubMarine Attrition Test (SMART), an additional screening metric, was developed utilizing logistic regression from the SUBSCREEN to better identify attrition risk for legal problems and non-judicial punishment (5). This utilization of logistic regression to predict adverse outcomes further highlights the importance of developing and applying statistical models to all A&S efforts within military special populations and serves as justification for the application of logistic regression to the dataset in the present research.

Special Operations Forces (SOF)

A&S procedures for special operations forces (SOF), such as the U.S. Navy SEa, Air, Land (SEAL) Teams, Army Special Forces (SF), Marine Corps Critical Skills Operators, and Air Force Pararescue, are necessarily well-guarded and even classified procedures (50). Despite this limitation, the importance of A&S for SOF is unquestioned because of the challenging nature of their missions, and some limited attrition research is available. All military A&S programs may benefit from increased dissemination of relevant research, as this project was designed to do.

Morgan and colleagues (46) provide some foundation for variables that are relevant to A&S in military programs when they found that dissociative symptoms (*e.g.*, changes in temporal-spatial awareness, out-of-body experiences, “spacing out”) assessed at baseline during U.S. Army Special Forces training was predictive of training failure, such that higher symptom endorsement increased the likelihood of training failure.

Banks (3) conveys that psychologists and psychological testing, such as the Minnesota Multiphasic Personality Inventory - 2 (MMPI-2), are used to assess for personality and psychological conditions that are predictive of poor performance in training for Army SF. Of particular note is that A&S psychologists are frequently integrated into the operational command, and this integration allows the psychologist(s) to make recommendations to the command based on a combination of assessments and interviews (3). However, this integration of psychological resources with training performance can be optimized to predict attrition through the development and application of statistical models that integrate this information.

Benefits of Improved Assessment and Selection

The purpose of the assessment and selection (A&S) efforts outlined above and others like them across the Department of Defense (DoD) is to reduce attrition. This reduction in attrition is sought because of the potential cost and manpower savings for the DoD. There are budgetary and mission constraints associated with bringing service members to training programs. For example, the cost of a single recruit at basic training ranged from $13,684-$20,473 (adjusted for inflation), and these costs are not recouped when a trainee attrites (52). Reducing failure rates has the potential to improve the cost-effectiveness of military training programs and to reduce administrative/training burden on the DoD more broadly (38). Even small improvements in predictive validity of screening procedures could save measureable financial, time, and manpower resources across the DoD. Within special training programs, smaller and more focused class sizes would improve the instructor-to-student ratio, which could potentially improve the quality of the training for the candidates who are statistically more likely to succeed.

Another important but often overlooked benefit of improved A&S is the benefit for the service members. Failing to qualify for or complete a special training program can result in career disruption. Service members may turn down other potential military career opportunities or they may move themselves or their families out of other successful billets with no gain for their career. Other service members may lose their military career track and wind up in undesired career designations or even separated from service. Improved A&S has even been called a “moral obligation” (p. 66) to service members (69). Whanger and colleagues (69) convey that keeping psychologically vulnerable service members out of potentially overwhelming training and operational environments protects their wellbeing as well as the safety of their fellow service members.

ATTRITION RISK FACTORS

There are several variables that have been studied as risk factors and predictors of adverse outcomes that may affect performance in military training settings. Utilizing the existing literature on demographic, psychological, and military/professional variables and adverse outcomes can provide the empirical foundation to integrate these variables in predictive statistical models to improve assessment and selection in a special military training command.

However, it should be noted that the majority of the literature outlined below provides information on variables primarily in isolation ("stovepiped"). There is little published research demonstrating the predictive impact of multiple variables considered simultaneously on attrition risk. The following variables, divided into demographic, psychological, and military/professional categories, comprise the areas assessed as part of the special military training program. It is from these areas that this project was derived with the intent of utilizing multiple variables simultaneously to increase attrition prediction above and beyond the ability of any single variable.

Demographic Variables

Demographic variables include age, gender, intelligence, and history of civilian legal infractions. Age has been found to be a risk factor for mental health diagnoses in active duty populations with deployments to either OEF or OIF, such that younger veterans had significantly increased risk for PTSD and alcohol use disorders (55).

Gender is associated with differential increases in risk for a number of concerning conditions, such as suicide risk with males at approximately four times greater risk for completed suicide (15), and depression and PTSD with females at greater risk following combat exposure (43). Intelligence, as measured on standardized tests such as the ASVAB, has been validated to predict school and training performance, attrition, and performance in military job performance, with higher ASVAB scores predicting better performance and less attrition (12; 68).

A history of legal violations, military and civilian, has been associated with higher attrition risk (40; 51). Individuals receiving moral waivers for most civilian legal infractions are more likely to attrite early from the military than those without waivers (21). Further, individuals with multiple, relatively minor infractions may be at greater risk for attrition (44). Service members with Uniformed Code of Military Justice (UCMJ) violations are more likely to be discharged from the military early (35; 51). A recent study of Army recruits found that the attrition risk was most clear by the end of the first enlistment (20). It is unknown what impact a history of civilian and military legal history will have on A&S for a special military program.

Military Variables

Military variables that may influence success in training for a special assignment include physical fitness, combat deployments, military occupational specialty, and rank. Taylor and colleagues (62) found that physical fitness was inversely related to trait anxiety, and that there may be a relationship in which physical fitness mediates the relationship between trait anxiety and military stress symptoms.

Deployment status to combat zones has a unique relationship to mental health and psychiatric diagnoses. Larson and colleagues (42) examined the incidence rates of psychiatric conditions in deployed and control groups, and they determined that the only condition positively associated with deployment was combat-related PTSD. In fact, the authors propose that psychologically unfit personnel are screened out before deployment, resulting in a psychologically healthier population deploying than those individuals who were not deployed.

Further, the majority of service members who deploy are resilient and do not develop PTSD (9). There does not appear to be a significant difference between service members who deploy a single time versus those who deploy multiple times in relation to remaining resilient (~83 vs. 85%, respectively) and those who worsen, approximately 6.7 vs 4.5%, respectively (9).

Military occupational specialty (MOS) likely plays an important role in attrition from or success in special training programs. Research has indicated that service members are better able to commit to and accomplish tasks when they have appropriate and relevant training for that task (10; 11). Presumably, service members with more training from a MOS relevant to the special program mission will be more successful than those with less relevant MOS training.

Finally, rank has been shown to be inversely associated with risk-taking behaviors such as alcohol abuse and dangerous behavior in motor vehicles (73; 75). Increasing rank is correlated with increasingly responsible behavior, though the direction of this relationship is not necessarily causal and is likely influenced by age.

Psychological Variables

Psychological variables that may predict performance outcomes include risky personal behaviors (*e.g.*, alcohol use), traumatic stress-related symptoms, and mental and emotional functioning based in personality factors, all of which are inversely related to positive outcomes. Alcohol use and financial responsibility patterns are predictive of occupational and safety concerns which increase the likelihood of attrition (73). Younger and lower ranking service members are more likely to engage in high-risk behaviors, and these behaviors have been found to be linked, such as drinking, excessive vehicular speeding, and low seatbelt use (4). Assessing drinking behavior can serve as a proxy for many of these other risky behaviors, given their frequent co-presence, again suggesting increased risk for attrition from special training programs in which mature behavior is requisite for success.

Although deploying service members are generally psychologically healthy, traumatic stress and mental health concerns following deployment significantly increase the risk of attrition from the military compared with those who do not report mental health concerns (30). The presence of a mental health diagnosis and psychiatric treatment and hospitalizations are associated with discharge and attrition from military service (30-32).

Exposure to childhood violence, such as childhood physical abuse, sexual assault, and domestic violence, increases risk for attrition during first-term enlistment (45). Attrition risk increases with each additional type of childhood experience of family violence. Additionally, childhood exposure to physical and sexual abuse resulted in elevated PTSD symptoms (56).

Finally, personality factors are proposed to contribute to success or failure in training settings. Feeley (22) identified the Sixteen Personality Factors Questionnaire (16

PF), a personality assessment, as the most congruent pen-and-paper assessment of desired attributes for U.S. Army Special Forces Assessment and Selection; further, the 16PF measured more than one third of the desired A&S attributes to a high degree. Stowers and Thompson (57) found that sub-scores and total composite score dimensions from the a version of the 16PF are significantly correlated with the final pass/fail recommendation for candidates assessed. Further, those dimensions as well as integrity/control were positively correlated with clinical interview scores for these candidates. The authors emphasize that utilizing clinical judgment in combination with the 16PF will likely yield the most accurate A&S when evaluating personnel for special assignments (57). This combination of subjective evaluations with objective psychological measures parallels the aims of this proposed project in that instructor ratings (subjective) and psychological measures (*e.g.*, 16-PF) were used.

In summary, current research suggests that a variety of demographic, military/occupational, and psychological variables are predictive of attrition when examined independently. It is likely that these demographic, military, psychological risk factors, and performance predictors interact in complex ways to determine individual training outcomes and attrition (*i.e.*, success or failure). Analyzing these variables in combination using statistical models will likely provide a useful tool to predict training outcomes of future trainees better than analyzing the variables in isolation. Further, there may be subgroups within the overall training population whose outcomes are similar because of similar characteristics. These subgroups are not apparent based on individual variables, but they may be identifiable using latent profile analysis. These identified subgroups can be compared against each other for training attrition to determine relative risk. This comparison of relative risk for attrition may be the basis for improved A&S for at-risk individuals in training programs.

STUDY RATIONALE: WHY PREDICT TRAINING ATTRITION?

Current military special training programs have high attrition rates among candidates. Statistical analyses of existing training databases may provide empirical support for current screening approaches and shed light on the qualities that make applicants particularly well- or ill-suited for a given training program. The U.S. military stands to save significant financial and manpower resources through increases in screening accuracy.

Special attention must be paid to the manner in which improvements in assessing and selecting (A&S) candidates are conducted. At one extreme, a training program could include a complete mock-up of the work environment and train and test the applicants in a one-to-one simulation of the eventual job setting. However, A&S gains must be balanced and optimized to the needs of the organization. If an A&S program is more costly (*e.g.*, time, resources) than the cost of current training attrition, then the increases in selection accuracy are not worth the organizational burden.

This study used retrospective statistical analyses of previously-collected training performance and outcome data. It did not require any increase in organizational burden to evaluate the ongoing assessment and selection program for the command from which the data were gathered. Rather, the use of existing data likely provides a valuable perspective about ways to enhance current procedures.

CHAPTER 2: Specific Aims and Hypotheses

This study had three specific aims. The first aim was to identify variables that optimally predict training attrition in a military special program training setting. The second aim was to identify variables that optimally predict training performance among special program training completers. The third aim was to identify, characterize, and validate subgroups within the sample population for whom different combinations of variables predict subgroup training attrition risk.

SPECIFIC AIM 1

To identify variables that optimally predict training attrition in a military special program training setting.

Hypothesis 1

The combination of demographic, military, and psychological variables included in the dataset will predict training success versus failure statistically better than chance.

Specific Aim 1 Rationale

This aim addresses the primary question of predicting success or failure in the training program. The variables included in the analysis were chosen based on practical and empirical considerations. From a practical consideration, there was a limited number of variables from which to select the predictor variables for use in the equation because the assessment and selection (A&S) program has been running for multiple years. No changes could be made to the previously-collected database. From an empirical standpoint, all variables selected for inclusion in the analysis have supporting literature suggesting that they impact attrition and/or assessment and selection. The rationale behind this aim was to determine which combination of these reasonable variables has the optimal predictive capability for training outcome in this program.

Logistic regression was the ideal analysis to answer this research question because it was a direct probability model that analyzes the mathematical relationship between predictor variables and a binary outcome variable; in this case, success or failure in the training program. The logistic regression also produced regression coefficients for the predictor variables that could be interpreted using the odds ratio, or the change in likelihood of being classified as a successful training outcome. From these regression coefficients, the likelihood of success or failure of new students in the training program was estimated, and these estimates could be used to inform future A&S decisions at this command.

SPECIFIC AIM 2

To identify variables that optimally predict relative performance in a military special program training setting among successful completers of the training program.

Hypothesis 2a

Some combination of demographic, military, and psychological variables will predict relative training performance in successful candidates as measured by their training program grade point average (GPA).

Hypothesis 2b

Some combination of demographic, military, and psychological variables will predict relative training performance in successful candidates as measured by their overall performance ratings from program instructors.

Specific Aim 2 Rationale

This aim addresses a secondary question of predicting relative performance among the successful completers of the training program. Identifying the variables that predict relative successful performance may improve the assessment and selection (A&S) of the command by improving its ability to "Select In" the students who are most likely to be highly successful in addition to completing the training. As with Specific Aim 1, the variables included in the analysis were chosen based on practical and empirical considerations. The rationale for this aim is to determine which of these variables have the greatest impact on objective and subjective training performance in this program.

Because the outcome variables (*i.e.*, program GPA, instructor rating) for these hypotheses are continuous, multiple regression was used to evaluate the association between the predictor variables and outcome variables. This aim was used to predict the degree of performance for candidates who completed the program. These outcomes (*i.e.*, GPA, and instructor ratings) may be predicted by the same or different variables from training program success (*i.e.*, Specific Aim 1). Including objective (GPA) and subjective (instructor ratings) evaluations of success as outcome variables provides a well-rounded assessment of performance that may lead to identification of a more comprehensive range of predictor variables to improve A&S.

SPECIFIC AIM 3

To identify, characterize, and validate subgroups within the sample population for whom different combinations of variables more accurately predict training attrition risk.

Hypothesis 3

The significant predictor variables identified in Specific Aim 1, and potentially 2, were included as indicators in the Latent Profile Analyses. Based on these indicator variables, it was hypothesized that three subgroups within the sample would be identified.

The attrition risk of these subgroups was hypothesized to include: low attrition risk, moderate attrition risk, and high attrition risk.

Specific Aim 3 Rationale

This aim utilized Latent Profile Analysis to characterize the sample into latent subgroups. These subgroups were maximally heterogeneous between groups and maximally homogenous within group. An examination of these groups based on their similar demographic and psychological characteristics provided valuable information for characterizing these groups based on differential attrition risk within the training setting.

CHAPTER 3: Methods

OVERVIEW

This study was a retrospective analysis of training performance and outcome data from an East Coast U.S. military training command. The data were collected from active duty, enlisted, U.S. service members in training for a special military assignment. The data were collected during the training and evaluation period prior to being assigned to the special program. The sample size for this study was N = 4241.

RESEARCH DESIGN AND PROCEDURES

Database

The research was conducted on data that were:

(1) existing/already collected; and

(2) de-identified.

Given these factors, the analyses in this project were determined by the Uniformed Services University of the Health Sciences Office of Research and Human Research Protections Program Office to be exempt from Institutional Review Board (IRB) review (Appendix A). The database contained training and demographic data for 4241 Service Members who were assigned to an East Coast U.S. military training command. The database used for the analyses included the following potential predictor variables:

<u>Personal</u>: age; sex; history of abuse; history of mental health treatment; first age of alcohol use.

<u>Military</u>: rank; military occupational specialty; Armed Services Vocational Aptitude Battery (ASVAB) General Technical (GT) Score; number of disciplinary counselings at training command; assessment of physical fitness;

total combat deployments.

Psychological: Alcohol Use Disorder Identification Test (AUDIT) score; Posttraumatic Checklist – Military (PCL-M) score; 16 Personality Factor Questionnaire (16 PF).

The database used for the analyses included the following outcome variables:

(1) Overall Training Outcome (pass/fail);

(2) Training Program GPA;

(3) Instructor Ratings of overall performance.

Participants

The participants were 4241 enlisted U.S. service members between the ages of 18-45 who were assigned to an East Coast U.S. military training command within the past decade. Written approval to access and analyze these data for the purposes of this doctoral research was obtained from the appropriate military chain of command. This documentation, along with additional information about the participant pool, has been provided to the Doctoral Dissertation Committee as part of the approval and defense proceedings; however, these approvals were not included in this document because of operational security considerations.

The military training program from which these data were collected lasts approximately eight weeks and includes classroom/academic instruction and practical training exercises. This program is open to enlisted personnel, and there are separate tracks for junior enlisted (E-3 to E-5) and senior enlisted personnel (E-6 and above). However, this project focused on students in the junior enlisted track. There was an average male to female ratio of 9:1. Class sizes ranged from 50-200 students per iteration across the data collection period. Prospective students were screened prior to arrival at the program, and this screening included information covering military service performance and medical readiness.

Procedures

All enrollees in the training program were administered psychological assessments, and demographic data were collected upon entry into the program. Participants were tracked throughout their time in the training program, and their eventual overall training outcome (*i.e.*, pass versus fail) within the program was linked to their assessment and demographic data. All participants included in the database had the baseline data and the primary outcome data (*i.e.*, pass versus fail). The participants who failed to complete the program did not have final GPA and instructor performance ratings.

Assessment Measures

1. The *Demographics Questionnaire* includes: rank; military occupational specialty (MOS); ASVAB General Technical (GT) Score; number of disciplinary counselings at training command; physical fitness (PT) assessment; and total combat deployments. The questionnaire also assesses non-military related personal characteristics, including: age; sex; history of abuse; history of mental health treatment; and first age of alcohol use.
2. The *Alcohol Use Disorders Identification Test* (2) is a 10-item test developed to "identify persons with hazardous and harmful patterns of alcohol consumption" (p. 2). The AUDIT demonstrates sensitivity (0.9 and above) as well as specificity (0.8 and above) for problematic drinking. The AUDIT demonstrates convergent

validity with other common screening measures ranging from 0.78-0.88, and it has test-retest reliability of $r = 0.86$ (2).

3. The *Posttraumatic Checklist – Military* (6) is a 17-item scale reflecting the symptoms of Posttraumatic Stress Disorder in the DSM-IV. Items are rated on a 5-point, Likert-type scale, ranging from “Not at all” (1) to “Extremely” (5). The *Posttraumatic Checklist* has demonstrated a high total-scale internal consistency coefficient (.97), convergent validity with other measures ranging from .77 to .93, reliability (.96) (6). Sensitivity and specificity for the PCL-M are inversely related based on the numerical threshold utilized for a “positive” screening. Bliese and colleagues (7) demonstrated that sensitivity is highest (.98) with an extremely low threshold of 17, but specificity is low (.33) at this value. In contrast, sensitivity drops dramatically (.24) with a threshold of 50, but sensitivity demonstrates a corresponding increase (.98). A clinical threshold of 30 or more has been used to identify individuals with at least moderately severe symptoms of posttraumatic stress disorder (48).
4. The *16 Personality Factor Questionnaire (16PF)* (14) is a 185-item measure that assesses five global personality factors and16 primary personality factors. The 16PF Global Factor scales include: extraversion, anxiety neuroticism, tough-mindedness, independence, and self-control. The 16PF Primary Factor scales include: warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. Subscale scores range from 0-10, with 0-3 indicating low, 4-7 indicating

moderate, and 8-10 indicating high on each personality factor. The 16PF demonstrates two-week test-retest reliability averages of 0.80 for the primary factors and 0.87 for the global factors (13). Internal consistency for the primary factors ranged from 0.66 to 0.86, and construct validity has been verified through correlations with other common personality inventories (13).

In addition to the global and primary factor scales, the 16PF has several additional scales that may be useful to characterize personality traits. The 16PF has three Response Style subscales, including impression management, infrequency, and acquiescence. The 16PF Protective Services Dimensions include: emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Pathology-Oriented Scales have also been generated and include: psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity.

Outcome Variables

1. *Overall training outcome*. The primary outcome variable predicted using the assessment measures was attrition (*i.e.*, failure) versus training completion (*i.e.*, pass/success). Attrition from the program may have occurred at any point during the training. This outcome variable is binary, and all participants were coded as either training attrition or training completion.
2. *Training program grade point average (GPA)*. GPA is a continuous outcome variable ranging from 0-100. GPA was based on student performance and

accuracy on mission-relevant tasks throughout the training, including academic tests and practical application exercises. Students who failed to complete the training program did not have GPA as an outcome measure.

3. *Instructor ratings*. Instructor ratings are a continuous outcome variable ranging from 0-5 on eight performance traits. These ratings were given by highly trained command staff members who were subject matter experts in the relevant domains being assessed in the training program. These eight ratings were averaged to yield a single instructor rating score per case. Students who failed to complete the training program did not have instructor ratings as an outcome measure.

DATA ANALYTIC PLAN

Data Management

Data were formatted for analysis in SPSS Version 22 (SPSS, Inc., Chicago, IL) and Mplus Version 7.31 (47). Data were then evaluated for accuracy of the data file. Descriptive statistics were used to determine whether the data were in range for the specific variables and if there were any implausible or impossible deviations in the documented responses. The missing data analytic plan is described in the following sections. The results from this analytic plan are detailed in the results section of this document (Chapter 4).

Missing Data

The data were then analyzed to determine which variables had greater than five percent missing data. If a single variable had a sizeable proportion (> 30%) of missing data, then the variable was considered for deletion from the analyses depending on its relative importance to the research and its relationship to other variables. For those variables with modest amounts (up to 29%) of missing data, SPSS Missing Values Analysis (MVA) was used to determine if the data were missing completely at random (MCAR), missing at random (MAR), or not missing at random (NMAR). If the missing data were random and there was one percent or fewer cases with missing data, the cases would be deleted from the analyses in lieu of deleting the variable.

The Multiple Imputation by Chained Equations (MICE), also known as the fully conditional specification (FCS) or the Markov Chain Montecarlo Method (MCMC), was used to generate univariate imputation models for each variable with missing data (54). Regression modeling estimated replacement means, standard deviations, covariances, and correlations of missing values based on their relationships to other variables in the data set. The use of 10 imputations was selected to reduce the likelihood of producing a Monte Carlo error (70). For each missing value, the mean of the replacement values from each of the 10 imputations was computed and included in the final imputed databases that were used for data analysis. As noted above, the analyses are detailed in Chapter 4 (Results) of this document.

Data Distribution

After evaluating and correcting for missing data, the distribution of the data was evaluated using SPSS version 22 (SPSS, Inc., Chicago, IL). Outliers were identified using frequency distributions, z-scores, and histograms. Outliers were analyzed to determine if they should be retained (*i.e.*, they were accurate and plausible entries) or Winsorized, or if the variable should have been transformed to increase normality.

Normality, linearity, and homoscedasticity were evaluated. Normality was assessed quantitatively using the Kolmogorov-Smirnov and Shapiro-Wilks tests.

Because these tests are highly sensitive with sample sizes greater than 1000, normality also was inspected visually using Q-Q and box plots (27). Skew and kurtosis were assessed quantitatively, with values of greater than or equal to 3 and 10 indicating abnormal skew and kurtosis, respectively. It is notable that the impact of differences from a normal distribution decrease as sample size increases (59); therefore, a statistically non-normal distribution in this dataset did not likely invalidate the analyses. Linearity was assessed using bivariate scatterplots in SPSS. Multicollinearity among continuous variables was assessed using bivariate Pearson's correlations; variables with correlations greater than or equal to 0.70 were identified, and one of the variables would be excluded from the analysis if this situation arose.

Sample Size and Power

The initial sample for this study was 4241 cases. Attrition was estimated to be 23.9% (n = 1013) of the sample. Different analyses have different conventions for ensuring enough power to detect statistical differences and to build generalizable models to predict the desired outcome. Logistic regression analyses should have 10-20 cases of the less common outcome (in this case, attrition) per variable included in the analysis (59). With 50 potential predictor variables and a conservative required 20 cases of attrition per variable, this sample was appropriately powered to run the analysis for Specific Aim 1. Recommendations for statistical regression with multiple regression have suggested cases-to-independent variables ratios of 40-to-1 (59). With 50 potential predictor variables, the sample of successful completers (n = 3228) was large enough to accommodate this recommended ratio; as such, analyses for Specific Aim 2 were adequately powered. Although there is no standard for sample size when conducting latent profile analyses, authors have recommended between 250-1000 participants (63). Therefore, the sample size of N = 4241 was adequate for this analysis.

Specific Aim 1

To identify variables that optimally predict training attrition in a military special program training command.

Hypothesis 1: Logistic regression was used to identify the optimal combination of continuous and categorical predictor variables that predict the discrete training outcome (*i.e.*, pass or fail). This logistic regression model provided a probabilistic prediction of training outcome based on the most parsimonious combination of these predictor variables.

Binary logistic regression was performed using SPSS Version 22 (SPSS, Inc., Chicago, IL) using statistical (stepwise) regression with backwards deletion for variable selection. 50 potential predictor variables spanning demographic, military, and psychological factors were evaluated in a series of logistic regression models. The dependent variable in all analyses was overall training outcome (coded as attrition/failure or pass/success). Omnibus tests were considered significant when p values were less than .05. Backwards deletion cutoff criterion also was set at $p < .05$. Following the completion of the stepwise regression steps (*i.e.*, in the final model), an α level of .005 was set *a priori* as the threshold for determining two-tailed statistical significance of individual predictors. This α level was chosen to adjust for the elevation in the family wise error rate associated with multiple comparisons and the potential bias associated with multiple imputation (37). However, individual predictor variables with p values between .01 and .005 were considered potentially significant, as this project was exploratory in nature.

The goodness-of-fit of the model identified by the logistic regression was evaluated using the Hosmer-Lemeshow statistic, in which a good-fitting model has a non-significant ($p > .05$) chi-square test from expected distribution of outcome. Additionally, the accuracy of classification of cases, or evaluating whether the model correctly or incorrectly predicted the cases' outcomes, was evaluated using the classification plot provided as part of the logistic regression output. Classification accuracy in logistic regression involves subjective assessment on the part of the researcher to determine which potential inaccurate classification (*i.e.*, false positive versus false negative) is acceptable given the potential costs of making the respective error (59). Because it is unreasonable to expect a perfect logistic regression model, there are potential benefits and drawbacks associated with both possible outcomes (false positive and false negative classification). As such, classification accuracy was considered acceptable if greater that 60% of the cases were correct, good if 75% or greater were correct, and excellent if 90% or greater were correct. In reaching these classification standards, it was taken for granted that improvements in predictor variables are the only way to improve classification (59). Cross-validation was considered for the logistic regression analyses but ultimately reject. It was determined that the loss of numbers of less frequent outcome (i.e., attrition) in an 80/20% split would be more of a detriment than an improvement to the model when predicting outcome.

Specific Aim 2

To identify variables that optimally predict relative performance in a military special program training setting among program completers.

Hypothesis 2a: Among training program completers, multiple regression was used to predict the final training program GPA from the predictor variables. This approach allowed a relative assessment of the predictive contribution of each independent variable towards program success as measured by performance on graded tasks.

Multiple regression was performed using SPSS Version 22 (SPSS, Inc., Chicago, IL) using statistical (stepwise) regression with backwards deletion for variable selection. 50 potential predictor variables spanning personal characteristics and psychological factors were evaluated in a series of multiple regression models. The dependent variable in all analyses was final training program GPA. It was hypothesized that intelligence and physical fitness would strongly predict performance because the GPA was based on physical and mental tasks, but the purpose of these analyses was to determine which variables, without theory, predicted relative performance in the program. Omnibus tests were considered significant when p values were less than .05. An α level of .005 was set *a priori* as the threshold for determining statistical significance of individual predictors to adjust for the elevation in the family wise error rate associated with multiple comparisons and the potential bias associated with multiple imputation (37).

Cross-validation was used to evaluate the generalizability of the results from this analysis. The data were randomly split 80/20%; the primary analyses were conducted using the training data (80%), and the results were tested on the validation (20%) set (59).

Hypothesis 2b: Among training program completers, multiple regression was used to predict the final training program instructor evaluation from the predictor variables. This approach allowed a relative assessment of the predictive contribution of each independent variable towards program success as measured by instructor evaluation of the student.

Multiple regression was performed using SPSS Version 22 (SPSS, Inc., Chicago, IL) using stepwise (statistical) regression with backwards deletion for variable selection. 50 potential predictor variables spanning personal characteristics and psychological factors were evaluated in a series of multiple regression models. The dependent variable in all analyses was final training program instructor rating. The purpose of these analyses was to determine which variables, without theory, predict relative performance in the program. Omnibus tests were considered significant when *p* values were less than .05. An α level of .005 was set *a priori* as the threshold for determining statistical significance of individual predictors to adjust for the elevation in the family wise error rate associated with multiple comparisons and the potential bias associated with multiple imputation (37).

Cross-validation was used to evaluate the predictive equation developed in this analysis. The data were randomly split 80/20%, the predictive equation was developed using the training data (80%), and it was tested on the cross-validation (20%) set (59).

Specific Aim 3

To identify, characterize, and validate subgroups within the sample population for whom different combinations of variables more accurately predict training attrition risk.

Hypothesis 3: Latent Profile Analysis (LPA) was used to analyze the significant predictors identified in Specific Aim 1 to divide the sample into subgroups. These subgroups were maximally heterogeneous between groups and maximally homogenous within group. The identified subgroups resulting from the analysis were compared to the existing training outcome data (pass/fail) to validate the subgroups according to their relative risk of training failure.

Latent profile analysis (LPA) was conducted using Mplus Version 7.31 (47). The number of latent subgroups within a sample was hypothesized to be discrete, and these subgroups ideally had large statistical distance between them. However, selection of the number of subgroups was based on statistics and theory, and this selection impacted the interpretation of the subgroups found in the analysis. Subgroups in the LPA were initially grounded in the *a priori* hypothesis that there were three attrition risk groups (low, moderate, and high) within the sample. Additionally, four selection and fit criteria were utilized in confirming or disconfirming the hypothesized three latent subgroup solution: Bayesian Information Criterion (BIC); adjusted Bayesian Information Criterion (aBIC); Akaike Information Criterion (AIC); and corrected Akaike Information Criterion (cAIC). Model entropy also was assessed to determine the number of subgroups, where entropy ranges from zero to one with higher entropy indicating greater certainty of classification. Although there was no empirically-derived consensus for the best fit criteria tests, researchers have found these methods to be reliable in accurately determining the number of latent subgroups in simulation studies (63).

After identifying the best-fitting LPA model, participants were assigned to a discrete latent subgroup based on their posterior probabilities of belonging to each potential subgroup. As such, the latent subgroup to which each individual was assigned reflected the subgroup in which they were most likely to belong (*i.e.*, had the highest posterior probability value), despite subgroup membership not being fixed in reality. This case assignment approach was used to facilitate the practical and clinical interpretation of the subgroups in their subsequent characterization and validation.

CHAPTER 4: Results

The following sections cover all steps taken from data preparation through the completed analyses. The first section discusses the overall pattern of missing data for participants and study variables. The second section covers the data imputation for the respective aims. The third section details the analyses and findings for specific aims 1 and 2 (Logistic Regression and Multiple Regression, respectively). The final portion of this chapter presents the results from specific aim 3 (Latent Profile Analyses).

Cleaning the Data

The original database contained 4241 cases. 42 cases were dropped because of missing overall training outcome, the most important dependent variable. One participant was dropped because the training outcome value was out of range (*i.e.*, letter instead of a number). Twelve additional participants were dropped due to out of range data (entries that were impossible) for some variables (*i.e.*, rank, GT score, age).

Two separate databases were created to run the logistic and multiple regression models. Moving forward, these databases are referred to as "Overall Training Outcome" (logistic regression) and "Successful Training Performance" (multiple regression) databases, respectively. The variables were then analyzed to determine the percentage of missing data for each variable (rather than for each case). For the Overall Training Outcome database, only three (of 50) variables were missing >30% of data: abuse history (45.4%), first age of alcohol use (59.5%), and PCL-M (38.8%). For the Successful Training Performance database, only four (of 52) variables were missing a large portion of data: instructor ratings (40.6%), abuse history (44.6%), first age of alcohol use (56.9%), and PCL-M (37.4%).

At this juncture, it was necessary to decide whether to keep these variables and lose hundreds to thousands of cases when conducting the analyses or drop these few variables and retain the maximum number of cases. This decision process is detailed below.

Analyzing Variables for Missing Data

All missing data analyses were conducted using IBM SPSS version 22 (SPSS, Inc., Chicago, IL). An α level of 0.05 was set *a priori* as the threshold for statistical significance when testing for missing data. All tests were two-tailed.

PCL-M

Overall Training Outcome Database: A 2x2 Pearson's chi-square analysis was conducted to examine the association between the PCL-M data missingness and training outcome. Individuals missing the PCL-M were significantly more likely to experience training failure (22.7%, n = 368) than those who were not missing the PCL-M (17.9%, n = 460), χ^2 (df =1) = 14.07, $p < .001$.

Among individuals with PCL-M data available (N = 2563), binary logistic regression analyses were conducted using training outcome as the dependent variable and PCL-M as the predictor variable. When the PCL-M score was continuous, there was a statistically significant, association with training outcome, OR = 1.09 (95% CI: 1.06, 1.12), χ^2 (1, N = 2563) = 39.25, $p < .001$, Nagelkerke $R^2 = 0.03$. Additionally, only 1.5% of training failures (n = 7 out of 460) were correctly identified, indicating that the continuous PCL-M score had poor sensitivity for predicting training outcome. When PCL-M was transformed into a dichotomous variable (coded as meeting the PTSD screening threshold of ≥30 or not), there was a significant association between PCL-M and training outcome, χ^2 (1, N = 2563) = 27.16, p < .001, Nagelkerke R^2 = 0.02. Individuals who met the PTSD screening threshold were significantly less likely to complete the training program successfully relative to those who were below the PTSD screening threshold, OR = 0.17 (95% CI: 0.09, 0.33). However, only 0.01% of the sample (n = 38) met the most liberal clinical threshold recommended for the PCL-M, suggesting that the statistic is not efficient with this great a difference between group sizes. Descriptively, of these 38 individuals, 21 (55.3%) were dropped from the training program. Of the 21 individuals dropped, 13 (61.9%) were dropped for psychological concerns with the other eight dropped for performance-related problems.

Successful Training Performance Database: Independent samples t-tests were conducted to determine if the final GPA and instructor ratings means were different for individuals with and without PCL-M data. Individuals with PCL-M data had significantly higher GPAs than those without PCL-M data (Present PCL-M: M = 93.74, SD = 2.97 vs. Missing PCL-M: M = 90.4, SD = 3.57), t(2246.33) = 27.64, p < .001, d = 1.17. However, there was no association between PCL-M data missingness and instructor ratings (Present PCL-M: M = 3.49, SD = 0.53 vs. Missing PCL-M: M = 3.56, SD = 0.59), t(1992) = -1.16, p = .25, d = -0.05.

Correlation analyses were conducted to examine associations among GPA, instructor rating, and PCL-M scores. The continuous PCL-M total score was significantly, negatively correlated with final GPA, r = -0.11, p < .01, but was not significantly related to instructor rating, r = -0.03, p = .19. When PCL-M was considered dichotomously (coded as meeting the PTSD screening threshold of ≥30 or not), point-biserial correlations revealed a significant relationship between the PCL-M and GPA, r =

-0.05, p = .015, but was not significantly related to instructor rating, r = -0.03, p = .24.

When considering the very modest effect size of the statistical relationships between the PCL-M and training outcome, the PCL-M score did not appear to have a functional impact on training outcome. The PCL-M also had very poor sensitivity in predicting training failures accurately. Furthermore, many cases would be retained in the analyses by dropping the variable. As such, the PCL-M was dropped from subsequent analyses. Overall, there is likely a very small relationship between the PCL-M and training outcome with limited predictive utility because the sample is relatively healthy, rather than a clinical sample, and participants should not have functionally impairing symptoms of PTSD.

First Age of Alcohol Use

Overall Training Outcome Database: A 2x2 Pearson's chi-square analysis was conducted to compare individuals with and without the variable "First Age of Alcohol Use (ETOH)" on training outcome. Individuals missing ETOH were significantly more likely to experience training failure (13.9%, n = 581) than those who were not missing ETOH (5.9%, n = 247), χ^2 (df =1) = 48.78 $p < .001$.

Among individuals with ETOH data available (N = 1695), a logistic regression analysis was conducted using training outcome as the dependent variable and ETOH as the predictor variable. There was no statistical relationship between ETOH and training outcome, OR = 0.96 (95% CI: 0.91, 1.00), χ^2 (1, N = 1695) = 3.54, p = .06, Nagelkerke R^2 = 0.004. Furthermore, zero training failures (out of 247) were correctly predicted.

Successful Training Performance Database: Independent samples t-tests were conducted to determine if the final GPA and instructor rating means were different for individuals with and without ETOH data. Individuals with ETOH data had significantly higher GPAs than those without ETOH data (Present ETOH: M = 94.01, SD = 2.75 vs. Missing ETOH: M = 91.34, SD = 3.73), t(3273.85) = 23.59, $p < .001$, $d = 0.82$. Similarly, individuals with ETOH data had significantly higher instructor ratings relative to those without ETOH data (Present ETOH: M = 3.51, SD = 0.54 vs. Missing ETOH: M = 3.44, SD = 0.51), t(1992) = 2.50, $p = .01$, $d = 0.11$.

Bivariate Pearson's correlations were conducted to examine associations among GPA, instructor rating, and the first age of ETOH. There was no statistically significant relationship between ETOH and the outcome variables, *p*s > .05.

There may be a statistically significant relationship between missinginess and training outcome; however, among individuals with ETOH data, this variable added no predictive value. Additionally, many cases would be retained in the analyses by dropping the variable, so First Age of Alcohol Use (ETOH) was dropped from the analyses. The loss of these data had a minimal impact on the intended concept (alcohol use patterns), as the AUDIT remains a variable in the analyses.

Abuse History

<u>Overall Training Outcome Database</u>: A 2x2 Pearson's chi-square analysis was conducted to compare individuals with and without the abuse history variable on training outcome. Individuals missing abuse history were significantly less likely to experience training failure (9.7%, n = 404) than those who were not missing this variable (10.1%, n = 424), χ^2 (df =1) = 4.66 $p < .05$.

Among individuals with abuse history data available (N = 2283), a logistic regression was conducted using training outcome as the dependent variable and abuse history (coded as present or absent) as the predictor variable. There was a significant association between abuse history and training outcome, χ^2 (1, N = 2283) = 73.17, p < .001, Nagelkerke R^2 = 0.04. Individuals endorsing a history of abuse were significantly less likely to complete the training program successfully relative to those without an abuse history, OR = 0.26 (95% CI: 0.19, 0.36). Despite this significant association, abuse history did not correctly identify any cases as training failures (0.00%), suggesting very poor sensitivity and limited utility in predicting overall training outcome.

<u>Successful Training Performance Database</u>: Independent samples t-tests were conducted to determine if the final GPA and instructor rating means were different for individuals with and without abuse history data. Individuals with abuse history data had significantly higher GPAs than those who were missing abuse history data (Present Abuse History Data: M = 93.46, SD = 3.07 vs. Missing Abuse History Data: M = 91.30, SD = 3.81), t(2829.47) = 17.63, p < .001, d = 0.66. Similarly, individuals with abuse history data had significantly higher instructor ratings relative to those who were missing abuse history data (Present Abuse History Data: M = 3.51, SD = 0.54 vs. Missing Abuse History Data: M = 3.43, SD = 0.51), t(1992) = 2.86, p = .004, d = 0.13.

Point-biserial correlations were conducted to examine the associations between abuse history and the final performance outcomes (*i.e.*, GPA, instructor rating). Abuse history and GPA were significantly, negatively correlated, r = -0.07, p < .01. However, there was no relationship between abuse history and instructor rating, r = -0.02, p = .45.

There appears to be a statistically significant relationship between missingness and training outcome; however, the usefulness of these data is questionable, and the effect size is small. When considering that these statistical relationships do not appear to have a large functional impact on training outcome and performance, and that many cases would be retained in the analyses by dropping the variable, abuse history was dropped from the analyses.

Instructor Rating

Overall Training Outcome Database: This variable was not included in this database and therefore no analyses were conducted to examine the impact of missingness.

Successful Training Performance Database: Because instructor rating was an outcome variable, this variable was retained, and all cases that had the data were included in the relevant analyses.

Analyzing Cases for Missing Data

To determine the extent of missing data per case, each of the remaining variables in the databases were recoded into binary variables (data present or absent). These binary variables were used to compute a new variable that was the percentage of data missing per case ([total number of missing variables / total variables] * 100). *A priori*, cases with missing 50% of data or greater were dropped. Using the 50% criteria, four cases were dropped from the Overall Training Outcome database and 25 cases were dropped from the Successful Training Performance database.

It was not feasible to develop a model for imputation of the 16PF variables because the 16PF scale scores generated from the individual items have incredibly complex scoring procedures and rely upon all individual data points being present. For example, if one 16PF item was missing, all of the 16PF scales were missing (*i.e.*, there are no partial data). As such, any case that was missing the 16PF was dropped from further analyses. In total, 125 cases were dropped from Overall Training Outcome database and 213 cases were dropped from the Successful Training Performance database due to missing 16PF data. Overall, the Overall Training Outcome database retained N_{OTO} = 3947 cases, and the Successful Training Performance database retained N_{STP} = 3228 cases. At this point, the data were ready for imputing the remaining missing values.

MULTIPLE IMPUTATION

Missing value analysis and data imputation were conducted using the Missing Values Add-On of IBM SPSS version 22 (SPSS, Inc., Chicago, IL). The Overall Training Outcome (N_{OTO} = 3947) and Successful Training Performance databases (N_{STP} = 3228) were examined independently because different variables required imputation for the respective study aims.

SPECIFIC AIM 1 – OVERALL TRAINING OUTCOME DATABASE

Analyzing Extent and Patterns of Missing Data

The following variables were examined to determine the extent of missing data: age, sex, GT score, military occupational specialty, total combat deployments, physical fitness score, disciplinary counseling statements, mental health treatment history, AUDIT total score, and 16PF scales. The overall patterns of missingness were as follows: 38.0% (n = 19) of the variables had at least 0.01% missing data; 25.9% (n = 1023) of individual participants had at least one missing data point; and less than 1% (n = 1362) of all potential values were missing. Variables that had greater than 3% missing values included: total combat deployments (12.7%, n = 500), physical fitness score (8.1%, n = 319), disciplinary counseling statements (6.9%, n = 272), and mental health treatment history (3.0%, n = 117).

Monotonicity of missing data was not observed. As such, missing values appeared to be missing in a random pattern. The missing value pattern (*i.e.*, the combination of variables participants were missing) frequencies were as follows: by far, the most common data pattern was for participants to have no missing values across all of the variables (74.1%). The next most common missing data pattern, albeit much less frequent, was for participants to only have missing data on their combat deployment history (~10%). The third and fourth most common patterns were missing only physical fitness score and only disciplinary counseling statement data, respectively (< 5%).

Independent samples t-tests and Pearson's chi-square tests were used to compare participants with and without any missing data values on continuous and categorical predictor variables, respectively. There were no significant group differences in age, sex, GT score, physical fitness score, total combat deployments, and AUDIT total score (*p*s > .05). Participants with missing data had more disciplinary counseling statements than those without missing data (Missing: M = 1.94, SD = 1.90 vs. Present: M = 1.51, SD = 1.60), t(1022.38) = -5.68, $p < .001$, $d = -0.36$. Significant group differences also emerged for rank (χ^2 (df = 3) = 30.30, $p < .001$, $\varphi = 0.09$), military occupational specialty (χ^2 (df = 2) = 6.62, $p = .037$, $\varphi = 0.04$), and mental health treatment history (χ^2 (df = 1) = 10.83, $p = .001$, $\varphi = 0.05$). However, the effect sizes were small and likely did not indicate functional or practical differences between individuals with and without missing data.

Overall, there were missing data in this dataset. However, the missing data appeared to be a very small percentage of the data in its entirety, and the data appeared to be missing at random. Further, there did not appear to be any meaningful differences between participants with and without missing data. As such, it was appropriate to impute the missing values.

Imputing Missing Data

Prior to conducting the data imputation procedures, the distribution of continuous variables in the dataset that were to be used as predictors but not imputed was examined for normality, linearity, and homoscedasticity. Five of the 16PF scales had abnormal skew and kurtosis across the original and imputed datasets: impression management, infrequency, liveliness, anxiety, and independence. Each of these variables had one impossible outlier; this outlier was recoded to be at the highest range of the possible values. After recoding to possible values, the skew and kurtosis of these variables were satisfactory.

Multiple Imputation by Chained Equations (MICE) was used to generate univariate imputation models for each variable with missing data, and regression modeling estimated replacement values. Variables included in the imputation models were: training outcome (pass/fail), age, sex, GT score, rank, total combat deployments, physical fitness score, disciplinary counseling statements, mental health treatment history, AUDIT total score, and the 16PF scales. Data were imputed for eight variables: age, sex, combat deployments, physical fitness, disciplinary counseling statements, mental health treatment history, AUDIT total score, and GT score. The imputation program was unable to estimate replacement values for military occupational specialty (military occupational specialty: $n = 19$) because of the large number of parameters (more than 100) required to estimate the replacement values. As such, these data were left missing for all subsequent analyses.

Follow-up diagnostic tests on the imputed data were conducted to assess the adequacy of the resulting imputation model. Independent samples t-tests or Pearson's chi-squared tests compared the distributions of the original data and the imputed data for each variable. No significant differences emerged. Therefore, it was concluded that the imputed data fit the dataset well, and it was appropriate to continue with the planned data analyses.

SPECIFIC AIM 2 – SUCCESSFUL TRAINING PERFORMANCE DATABASE

Analyzing Extent and Patterns of Missing Data

The following variables from the Successful Training Performance database were examined to determine the extent of missing data: GPA, instructor rating, age, sex, GT score, military occupational specialty, total combat deployments, physical fitness score, disciplinary counseling statements, mental health treatment history, AUDIT total score, and the 16PF scales. The overall patterns of missingness were as follows: 38.5% (n = 20) of the variables had at least 0.01% missing data; 52.0% (n = 1677) of individual participants had at least one missing data point; and 1.3% (n = 2194) of all potential values were missing. Variables that had greater than 3% missing values included: instructor rating (38.5%, n = 1243), total combat deployments (12.1%, n = 390), physical fitness score (5.2%, n = 167), and disciplinary counseling statements (4.8%, n = 156).

Monotonicity of missing data was not observed. As such, missing values appeared to be missing in a random pattern. The missing value pattern frequencies were as follows: the most common data pattern was for participants to have no missing values across all of the variables (48.1%). The next most common missing data pattern was for participants to only have missing data on their instructor rating (~30%). The third most common missing data patterns was for participants to be missing only combat deployment history (~10%).

Independent samples t-tests and Pearson's chi-square tests were used to compare participants with and without any missing data values on continuous and categorical predictor variables, respectively. There were no significant group differences in sex, GT score, military occupational specialty, mental health treatment history, and AUDIT total score (ps > .05). Relative to those without any missing data, participants with missing data were younger (Missing: M = 22.29, SD = 2.21 vs. Present: M = 22.46, SD = 2.31), t(3177.88) = 2.24, $p = .025$, $d = 0.08$; had more disciplinary counseling statements (Missing: M = 1.67, SD = 1.58 vs. Present: M = 1.35, SD = 1.41), t(3013.06) = -6.07, $p < .001$, $d = -0.22$; poorer physical fitness scores (Missing: M = 250.79, SD = 28.73 vs. Present: M = 256.38, SD = 27.04), t(3033.78) = 5.55, $p < .001$, $d = 0.22$; and a higher number of combat deployments (Missing: M = 0.73, SD = 0.87 vs. Present: M = 0.42, SD = 0.68), t(2392.57) = -10.40, $p < .001$, $d = -0.41$. Significant group differences also emerged for rank (χ^2 (df = 2) = 25.57, $p < .001$, $\varphi = 0.09$). This pattern of results potentially suggested that participants with missing data who completed the training program had slightly more adverse performance than program completers who were not missing data. Despite the statistical significance of these differences, the effect sizes were small and therefore may be negligible.

Overall, there were missing data in this dataset. They appeared to be a small percentage of the data in its entirety, and the data appeared to be missing at random. Further, based on the preceding analyses, there may be few meaningful differences between participants with and without missing data. As such, it was appropriate to impute the missing values.

Imputing Missing Data

Multiple Imputation by Chained Equations (MICE) was used to generate 10 univariate imputation models for each variable with missing data, and regression modeling estimated replacement values. Variables included in the imputation models were: GPA, instructor rating, age, sex, GT score, total combat deployments, physical fitness score, disciplinary counseling statements, mental health treatment history, AUDIT total score, and the 16PF scales. Data were imputed for 10 variables: GPA, instructor rating, age, sex, GT score, total combat deployments, physical fitness score, disciplinary counseling statements, mental health treatment history, and AUDIT total score. The imputation program was unable to estimate replacement values for military occupational specialty due to the small number of missing values (n = 13) and the large number of parameters (more than 100) required to estimate the replacement values.

As described above with the Overall Training Outcome database, the AUDIT total score presented several challenges to multiple imputation. Data could not be imputed when constraints were placed upon the possible values of the AUDIT total score (0-40). As such, data imputation models were run without placing constraints on the AUDIT total score, which yielded between one to eight (out of 42 total) negative replacement values (*e.g.*, -10) in each of the 10 imputations. These issues persisted when data imputation models were re-run with a logarithm-transformed AUDIT total score in an attempt to normalize the data. As such, the original, untransformed AUDIT total score values were used in the data imputation models, and the negative replacement values were recoded to zero so they represented a possible value in each imputed dataset.

Combat deployment history presented similar challenges to imputation as the

AUDIT total score data. Data could not be imputed when constraints were placed upon the possible values of the total number of past combat deployments (0-10). As such, data imputation models were run without placing constraints on the combat deployment frequency, which yielded between 42 to 61 (out of 390 total) negative replacement values (*e.g.*, -10) across the 10 imputed datasets. These issues persisted when data imputation models were re-run with a logarithm-transformed combat deployment frequency values in an attempt to normalize the data. As such, the original, untransformed combat deployment history data were used in the imputation models, and the negative replacement values were recoded to zero so they represented a possible value in all of the imputed datasets.

Follow-up diagnostic tests on the imputed data were conducted to assess the adequacy of the result imputation model. Independent samples t-tests or Pearson's chi-squared tests compared the distributions of the original data and the imputed data for each variable. No significant differences emerged. Therefore, it was concluded that the imputed data fit the dataset well, and it was appropriate to continue with the planned data analyses.

RESULTS: SPECIFIC AIMS 1 & 2

The following sections detail the data distributions for the databases as well as the primary analyses for specific aim 1 (overall training outcome) and specific aim 2 (successful training performance). Because the hypotheses were statistically-driven (vs. theoretically driven), multiple statistical models are presented for each aim. All analyses were conducted using IBM SPSS version 22 (SPSS, Inc., Chicago, IL). Omnibus tests were considered significant when p values were less than .05. An α level of .005 was set *a priori* as the threshold for determining statistical significance of individual predictor variables to adjust for the elevation in the family wise error rate associated with multiple comparisons and the bias associated with multiple imputation.

Overall Training Outcome Database – Data Distribution

Following the data imputation procedures, the distribution of continuous variables in the dataset was examined for normality, linearity, and homoscedasticity. The skew and kurtosis for all variables were satisfactory.

The Kolmogorov-Smirnov and Shapiro-Wilks tests were significant for all continuous variables (*p*s < .05), suggesting that these distributions statistically deviated from normality. However, these tests are highly sensitive with sample sizes greater than 1000, and therefore it was appropriate to conduct a visual inspection of graphical plots alongside quantitative results (27). Visual inspection of the Q-Q plots and box plots indicated adequately normal distributions for all continuous variables. As such, no additional transformations were applied to the data.

Based on these inspections of the data, it was determined that it was appropriate to continue with the logistic regressions as planned.

Specific Aim 1 Results – Overall Training Outcome

Binary logistic regression models with backward elimination (likelihood ratio) were conducted to examine associations among demographic, military, and psychological factors (*i.e.*, predictor variables) and overall training outcome (*i.e.*, dependent variable) in a military special program training command. The primary objective was to identify the most parsimonious model to predict attrition (vs. success) in the training program. Predictor variables included age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness scores, mental health treatment history, total combat deployments, AUDIT total score, and the 16PF scales. The binary dependent variable was training outcome, coded as pass or fail.

The following subsections present the database sample characteristics followed by findings from the individual regression models conducted. These sections describe: predictors entered in the model; model significance and fit; classification accuracy; and identified significant predictors of training outcome. Of note, the classification cutoff utilized by SPSS is 0.5. This cutoff, while commonly used, is arbitrary, and the likelihood of committing a Type I versus Type II error changes based on the cutoff value. This standard classification cutoff of 0.5 decreases the likelihood of finding lower probability outcomes (*e.g.*, training failure).

Overall Training Outcome Database – Sample Characteristics

Data from 3947 service members (92.6% male; age: M = 22.38, SD = 2.28 years) were included in analyses to examine predictors of attrition from the training program. The sample breakdown by military occupational specialty was: 23.1% (N = 912) air, 25.0% (N = 987) ground, and 51.4% (N = 2029) support. The sample breakdown by rank was: 0.1% (N = 2) E2, 24.1% (N = 950) E3, 55.6% (N = 2195) E4, and 20.3% (N = 800) E5. Service members had between 0 and 5 prior combat deployments (M = 0.58, SD = 0.75). In general, service members had an average GT score (M = 109.11, SD = 11.82) and above average physical fitness score (M = 252.04, SD = 27.94). They had an average of 1.63 (SD = 1.62) prior disciplinary counseling statements at the command. Approximately 13.3% (N = 526) of service members reported previously seeking treatment for a mental health problem. The training program attrition rate was 18.2%

(N = 718).

Model 1: Participant Characteristics Only

Predictors included age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, total combat deployments, and AUDIT total score. Combat deployments (step 2) and age (step 3) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (10, N = 3947) = 319.40, $p < .001$. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3947) = 10.44, $p = .236$, further suggesting an adequate model fit. The overall fit of the model was fair, with pseudo-R^2 values of 0.078 (Cox and Snell R^2) and 0.128 (Nagelkerke R^2). Overall classification was very good, as the model predicted 82.5% of total responses correctly. Classification of passes (99.0% correctly classified) was excellent, while classification of failures (8.7% correctly classified) was poor.

Table 1.1 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals who were male (vs. female) or had previous mental health treatment (vs. none) were approximately 1.7-1.8 times more likely to fail out of the training program (*p*s < .001). Relative to individuals with an E-2 and E-3 rank, individuals with an E-4 rank ($p < .001$), but not those with an E-5 ($p = .15$) rank, were more likely to fail out of the training program. Lower GT scores, poorer physical fitness scores, more prior disciplinary counseling statements, and higher AUDIT total scores were also associated with a greater likelihood of attrition (*p*s < .001). Military occupational specialty was not significantly related to training program outcome (*p*s = .024-.812).

Model 2: 16PF Response Style Subscales Only

Predictors included impression management, infrequency, and acquiescence. Acquiescence (step 2) was removed through the backward elimination procedure. The final overall model was statistically significant, χ^2 (2, N = 3946) = 63.18, $p < .001$. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3946) = 10.23, $p = .25$, further suggesting an adequate model fit. The overall fit of the model was poor, with pseudo-R^2 values of 0.016 (Cox and Snell R^2) and 0.026 (Nagelkerke R^2). Overall classification was very good, as the model predicted 81.8% of total responses correctly. Classification of passes (100% correctly classified) was excellent, while classification of failures (0.0% correctly classified) was very poor.

Table 1.2 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Lower impression management scores were associated with a greater likelihood of failure to complete the training program ($p < .001$), whereas higher scores on infrequently endorsed items were associated with an increased likelihood of attrition at a trend level ($p = .006$).

Model 3: 16PF Global Factor Patterns Only

Predictors included extraversion, anxiety, tough-mindedness, independence, and self-control. Tough-mindedness (step 2), extraversion (step 3), and independence (step 4) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (2, N = 3943) = 137.85, $p < .001$. The Hosmer-Lemeshow test was significant, χ^2 (8, N = 3943) = 22.09, $p = .005$, which suggests that the final model was not an adequate fit for predicting overall training outcome. The overall fit of the model was poor, with pseudo-R^2 values of 0.034 (Cox and Snell R^2) and 0.056 (Nagelkerke R^2). Overall classification was very good, as the model predicted 81.9% of total responses correctly. Classification of passes (99.9% correctly classified) was excellent, while classification of failures (1.0% correctly classified) was very poor.

Table 1.3 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals reporting higher anxiety and/or lower self-control were more likely to fail out of the training program (ps < .001).

Model 4: 16PF Primary Factor Profile Subscales Only

Predictors included warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. Dominance (step 2), warmth (step 3), sensitivity (step 4), privateness (step 5), openness to change (step 6), liveliness (step 7), apprehension (step 8), vigilance (step 9), and perfectionism (step 10) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (7, N = 3945) = 240.43, $p < .001$. The Hosmer-Lemeshow test was significant, χ^2 (8, N = 3945) = 5.67, p =.68, which suggests that the final model was an adequate fit for predicting overall training outcome. The overall fit of the model was poor, with pseudo-R^2 values of 0.059 (Cox and Snell R^2) and 0.096 (Nagelkerke R^2). Overall classification was very good, as the model predicted 82.2% of total responses correctly. Classification of passes (99.5% correctly classified) was excellent, while classification of failures (4.3% correctly classified) was poor.

Table 1.4 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Lower reasoning capacity ($p < .001$), emotional stability ($p < .001$), and rule-consciousness ($p = .002$) were associated with a greater likelihood of attrition from the training program. Individuals reporting higher abstractedness ($p < .001$) and/or self-reliance ($p = .004$) were also more likely to fail out of the training program. Higher social boldness was associated with an increased likelihood of attrition at a trend level ($p = .008$), whereas tension was not significantly related to overall training outcome ($p = .10$)

Model 5: 16PF Protective Services Dimensions Only

Predictors included emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Interpersonal relations (step 2) and integrity/control (step 3) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (2, N = 3946) = 185.62, $p < .001$. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3946) = 12.37, $p =.14$, further suggesting an adequate model fit. The overall fit of the model was poor, with pseudo-R^2 values of 0.046 (Cox and Snell R^2) and 0.075 (Nagelkerke R^2). Overall classification was very good, as the model predicted 81.9% of total responses correctly. Classification of passes (99.6% correctly classified) was excellent, while classification of failures (2.1% correctly classified) was poor.

Table 1.5 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Lower emotional adjustment and intellectual efficiency scores were associated with an increased likelihood of failure to complete the training program (ps < .001).

Model 6: 16PF Pathology-Oriented Scales Only

Predictors included psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity. Obsessional thinking (step 2), psychological inadequacy (step 3), self-reproach (step 4), thrill seeking (step 5), anxious depression (step 6), and paranoid ideation (step 7) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (6, N = 3943) = 232.11, $p < .001$. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3943) = 4.19, $p = .84$, further suggesting an adequate model fit. The overall fit of the model was poor, with pseudo-R^2 values of 0.057 (Cox and Snell R^2) and 0.093 (Nagelkerke R^2). Overall classification was very good, as the model predicted 81.9% of total responses correctly. Classification of passes (98.9% correctly classified) was excellent, while classification of failures (5.3% correctly classified) was poor.

Table 1.6 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Reports of higher health concerns ($p < .001$), lower energy state ($p = .005$), and greater alienation/perceptual distortion ($p < .001$) were associated with an increased likelihood of attrition from the training program. Additionally, at a trend level, individuals with higher reports of suicidal thinking were more likely to fail out of the training program ($p = .006$). Apathetic withdrawal and threat immunity scores were not significantly related to training outcome (ps = .01).

Model 7: Significant and Marginally Significant Predictors from Models 1-6

Predictors included sex, rank, GT score, disciplinary counseling statements, physical fitness scores, mental health treatment history, AUDIT total score, impression management, infrequency, anxiety, self-control, reasoning, emotional stability, rule-consciousness, abstractedness, self-reliance, social boldness, emotional adjustment, intellectual efficiency, health concerns, low energy state, health concerns, alienation/perceptual distortion, and suicidal thinking. Prior to running this model, bivariate Pearson correlations were conducted to evaluate the collinearity among predictor variables. The vast majority of variable did not demonstrate multicollinearity, as they had r values less than 0.80. However, emotional adjustment was highly correlated with anxiety (r = -0.85, p < .001) and emotional stability (r = 0.82, p < .001). Additionally, intellectual efficiency was highly correlated with reasoning (r = 0.87, p < .001). As such, emotional adjustment and intellectual efficiency were not included in the logistic regression model so that there would be no multicollinearity among the predictor variables. Rule-consciousness (step 2), impression management (step 3), infrequency (step 4), abstractedness (step 5), anxiety (step 6), low energy state (step 7), and suicidal thinking (step 8) were removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (16, N = 3940) = 483.73, p < .001. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3940) = 14.32, p =.07, further suggesting an adequate model fit. The overall fit of the model was fair, with pseudo-R^2 values of 0.115 (Cox and Snell R^2) and 0.188 (Nagelkerke R^2). Overall classification was very good, as the model predicted 83.1% of total responses correctly. Classification of passes (97.9% correctly classified) was excellent, while classification of failures (16.7% correctly classified) remained poor but was an improvement over prior models.

Table 1.7 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals who were male (vs. female) or had previous mental health treatment (vs. none) were significantly more likely to fail out of the training program (*p*s < .001). Relative to individuals with an E-2 or E-3 rank, individuals with an E-4 rank (*p* = .001), but not those with an E-5 rank (*p* = .24), were more likely to fail out of the training program. Lower GT scores, poorer physical fitness scores, more prior disciplinary counseling statements, and higher AUDIT total scores were also associated with a greater likelihood of attrition (*p*s ≤ .001). Individuals reporting lower reasoning ability had a greater likelihood of program attrition (*p* < .001). Those who reported higher self-reliance (*p* = .001) and more health concerns (*p* < .001) were also more likely to fail out of the training program. Greater ratings of alienation/perceptual distortion were associated with a higher likelihood of attrition on a trend level (*p* = .006). Self-control (*p* = .039), emotional stability (*p* = .08), social boldness (*p* = .03), and threat immunity (*p* = .02) were not significantly related to overall training outcome.

Model 8: All Potential Predictors (Personal, Military, and Psychological Factors)

Prior to running this model, bivariate Pearson correlations were conducted to evaluate the collinearity among predictor variables. The vast majority of variables did not demonstrate multicollinearity, as they had r values less than 0.80. However, the following variables were highly correlated with each other (*p*s < .001): integrity/control with self-control (*r* = .93) and rule-consciousness (*r* = 0.91); self-control with rule-consciousness (*r* = 0.82); dominance with independence (*r* = 0.82); emotional adjustment with emotional stability (*r* = 0.82) and anxiety (*r* = -0.85); intellectual efficiency with reasoning (r = 0.87); and interpersonal relations with extraversion (r = 0.96). As such, rule-consciousness, emotional adjustment, intellectual efficiency, dominance, and extraversion were not included in the logistic regression model so that there would be no multicollinearity among the predictor variables.

The stepwise logistic regression model used 21 steps to reach the final model. In sequential order, the following variables were removed one at a time through the backward elimination procedure: tough-mindedness, impression management, independence, paranoid ideation, anxious depression, self-reliance, perfectionism, infrequency, obsessional thinking, openness to change, self-reproach, privateness, emotional stability, vigilance, tension, acquiescence, abstractedness, low energy state, and apathetic withdrawal.

The final overall model was statistically significant, χ^2 (25, N = 3913) = 504.03, $p < .001$. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3913) = 7.62, p =.47, further suggesting an adequate model fit. The overall fit of the model was fair, with pseudo-R^2 values of 0.121 (Cox and Snell R^2) and 0.197 (Nagelkerke R^2). Overall classification was very good, as the model predicted 83.5% of total responses correctly. Classification of passes (98.1% correctly classified) was excellent, while classification of failures (17.8% correctly classified) was an improvement over prior models.

Table 1.8 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals who were male (vs. female) and/or had prior mental health treatment (vs. none) were significantly more likely to fail out of the training program (ps < .001). Relative to individuals with an E-2 or E-3 rank, individuals with an E-4 rank (p = .005), but not those with an E-5 rank (p = .31), were more likely to fail out of the training program. Lower GT scores, poorer physical fitness scores, more prior disciplinary counseling statements, and higher AUDIT total scores were also associated with a greater likelihood of attrition (*p*s $\leq$ .001). Neither age (*p* = .013) nor military occupational specialty (*p* = .03) was associated with overall training outcome. Individuals reporting lower reasoning (*p* = .001) and interpersonal relations (*p* = .002) had a greater likelihood of program attrition. Similarly, although on a trend level, lower ratings of self-control were associated with an increased likelihood of attrition (*p* = .008). Those who reported higher social boldness and greater health concerns were also more likely to fail out of the training program (*p*s $\leq$ .001). Similarly, higher ratings of alienation/perceptual distortion were associated on a trend level with a greater likelihood of failure from the training program (*p* = .009). The following personality factors were not significantly associated with overall training program outcome: anxiety (*p* = .05), warmth (*p* = .07), liveliness (*p* = .06), sensitivity (*p* = .10), apprehension (*p* = .03), suicidal thinking (*p* = .06), thrill seeking (*p* = .08), and threat immunity (*p* = .03).

Model 9: Significant Predictors from Models 7 and 8

Predictors included sex, rank, GT score, disciplinary counseling statements, physical fitness scores, mental health treatment history, AUDIT total score, reasoning, social boldness, self-reliance, interpersonal relations, health concerns, alienation/perceptual distortion, and self-control. Interpersonal relations (step 2) was removed through the backward elimination procedure.

The final overall model was statistically significant, χ^2 (15, N = 3946) = 480.46, *p* < .001. The Hosmer-Lemeshow test was non-significant, χ^2 (8, N = 3946) = 7.19, *p* =.52, further suggesting an adequate model fit. The overall fit of the model was fair, with pseudo-R^2 values of 0.115 (Cox and Snell R^2) and 0.187 (Nagelkerke R^2). Overall classification was very good, as the model predicted 83.0% of total responses correctly. Classification of passes (98.0% correctly classified) was excellent, while classification of failures (15.9% correctly classified) was comparable to model 7.

Table 1.9 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals who were male (vs. female) and/or had prior mental health treatment (vs. none) were significantly more likely to fail out of the training program (*p*s < .001). Lower GT scores, poorer physical fitness scores, more prior disciplinary counseling statements, and higher AUDIT total scores were also associated with a greater likelihood of attrition (*p*s ≤ .001). Rank ($p = .02$) and combat deployments ($p = .04$) were not significantly associated with overall training program outcome. Individuals reporting lower self-control ($p = .004$) and reasoning ability ($p < .001$) had a greater likelihood of program attrition. Those who reported higher self-reliance ($p < .001$), more health concerns ($p < .001$), and greater alienation/perceptual distortion ($p = .003$) were also more likely to fail out of the training program. Greater social boldness was related to a higher likelihood of program attrition at a trend level ($p = .008$).

Specific Aim 1 – Logistic Regression Model Summary

All models tested were statistically significant with overall correct classification of cases above 80%. However, the classification of training failures in most of the models was poor, ranging from 0.0% to 17.8%. Model 8 had the highest classification of training failures (17.8%), and its predictors accounted for the most variance in overall training outcome (pseudo-R^2: 12.1 – 19.7%). Models 7 and 9 had the next highest failure classification accuracy (16.7 and 16.1%, respectively) and model fit as assessed by pseudo-R^2 (11.5-18.8 and 11.5-18.7%, respectively). All three models shared nine common predictor variables, including: sex, rank, mental health treatment history, GT scores, physical fitness scores, disciplinary counseling statements, AUDIT total scores, the 16PF reasoning scale, the 16PF health concerns scale, and the 16PF alienation/perceptual distortion scale. Models 7 and 9 also identified the 16PF self-reliance scales as significant predictors of training outcome, whereas Models 8 and 9 identified the 16PF self-control and social boldness scales as significant predictors. Additionally, Model 8 uniquely identified interpersonal relations as a significant predictor, and rank significantly predicted overall training outcome in Models 7 and 8 but not in Model 9. The significant predictors identified in this section were later used as indicators in the latent profile analyses to identify subgroups.

Successful Training Performance Database – Data Distribution

Following the data imputation procedures, the distribution of continuous variables was examined for normality, linearity, and homoscedasticity. The following variables had abnormal skew and kurtosis: AUDIT total score, 16PF impression management, 16PF infrequency, 16PF independence, and 16PF liveliness. Each of these variables had one impossible positive outlier, which was recoded to the highest range of the possible values. After recoding to possible values, the skew and kurtosis of these variables were satisfactory.

The Kolmogorov-Smirnov and Shapiro-Wilks tests were significant for all continuous variables (ps $< .05$), suggesting that these distributions statistically deviated from normality. However, these tests are highly sensitive with sample sizes greater than

1000, and therefore it is recommended to consider visual inspection of graphical plots alongside quantitative results (27). Visual inspection of the Q-Q plots and box plots indicated normal distributions for all continuous variables. As such, no additional transformations were applied to the data.

Based on these inspections of the data, it was determined that it is appropriate to continue with the multiple regressions as planned.

Specific Aim 2 – Successful Training Performance

Stepwise linear regression models with backward elimination were conducted to examine associations among personal, military, and psychological factors and training performance among successful program completers. The primary objective was to identify which factors formed the most parsimonious model to predict training performance. Predictor variables included age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness scores, mental health treatment history, combat deployment history, AUDIT total score, and the 16PF scales. Dependent variables included final GPA and instructor rating.

The following subsections present the findings from the regression models tested. These sections describe: predictors entered in the model; model significance; percentage of variation explained by only the independent variables that actually affect the dependent variable (adjusted R^2); and identified significant predictors of training outcome. The first subsection (Hypothesis 2a) contains the models with GPA as the dependent variable, and the second section (Hypothesis 2b) contains the models with Instructor Rating as the dependent variable.

Successful Training Performance Database – Sample Characteristics

Data from 3228 service members (99.3% male; age: M = 22.37, SD = 1.20 years) were included in analyses to examine predictors of training performance among successful program completers. The sample breakdown by military occupational specialty was: 24.2% (N = 780) air, 25.1% (N = 810) ground, and 50.3% (N = 1625) support. The sample breakdown by rank was: 10.6% (N = 342) E3, 80.6% (N = 2603) E4, and 8.8% (N = 283) E5. Service members had between 0 and 4 prior combat deployments (M = 0.56, SD = 0.74). In general, service members had an average GT score (M = 109.81, SD = 5.34) and above average physical fitness score (M = 253.47, SD = 11.94). They had an average of 1.52 (SD = 0.61) prior disciplinary counseling statements at the command. Approximately 2.9% (N = 93) of service members reported previously seeking treatment for a mental health problem. The average final GPA was 92.56 (SD = 1.56) out of a possible 100 percent, while the mean instructor rating was 3.47 (SD = 0.23) on a 5-point scale.

Hypothesis 2A: Predicting GPA

Model 2A.1: Participant Characteristics Only

Predictors included age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness scores, mental health treatment history, combat deployment history, and AUDIT total score. The dependent variable was final GPA. Military occupational specialty (step 2), sex (step 3), rank (step 4), and mental health treatment history (step 5) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(6, 3208) = 202.43, p <

.001, with adjusted R^2 = 0.273 for GPA. The Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.05 – 1.55; Tolerance = 0.65 – 0.96).

Table 2A.1 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores had higher GPAs (*p*s < .001). Those with fewer prior disciplinary counseling statements and combat deployments, as well as lower AUDIT total scores, also had higher GPAs (*p*s < .001). Age was not significantly related to GPA (p = .01).

Model 2A.2: 16PF Response Style Subscales Only

Predictors included impression management, infrequency, and acquiescence. Dependent variable was final GPA. Infrequency (step 2) and acquiescence (step 3) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(1, 3225) = 14.86, $p < .001$, but adjusted R^2 = 0.004 for GPA.

Table 2A.2 depicts the statistical contribution of the lone predictor variable retained in the final regression model. Individuals reporting higher impression management tendencies had higher GPAs ($p < .001$).

Model 2A.3: 16PF Global Factor Patterns Only

Predictors included: extraversion, anxiety, tough-mindedness, independence, and self-control. Tough-mindedness (step 2), independence (step 3), and extraversion (step 4) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(2, 3222) = 10.70, $p < .001$, but adjusted R^2 = 0.006 for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.17; Tolerance = 0.85).

Table 2A.3 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals reporting higher self-control had higher GPAs (p = .001). Anxiety was not significantly related to GPA (p = .09).

Model 2A.4: 16PF Primary Factor Profile Subscales Only

Predictors included: warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. The following predictors were sequentially removed in 14 steps through the backward elimination procedure: sensitivity, warmth, abstractedness, perfectionism, social boldness, emotional stability, vigilance, reasoning, openness to change, dominance, apprehension, privateness, and self-reliance.

The final overall model was statistically significant, F(3, 3222) = 10.90, $p < .001$, but adjusted R^2 = 0.009 for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.03-1.19; Tolerance = 0.85-0.97).

Table 2A.4 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals reporting less tension had higher GPAs (p = .003). Liveliness and rule-consciousness were not significantly related to GPA (ps = .01).

Model 2A.5: 16PF Protective Services Dimensions Only

Predictors included: emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Intellectual efficiency (step 2) and interpersonal relations (step 3) were removed through the backward elimination procedure.

The final overall model was statistically significant, $F(3, 3224) = 9.21$, $p < .001$, but adjusted $R^2 = 0.005$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.68; Tolerance = 0.59).

Table 2A.5 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Neither emotional adjustment nor integrity/control was significantly associated with GPA (*p*s = .07).

Model 2A.6: 16-PF Pathology-Oriented Scales Only

Predictors included psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity. The following predictors were sequentially removed in 11 steps through the backward elimination procedure: apathetic withdrawal, low energy state, self-reproach, anxious depression, psychological inadequacy, threat immunity, health concerns, obsessional thinking, paranoid ideation, and suicidal thinking.

The final overall model was statistically significant, $F(2, 3222) = 7.57$, $p = .001$, but adjusted $R^2 = 0.004$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.09; Tolerance = 0.91).

Table 2A.6 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Neither alienation/perceptual distortion ($p = .01$)

nor thrill seeking (p = .03) was significantly associated with GPA.

Model 2A.7: Significant Predictors from Models 1-6

Predictors included: GT score, disciplinary counseling statements, physical fitness scores, total combat deployments, AUDIT total score, impression management, self-control, and tension. The following variables were also included (despite not being significant based on the *a priori* α level of .005) to be thorough: age, liveliness, rule-consciousness, alienation/perceptual distortions, and thrill seeking. Thrill seeking (step 2), impression management (step 3), rule-consciousness (step 4), AUDIT total score (step 5), and alienation/perceptual distortions (step 6) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(8, 3217) = 169.06, $p <$ .001, and adjusted R^2 = 0.294 for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.05-1.52; Tolerance = 0.65-0.96).

Table 2A.7 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (ps < .001). Age was not significantly related to GPA (p = .05). Higher self-control (p < .001) and lower liveliness (p = .003) scores were associated with higher GPAs. Self-reported tension was not significantly related to GPA (p = .02).

Model 2A.8: All Potential Personal, Military, and Psychological Predictors

The final regression model was achieved in 38 steps through the backward elimination procedure. The following participant characteristic predictors were removed:

rank, sex, and mental health treatment history. All of the 16PF scales were removed with the exception of: acquiescence, emotional stability, liveliness, apprehension, emotional adjustment, intellectual efficiency, and self-reproach.

The final overall model was statistically significant, F(13, 3191) = 104.99, $p < .001$, and adjusted $R^2 = 0.297$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.05-5.05; Tolerance = 0.30-0.95).

Table 2A.8 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (*p*s < .001). Neither age ($p = .04$) nor AUDIT total score ($p = .07$) was significantly related to GPA. Reports of lower emotional stability ($p < .001$), liveliness ($p < .001$), and intellectual efficiency ($p = .002$) were associated with a higher GPA. Acquiescence ($p = .07$) and self-reproach ($p = .06$) were not significantly related to GPA.

Model 2A.9: Significant Predictors from Models 7 and 8

Predictors included: GT score, disciplinary counseling statements, physical fitness scores, combat deployments, self-control, emotional stability, liveliness, apprehension, emotional adjustment, and intellectual efficiency. Self-control (step 2) was removed through the backward elimination procedure.

The final overall model was statistically significant, F(9, 3218) = 150.43, $p < .001$, and adjusted $R^2 = 0.294$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.05-4.78; Tolerance = 0.21-0.96).

Table 2A.9 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (*p*s < .001). Reports of lower emotional stability ($p < .001$), liveliness ($p < .001$), and intellectual efficiency ($p = .002$) were associated with a higher GPA. Individuals reporting higher apprehension and emotional adjustment also had higher GPAs (*p*s < .001).

Hypothesis 2A (Grade Point Average) Model Summary

All models tested were statistically significant, indicating that the models were able to predict GPA from the included variables. However, the impact of psychological independent variables on the dependent variables was minimal. Models 7A, 8A, and 9A had the largest adjusted R^2 values (0.294, 0.297, and 0.294, respectively). The five common predictor variables across these three models included: GT score, combat deployment history, prior disciplinary counseling statements, physical fitness score, and the 16PF liveliness scale. Overall, psychological factors entered into the regression models appeared to have a minimal role in predicting program GPA.

Hypothesis 2B: Predicting Instructor Ratings

Model 2B.1: Participant Characteristics Only

Predictors included: age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, combat deployment history, and AUDIT total score. The dependent variable was instructor rating. Mental health treatment (step 2) and military occupational specialty (step 3) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(8, 3206) = 149.93, $p <$ .001, and adjusted R^2 = 0.270 for instructor rating. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.06-2.60; Tolerance = 0.38-0.94).

Table 2B.1 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (*p*s ≤ .001). The AUDIT total score was not significantly associated with instructor rating ($p = .05$).

Model 2B.2: 16PF Response Style Subscales Only

Predictors: included impression management, infrequency, and acquiescence. Infrequency (step 2), acquiescence (step 3), and impression management (step 4) were all removed through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.3: 16PF Global Factor Patterns Only

Predictors: included extraversion, anxiety, tough-mindedness, independence, and self-control. Extraversion (step 2), independence (step 3), self-control (step 4), anxiety (step 5), and tough-mindedness (step 6) were all removed through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.4: 16PF Primary Factor Profile Subscales Only

Predictors: included warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. The following predictors were sequentially removed in 16 steps through the backward elimination procedure: privateness, perfectionism, tension, self-reliance, vigilance, warmth, social boldness, emotional stability, liveliness, rule-consciousness, abstractedness, apprehension, reasoning, sensitivity, and dominance.

The final overall model was statistically significant, F(1, 3224) = 6.25, $p = .012$, but adjusted $R^2 = 0.002$ for instructor rating.

Table 2B.4 depicts the statistical contribution of the lone predictor variable retained in the final regression model. However, openness to change was not significantly associated with instructor rating ($p = .01$).

Model 2B.5: 16PF Protective Services Dimensions Only

Predictors included emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Integrity/control (step 2), emotional adjustment (step 3), interpersonal relations (step 4), and intellectual efficiency (step 5) were all removed through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.6: 16PF Pathology-Oriented Scales Only

Predictors included psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity. All of the predictors were sequentially removed in 13 steps through the backward elimination procedure in the following order: obsessional thinking, alienation/perceptual distortion, anxious depression, health concerns, threat immunity, suicidal thinking, low energy state, paranoid ideation, self-reproach, thrill seeking, apathetic withdrawal, and psychological inadequacy. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.7: Significant Predictors from Models 1-6

Predictors included rank, GT score, age, sex, disciplinary counselings, physical fitness, combat deployments, and openness to change. Although openness to change was not significant based on the *a priori* α level, it was included in this model because it was the only potential psychological predictor to emerge. Openness to change was removed through the backward elimination procedure.

The final overall model was statistically significant, F(1, 3220) = 170.59, $p <$ .001, and adjusted R^2 = 0.269 for instructor rating. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.06-2.54; Tolerance = 0.39-0.94).

Table 2B.7 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (*p*s < .001).

Model 2B.8: All Potential Predictors

The stepwise multiple regression used 41 steps to reach the final model. The following participant characteristic predictors were removed: military occupational specialty and mental health treatment history. All of the 16PF scales were removed through the backward elimination, with the exception of tough-mindedness and openness to change.

The final overall model was statistically significant, F(10, 3194) = 119.56, $p < .001$, and adjusted $R^2 = 0.270$ for instructor rating. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.07-2.42; Tolerance = 0.38-0.94).

Table 2B.8 depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (ps $\leq .001$). AUDIT total score ($p = .05$), tough-mindedness ($p = .08$), and openness to change ($p = .03$) were not significantly associated with instructor ratings.

Hypothesis 2B (Instructor Rating) Model Summary

Only models 1B, 4B, 7B, and 8B were statistically significant, indicating that they were able to predict instructor rating. However, the percentage of variation explained by only the independent variables that actually affect the dependent variable for model 4B was negligible (adjusted $R^2 = 0.002$). Models 1b, 7b, and 8b were functionally identical, as they included the same variables and had the largest percentage of variation explained by only the included predictors related to instructor rating (adjusted $R^2 = 0.270$, 0.269, and 0.270, respectively). Model 9B (significant predictors from Models 7B & 8B) was not conducted, as the predictors for those models were identical. These models indicated that instructor rating is best accounted for by the following personal and military characteristics: age, sex, GT score, combat deployment history, prior disciplinary counseling statements, and physical fitness score. Psychological characteristics did not play a significant role in predicting instructor ratings among successful completers of this training program.

Specific Aim 2 – Successful Training Performance – Cross-Validation Models

Stepwise linear regression models with backward elimination were conducted to examine associations among personal, military, and psychological factors and training performance among successful program completers. The primary objective was to identify which factors formed the most parsimonious model to predict training performance. Predictor variables included age, sex, rank, military occupational specialty, GT score, prior disciplinary counseling statements, physical fitness scores, mental treatment history, combat deployment history, AUDIT total score, and the 16-PF subscales. Dependent variables included final GPA and instructor rating.

Cross-validation was determined to be an appropriate approach to evaluate regression model fit and generalizability to a potential, future independent sample. 80% of the full regression model dataset (N = 2589; “training sample”) was randomly selected using a random number generator to conduct all linear regression models. The other 20% (N = 639) of the sample was used as the cross-validation sample.

The following subsections present the findings from the individual models tested and are structured identically to the previous multiple regression sections. These sections include: predictors entered in the model; model significance; percentage of variation explained by only the independent variables that actually affect the dependent variable (adjusted R^2); and identified significant predictors of training outcome. The first subsection (Hypothesis 2a) contains the models with GPA as the dependent variable, and the second section (Hypothesis 2b) contains the models with Instructor Rating as the dependent variable.

Successful Training Performance CVM Database – Sample Characteristics

Data from 2589 service members (98.4% male; age: M = 22.37, SD = 1.20 years) were included in analyses of the training sample to examine predictors of training performance among successful program completers. The sample breakdown by military occupational specialty was: 23.8% (N = 616) air, 25.0% (N = 646) ground, and 50.9% (N = 1317) support. The sample breakdown by rank was: 10.9% (N = 283) E3, 80.2% (N = 2076) E4, and 8.9% (N = 230) E5. Service members had between 0 and 4 prior combat deployments (M = 0.56, SD = 0.74). In general, service members had an average GT score (M = 109.81, SD = 5.31) and above average physical fitness score (M = 253.54, SD = 11.84). They had an average of 1.53 (SD = 0.61) prior disciplinary counseling statements at the command. Approximately 3.1% (N = 80) of service members reported previously seeking treatment for a mental health problem. The average final GPA was 92.57 (SD = 1.56) out of a possible 100 percent, while the mean instructor rating was 3.47 (SD = 0.23) on a 5-point scale.

Hypothesis 2a: Predicting GPA (Cross Validation Models)

Model 2A.1 (CVM): Participant Characteristics Only

Predictors included: age, sex, rank, military occupational specialty, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, combat deployment history, and AUDIT total score. Military occupational specialty (step 2), rank (step 3), age (step 4), and sex (step 5) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(6, 2572) = 172.82, $p <$ .001, and with adjusted R^2 = 0.286 for GPA. The Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.01 – 1.19; Tolerance = 0.84 – 0.99).

Table 2A.1 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores had higher GPAs (*p*s < .001). Those with fewer prior disciplinary counseling statements and combat deployments also had higher GPAs (*p*s < .001). Higher AUDIT total scores were associated with lower GPAs at a trend level (p = .006) Mental health treatment history was not significantly related to GPA (p = .08).

Model 2A.2 (CVM): 16PF Response Style Subscales Only

Predictors included impression management, infrequency, and acquiescence. Acquiescence (step 2) and infrequency (step 3) were removed through the backward elimination procedure. The final overall model was statistically significant, F(1, 2586) = 8.48, p = .004, but with adjusted R^2 = 0.003 for GPA. **Table 2A.2 (CVM)** depicts the statistical contribution of the lone predictor variable retained in the final regression model. Individuals reporting higher impression management tendencies had higher GPAs (p = .004).

Model 2A.3 (CVM): 16PF Global Factor Patterns Only

Predictors included extraversion, anxiety, tough-mindedness, independence, and self-control. Tough-mindedness (step 2), independence (step 3), extraversion (step 4), and anxiety (step 5) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(1, 2584) = 12.75, $p < .001$, but with adjusted $R^2 = 0.005$ for GPA. **Table 2A.3 (CVM)** depicts the statistical contribution of the lone predictor variable retained in the final regression model. Individuals reporting higher self-control had higher GPAs ($p = .001$).

Model 2A.4 (CVM): 16PF Primary Factor Profile Subscales Only

Predictors included: warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. The following predictors were sequentially removed in 14 steps through the backward elimination procedure: perfectionism, reasoning, abstractedness, social boldness, emotional stability, sensitivity, vigilance, openness to change, warmth, self-reliance, apprehension, liveliness, and dominance.

The final overall model was statistically significant, F(3, 2583) = 7.19, $p < .001$, but with adjusted $R^2 = 0.007$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.03-1.17; Tolerance = 0.86-0.97).

Table 2A.4 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals reporting less tension had higher GPAs on a trend level ($p = .007$). Liveliness ($p = .01$) and rule-consciousness ($p = .08$) were not significantly related to GPA.

Model 2A.5 (CVM): 16PF Protective Services Dimensions Only

Predictors included: emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Intellectual efficiency (step 2), interpersonal relations (step 3), and emotional adjustment (step 4) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(1, 2586) = 11.06, p = .001, but with adjusted R^2 = 0.004 for GPA. **Table 2A.5 (CVM)** depicts the statistical contribution of the lone predictor variable retained in the final regression model. Individuals reporting higher integrity/control scores had significantly higher GPAs (p = .001).

Model 2A.6 (CVM): 16PF Pathology-Oriented Scales Only

Predictors included: psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity. The following predictors were sequentially removed in 12 steps through the backward elimination procedure: low energy state, self-reproach, psychological inadequacy, health concerns, apathetic withdrawal, anxious depression, paranoid ideation, thrill seeking, obsessional thinking, suicidal thinking, and threat immunity.

The final overall model was statistically significant, F(1, 2585) = 7.20, p = .007, but with adjusted R^2 = 0.002 for GPA.

Table 2A.6 (CVM) depicts the statistical contribution of the lone predictor variable retained in the final regression model. Higher Alienation/perceptual distortion scores were associated with lower GPAs on a trend level (p = .007).

Model 2A.7 (CVM): Significant Predictors from Models 1-6

Predictors included: GT score, prior disciplinary counseling statements, physical fitness scores, combat deployment history, impression management, self-control, and integrity/control. The following variables were also included (despite not being significant based on the *a priori* α level) to be thorough: AUDIT total score and alienation/perceptual distortion. Impression management (step 2), integrity/control (step 3), and AUDIT total score (step 4) were removed through the backward elimination procedure.

The final overall model was statistically significant, $F(6, 2580) = 188.70$, $p < .001$, and with adjusted $R^2 = 0.303$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.05-1.18; Tolerance = 0.84-0.95).

Table 2A.7 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (ps < .001). Higher self-control scores were associated with higher GPAs ($p < .001$). Self-reported alienation/perceptual distortion was not significantly related to GPA ($p = .10$).

Model 2A.8 (CVM): All Potential Predictors

The final model was reached in 38 steps through the backward elimination procedure. The following participant characteristic predictors were removed: age, sex, rank, military occupational specialty, and AUDIT total score. All of the 16PF scales were removed through the backward elimination procedure, with the exception of anxiety, liveliness, vigilance, tension, emotional adjustment, intellectual efficiency, paranoid ideation, and alienation/perceptual distortion.

The final overall model was statistically significant, F(13, 2556) = 88.87, $p < .001$, and with adjusted $R^2 = 0.308$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.01-9.00; Tolerance = 0.11-0.99).

Table 2A.8 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (ps < .001). Mental health treatment history was not associated with final GPA ($p = .07$). Reports of higher liveliness, tension, and intellectual efficiency were associated with lower GPAs (ps ≤ .001). By contrast, reports of higher anxiety and emotional adjustment were associated with higher GPAs (ps < .001). Vigilance ($p = .01$), paranoid ideation ($p = .09$), and alienation/perceptual distortion ($p = .05$) scores were not significantly related to GPA.

Model 2A.9 (CVM): Significant Predictors from Models 7 and 8

Predictors included: GT score, disciplinary counseling statements, physical fitness scores, combat deployment history, and the 16PF scales of self-control, anxiety, liveliness, tension, emotional adjustment, and intellectual efficiency. Self-control (step 2) was removed through the backward elimination procedure.

The final overall model was statistically significant, F(9, 2577) = 126.52, $p < .001$, and with adjusted $R^2 = 0.304$ for GPA. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.04-5.58; Tolerance = 0.18-0.95).

Table 2A.9 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Individuals with higher GT scores and physical fitness scores, as well as fewer prior counseling statements and combat deployments, had higher GPAs (ps < .001). Reports of higher anxiety and emotional adjustment, but lower liveliness and tension, were associated with higher GPAs (ps ≤ .002). Reports of higher intellectual efficiency was associated with a lower GPA on a trend level (p = .006).

Hypothesis 2A (Grade Point Average) CVM - Model Summary

All models tested were statistically significant, indicating that the models were able to predict GPA in successful program completers from the included variables. However, the percentage of variation explained by only the psychological independent variables that actually affect the dependent variable was negligible. CVM models 7A, 8A, and 9A had the percentage of variation explained by the independent variables that affect the dependent variable (adjusted R^2 = 0.303, 0.308, and 0.304, respectively). The four common predictor variables across these three models included: GT score, combat deployment history, prior disciplinary counseling statements, and physical fitness score. Models 8A and 9A shared the following psychological factors in common for predicting GPA: anxiety, liveliness, emotional adjustment, and tension. By contrast, model 7A uniquely identified the 16PF self-control scale as a significant predictor of GPA. Overall, psychological factors appeared to have a negligible role in predicting program GPA in successful program completers, accounting for a small increase in adjusted-R^2 beyond intelligence and military factors.

Hypothesis 2B: Predicting Instructor Ratings (CVM)

Model 2B.1 (CVM): Participant Characteristics Only

Predictors included: age, sex, rank, military occupational specialty, GT score, prior disciplinary counseling statements, physical fitness score, mental health treatment history, combat deployment history, and AUDIT total score. Mental health treatment history (step 2), military occupational specialty (step 3), and AUDIT total score (step 4) were removed through the backward elimination procedure.

The final overall model was statistically significant, F(7, 2571) = 138.82, $p < .001$, and with adjusted $R^2 = 0.272$ for instructor rating. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.06-2.60; Tolerance = 0.38-0.94).

Table 2B.1 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (ps $\leq .001$).

Model 2B.2 (CVM): 16PF Response Style Subscales Only

Predictors included impression management, infrequency, and acquiescence. Infrequency (step 2) and acquiescence (step 3) were removed through the backward elimination procedure. The final overall model was not statistically significant, F(1, 2586) = 3.136, $p = .077$, with adjusted $R^2 = 0.001$ for instructor rating.

Table 2B.2 (CVM) depicts the statistical contribution of the lone predictor variable retained in the final regression model. Impression management scores were not significantly related to instructor ratings (p = .08).

Model 2B.3 (CVM): 16PF Global Factor Patterns Only

Predictors included: extraversion, anxiety, tough-mindedness, independence, and self-control. Self-control (step 2), independence (step 3), extraversion (step 4), anxiety (step 5), and tough-mindedness (step 6) were all removed through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.4 (CVM): 16PF Primary Factor Profile Subscales Only

Predictors included: warmth, reasoning, emotional stability, dominance, liveliness, rule-consciousness, social boldness, sensitivity, vigilance, abstractedness, privateness, apprehension, openness to change, self-reliance, perfectionism, and tension. The following predictors were sequentially removed in 14 steps through the backward elimination procedure: vigilance, self-reliance, emotional stability, social boldness, privateness, rule-consciousness, tension, sensitivity, warmth, perfectionism, abstractedness, liveliness, and apprehension.

The final overall model was statistically significant, F(3, 2583) = 3.16, p = .02, with adjusted R^2 = 0.002 for instructor rating.. **Table 2B.4 (CVM)** depicts the relative statistical contribution of each predictor variable retained in the final regression model. Reasoning, (p = .09), dominance (p = .06), and openness to change (p = .03) were not significantly associated with instructor ratings.

Model 2B.5 (CVM): 16PF Protective Services Dimensions Only

Predictors included: emotional adjustment, integrity/control, intellectual efficiency, and interpersonal relations. Integrity/control (step 2), emotional adjustment (step 3), intellectual efficiency (step 4), and interpersonal relations (step 5) were removed through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.6 (CVM): 16PF Pathology-Oriented Scales Only

Predictors included psychological inadequacy, health concerns, suicidal thinking, anxious depression, low energy state, self-reproach, apathetic withdrawal, paranoid ideation, obsessional thinking, alienation/perceptual distortion, thrill seeking, and threat immunity. All predictors were sequentially removed in 13 steps through the backward elimination procedure. As such, there were no summary statistics because the model was a poor fit for the data.

Model 2B.7 (CVM): Significant Predictors from Models 1-6

Predictors included: age, sex, rank, GT score, disciplinary counseling statements, physical fitness score, combat deployment history, and openness to change. Although openness to change was not significant based on the *a priori* α level, this variable was included in this model because it was the only potential psychological predictor to emerge. Openness to change was removed through the backward elimination procedure.

The final overall model was statistically significant, $F(7, 2581) = 139.03$, $p < .001$, and with adjusted $R^2 = 0.272$ for instructor rating. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.07-2.60; Tolerance = 0.39-0.94).

Table 2B.7 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (*p*s < .001).

Model 2B.8 (CVM): All Potential Predictors

The final model was reached in 44 steps. The following participant characteristic predictors were removed: military occupational specialty, mental health treatment history, and AUDIT total score. All of the 16-PF scales were removed through the backward elimination procedure.

The final overall model was statistically significant, F(7, 2562) = 137.55, $p < .001$, and with adjusted R^2 = 0.271 for instructor rating.. Variance Inflation Factor (VIF) and Tolerance values suggested that the multicollinearity assumption was not violated (VIF = 1.07-2.60; Tolerance = 0.39-0.93).

Table 2B.8 (CVM) depicts the relative statistical contribution of each predictor variable retained in the final regression model. Males were more likely to have lower instructor ratings than females ($p < .001$). Individuals who were older, had higher GT scores, more prior combat deployments, fewer prior disciplinary counseling statements, and better physical fitness scores also had higher instructor ratings (*p*s < .001).

Hypothesis 2B (Instructor Rating) CVM - Model Summary

Only models 1B, 2B, 4B, 7B, and 8B were statistically significant, indicating that they were able to predict instructor rating. However, the adjusted R^2 for models 2B and 4B were negligible (0.001 and 0.002, respectively). Models 1B, 7B, and 8B were functionally identical, as they included the same variables and had nearly identical adjusted R^2 values (0.272, 0.272, and 0.271, respectively). Model 9B (significant predictors from Models 7B & 8B) was not conducted, as the predictors for those models were identical.

These models indicated that a significant portion of instructor rating could be accounted for by the following personal and military characteristics: age, sex, GT score, combat deployment history, prior disciplinary counseling statements, and physical fitness score. Psychological characteristics did not play a significant role in predicting instructor ratings among successful completers of this training program.

Evaluating the Cross-Validation Models

The predictive regression models identified using the training (80%) sample for GPA and Instructor Ratings are detailed below. These regression equations were used to generate predicted scores for all participants. The association between the observed and predicted scores for the training (80%) and cross-validation (20%) samples were then evaluated. More generalizable models will have higher correlations when examining associations between the training and cross-validation samples. All analyses were conducted using IBM SPSS version 22 (SPSS, Inc., Chicago, IL).

Successful Training Performance CVM Database – Cross-Validation Sample Characteristics

Data from 639 service members (98.6% male; age: M = 22.39, SD = 1.17 years) were included in the cross-validation sample. The sample breakdown by military occupational specialty was: 25.7% (N = 164) air, 25.7% (N = 164) ground, and 48.2% (N = 308) support. The sample breakdown by rank was: 9.2% (N = 59) E3, 82.5% (N = 527) E4, and 8.3% (N = 53) E5. Service members had between 0 and 4 prior combat deployments (M = 0.54, SD = 0.71). In general, service members had an average GT score (M = 109.82, SD = 5.49) and above average physical fitness score (M = 253.54, SD = 11.84). They had an average of 1.50 (SD = 0.61) prior disciplinary counseling statements at the command. Approximately 2.0% (N = 13) of service members reported previously seeking treatment for a mental health problem. The average final GPA was 92.51 (SD = 1.57) out of a possible 100 percent, while the mean instructor rating was 3.48 (SD = 0.23) on a 5-point scale. There were no significant differences between the training and cross-validation samples on these demographic and military characteristics (*p*s > .05).

Hypothesis 2A: Final GPA CVM

The best-fitting model generated using the training sample (80% of the original sample) was deemed to be the final step of "Model 9B." The regression equation generated from this model was as follows:

Predicted GPA = 80.96 + 0.087(GT Score) - 0.904(Disciplinary Counselings) + 0.11(Physical Fitness) - 0.318(Combat Deployments) + 0.121(Anxiety) – 0.062(Liveliness) – 0.083(Tension) + 0.183(Emotional Adjustment) – 0.050(Intellectual Efficiency).

This formula was used to calculate predicted values of GPA for each participant in both the training and cross-validation samples. Bivariate Pearson's correlations were conducted to examine the association between the actual and predicted values of GPAs. There was a large, positive association between actual and predicted GPAs among the training sample. $r = 0.554$, $p < .001$. Similarly, there was a large, positive association between actual and predicted instructor ratings among the cross-validation sample, $r = 0.498$, $p < .001$. R^2 was 0.307 and 0.248 for the training and cross-validation samples, respectively.

The most parsimonious regression equation for predicting GPA in successful program completers appears to account for a comparable proportion of the variance in both the training and cross-validation samples. This finding suggests that the identified regression equation is generalizable and is not over-fitted to the training sample. However, some caution may be warranted regarding the generalizability of the data to future independent samples given the difference in R^2 values between the training and cross-validation samples.

Hypothesis 2B: Instructor Rating CVM

The most parsimonious, best-fitting model generated with the training sample (80% of the original sample) was deemed to be the final step of "Model 7A." The regression equation generated from this model was as follows:

Predicted Instructor Rating = 1.159 + 0.127(Rank) + 0.003(GT Score) + 0.03(Age) – 0.174(Sex) – 0.087(Disciplinary Counselings) + 0.004(Physical Fitness) + 0.037(Combat Deployments).

This formula was used to calculate predicted values for instructor ratings for each participant in both the training and cross-validation samples. Bivariate Pearson's correlations were conducted to examine the association between the actual and predicted values of instructor ratings. There was a large, positive association between actual and predicted instructor ratings among the training sample, $r = 0.523$, $p < .001$. Similarly, there was a large, positive association between actual and predicted instructor ratings among the cross-validation sample, $r = 0.506$, $p < .001$. R^2 was 0.274 and 0.256 for the training and cross-validation samples, respectively. As such, the identified regression equation appears to have acceptable generalizability and is not over-fitted to the training sample.

Comparison of Multiple Regression Models from the Original and Training Samples

Specific Aim 1 – Predicting GPA: There was no difference in the statistical significance of the regression models conducted using the original (100%) and training (80%) samples (*i.e.*, all were statistically significant), and the adjusted R^2 values for each model were comparable. Further, models 2A, 3A, and 6A were identical in the significant predictors identified across both samples. The majority of the models were extremely similar, and the vast majority of the variables in all models were identical.

However, five psychological variables that demonstrated differences between the most-parsimonious regression models generated using the full original sample (100% of the original sample) and training sample (80% of the original sample): emotional stability, apprehension, anxiety, tension, and intellectual efficiency. Emotional stability, apprehension, and intellectual efficiency were significant predictors of GPA in the original sample, but not in the training sample. By contrast, anxiety and tension were identified as significant predictors of GPA in the training sample, but not in the original sample.

Overall, the general convergence of models indicates that the variables identified in the preceding analyses were the best statistical predictors of GPA.

Specific Aim 2 – Predicting Instructor Ratings: No differences were found for significant predictors between the regression models conducted using the original and training samples. The difference between adjusted R^2 in regression models using the original sample and training sample ranged from .001 to .003 with models derived from the training sample having the higher values. Overall, these minimal differences indicate that the models were extremely similar in predicting instructor rating. This convergence of models suggests that the variables identified in these analyses were the best statistical predictors of Instructor Rating.

LATENT PROFILE ANALYSIS (LPA) SUMMARY: SPECIFIC AIM 3

The purpose of Specific Aim 3 was to identify, characterize, and validate subgroups within the sample population for whom different combinations of variables more accurately predict training attrition risk. This aim was exploratory in design. The initial model was developed based on the significant predictors of overall training outcome (success versus failure) identified in Models 7, 8, and 9 of **Specific Aim 1**. These indicators included: sex, rank, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, AUDIT total score, self-control, health concerns, reasoning, self-reliance, alienation/perceptual distortion, social boldness, and interpersonal relations. Age, combat deployments, and emotional stability also were considered as potential, secondary indicators because they significantly predicted GPA and/or instructor rating among successful training completers in **Specific Aim 2**.

To aid in interpretation of the latent subgroups, all indicator variables were coded categorically. The cut points for certain indicator variables were based on data distribution (e.g., disciplinary counseling statements, combat deployments), and other variables were divided based on inherent dividing points (e.g., AUDIT clinical cutoff, 16PF scale divisions in the manual). The indicator variables were categorically coded as follows:

- Sex: Male, Female
- Rank: E2/E3, E4, E5
- GT score: Average (80-119), Above Average (120+)
- Disciplinary counseling statements: 0, 1, 2+
- Physical fitness score: High (225+), Lower (100-224)
- Combat deployments: 0, 1, 2+
- Mental health treatment history: Yes, No
- AUDIT total score: Minimal (0-6), Moderate Drinking (7+)
- Age: Young (21 and below), Middle (21.1-23), Older (23.1 and above)
- 16PF scales: Low (0-3), Average (4-7), High (8-10)

LPA Model 1

Indicator variables included in the first latent profile analysis were the **shared** significant predictors from Models 7, 8 & 9 of Specific Aim 1. Specifically, these indicator variables were: sex, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, AUDIT total score, reasoning, health concerns, and alienation/perceptual distortion. Inspection of the fit indices suggested that a three-subgroup solution yielded the best fit for the data (see **Table 3.1** for a summary of fit indices).

<u>Subgroup 1</u> ("Psychological Concerns") was comprised of individuals with maladaptive psychosocial functioning (3.0%, n = 119). Relative to other subgroups, this psychologically concerning subgroup had a greater proportion of individuals reporting prior mental health treatment (22.0%; see **Figure 1.5**), moderate to concerning alcohol use patterns (18.8%; see **Figure 1.6**), and high health concerns (26.8%; see **Figure 1.8**).

This vast majority of individuals in this subgroup reported high alienation/perceptual distortion (62.8%; see **Figure 1.7**). Additionally, this psychologically concerning subgroup reported the highest percentage of individuals reporting low reasoning ability (40.2%; see **Figure 1.9**) in comparison to other subgroups.

Subgroup 2 ("Control/Average") was primarily average across all indicators, thereby serving as a reference group with which to compare the other subgroups (75.0%, n = 2962). The vast majority had average intellectual functioning (98.7%), no prior mental health treatment (87.3%), and minimal alcohol use (93.7%). The vast majority of the subgroup had average health concerns (100.0%), reasoning capabilities (79.1%), and alienation/perceptual distortion (75.9%).

Subgroup 3 ("High IQ") represented a subgroup with high intellectual functioning (21.9%, n = 866). Notably, the vast majority of individuals from this highly intelligent subgroup had above average intelligence scores (77.8%; see **Figure 1.3**). Additionally, this intelligent subgroup had approximately 25% of individuals reporting high reasoning abilities, which starkly contrasted with other subgroups reporting a negligible percentage of high reasoning abilities (see **Figure 1.9**).

Graphical plots of the proportion of endorses responses for each item, across each subgroup, were visually inspected (see **Figures 1.1-1.9**). Subgroups did not markedly differ with respect to sex (**Figure 1.1**), or mental health treatment history (**Figure 1.5**). Across all three subgroups, the vast majority of individuals were males and had no prior mental health treatment history. As such, latent profile analysis was conducted without these variables in subsequent models.

LPA Model 2

A revised LPA model was conducted by dropping the indicator variables from LPA Model 1 that did not contribute to the apparent differentiation of the subgroups (i.e., sex, mental health treatment history) and adding the other significant variables identified across Models 7, 8, & 9 from Specific Aim 1. As such, indicator variables for this model included: rank, GT score, disciplinary counseling statements, physical fitness score, AUDIT total score, self-control, health concerns, reasoning, self-reliance, alienation/perceptual distortion, social boldness, and interpersonal relations. Inspection of the fit indices suggested that a six-subgroup solution yielded the best fit for the data (see **Table 3.2** for a summary of fit indices).

<u>Subgroup 1</u> ("Introverts") represented a subgroup of introverted individuals (8.1%, n = 318). Notably, all of the individuals from this introverted subgroup reported low interpersonal relations, which contrasted starkly with all other subgroups (see **Figure 2.10**). A much higher proportion of individuals in this subgroup also reported high self-reliance (37.8%; see **Figure 2.8**) and low social boldness (44.4%; see **Figure 2.9**) than other subgroups, with the exception of the psychologically concerning subgroup 5.

<u>Subgroup 2</u> ("High IQ") represented a subgroup with high intellectual functioning (12.9%, n = 510). Notably, the vast majority of individuals from this highly intelligent subgroup had above average intelligence scores (99.6%; see **Figure 2.2**). Additionally, this intelligent subgroup had 32.5% of individuals reporting high reasoning abilities, which was greater than all other subgroups, and was the only subgroup without individuals with low reasoning (see **Figure 2.6**).

<u>Subgroup 3</u> ("Self-Sufficient") was comprised of independent, self-controlled

participants (19.6%, n = 775). This subgroup had, by far, the highest percentage of individuals reporting high self-control (52.9%); see **Figure 2.7**), which appeared to be the key distinguishing indicator. In addition, this subgroup also had the highest percentage of individuals reporting low alienation/perceptual distortion in comparison to other subgroups (47.2%; see **Figure 2.4**). Individuals from this self-sufficient subgroup also reported a higher proportion of those with low self-reliance (22.9%; see **Figure 2.8**) than all subgroups except the extroverted subgroup 6. This self-sufficient subgroup had the lowest percentage of moderate to concerning alcohol use patterns (1.0%; see **Figure 2.3**).

Subgroup 4 ("Control/Average") was generally average across all indicators, thereby serving as a control group (39.9%, n = 1573). This subgroup was comprised of 18.8% E5, 57.1% E4, and 24.1% E2/E3 individuals (see **Figure 2.1**). The vast majority had minimal alcohol use (90.7%) and average health concerns (99.6%). Additionally, most individuals from this subgroup reported average levels of alienation/perceptual distortion (88.7%), self-control (87.0%), reasoning (72.9%), self-reliance (91.6%), social boldness (88.9%), and interpersonal relations (99.7%). Nearly all of the individuals from this subgroup had average intellectual functioning (98.5%); the negligible amount of individuals with above average intellectual functioning (1.5%) represented the lowest proportion in comparison to other subgroups (see **Figure 2.2**).

Subgroup 5 ("Psychologically Concerning") was comprised of individuals with maladaptive psychosocial functioning (1.8%, n = 72). Relative to other subgroups, this psychologically concerning subgroup had a greater proportion of individuals reporting moderate to concerning alcohol use patterns (16.8%; see **Figure 2.3**), multiple disciplinary counseling statements (59.9%; see **Figure 2.12**), and high health concerns

(42.0%; see **Figure 2.5**). Additionally, this psychologically concerning subgroup reported the highest percentage of individuals reporting high alienation/perceptual distortion in comparison to other subgroups that had much fewer individuals reporting this trait (77.5%; see **Figure 2.4**). Low reasoning ability (33.5%; see **Figure 2.6**), low self-control (30.9%; see **Figure 2.**7), high self-reliance (45.2%; see **Figure 2.8**), high social boldness (42.8%; see **Figure 2.9**), and low interpersonal relations scores (52.8%; see **Figure 2.10**) were also more likely to be endorsed by individuals in this psychologically concerning subgroup relative to all other subgroups except for one. This subgroup also had a higher proportion of individuals with E2 or E3 ranks (see **Figure 2.1**) and a lower proportion of individuals with a first class physical fitness score (see **Figure 2.11**) in comparison to other subgroups.

Subgroup 6 ("Extrovert") represented a sociable subgroup of individuals (17.7%, n = 699). All of the participants from this extroverted subgroup reported high interpersonal relations scores (100.0%; see **Figure 2.10**), which represented a stark contrast from other subgroups. This subgroup also had the highest percentage of individuals reporting low self-reliance (65.1%; see **Figure 2.8**) and high social boldness (63.4%; see **Figure 2.9**) relative to other subgroups.

LPA Model 3

Since the fit indices of LPA Models 1 and 2 did not converge perfectly, a new model was developed incorporating the differentiating indicators from Model 2 while adding significant indicators from Specific Aim 2. The added indicators were: age, combat deployments, and emotional stability. The indicators dropped from the model included: rank, physical fitness scores, disciplinary counseling statements, and social boldness scale scores. A 7-subgroup solution yielded the best for the data; **Table 3.3** depicts a summary of the fit indices and classification accuracy.

Subgroup 1 ("Intelligent") represented a subgroup with high intellectual functioning (14.5%, n = 572). Notably, all of the individuals from this highly intelligent subgroup had above average intelligence scores (100.0%; see **Figure 3.1**). By contrast, only a small percentage of individuals from other subgroups had above average intellectual functioning. This intelligent subgroup also had a much higher proportion of individuals reporting high reasoning abilities (33.1%; see **Figure 3.5**) than other subgroups. Additionally, this intelligence subgroup had 26.5% of individuals reporting low alienation/perceptual distortion (see **Figure 3.8**) and 40.6% reporting high emotional stability (see **Figure 3.9**), which was greater than several other subgroups.

Subgroup 2 ("Well Regulated") was comprised of self-controlled, emotionally stable participants (17.1%, n = 675). This subgroup had, by far, the highest percentage of individuals reporting high self-control (59.3%); see **Figure 3.10**), which appeared to be the key distinguishing indicator. The vast majority of individuals from this subgroup reported having high emotional stability (74.1%; see **Figure 3.9**), which occurred at a much higher percentage than all other subgroups except for subgroup four. In addition, this subgroup also had the highest percentage of individuals reporting low alienation/perceptual distortion in comparison to all other subgroups except for the fourth (51.8%; see **Figure 3.8**). This well-regulated subgroup had the lowest percentage of moderate to concerning alcohol use patterns (0.9%; see **Figure 3.4**).

Subgroup 3 ("Young/Average") was comprised of younger individuals who were average across all indicators (28.1%, n = 1110). The vast majority were individuals who were below 21 years old (75.7%), with no individuals being older than 23 years old (see **Figure 3.1**). Most individuals had not previously experienced a combat deployment (62.5%; see **Figure 3.2**). The vast majority had average intelligence (95.5%), reasoning abilities (77.3%), self-reliance (91.0%), health concerns (99.8%), alienation/perceptual distortion (91.0%), emotional stability (73.4%), self-control (85.4%), and interpersonal relations (89.5%).

Subgroup 4 ("Extroverted") represented a sociable subgroup of individuals (13.5%, n = 531). Nearly all of the participants from this social subgroup reported high interpersonal relations scores (88.9%; see **Figure 3.11**) and low self-reliance (83.5%; see **Figure 3.6**), which represented a stark contrast from other subgroups. This subgroup also had the highest percentage of individuals reporting low alienation/perceptual distortion (46.5%) relative to other subgroups except for the second (see **Figure 3.8)**. Most individuals in this sociable subgroup reported high emotional stability (65.9%; see **Figure 3.9**).

Subgroup 5 ("Psychologically Concerning") was comprised of individuals with maladaptive psychosocial functioning (2.0%, n = 80). Relative to other subgroups, this psychologically concerning subgroup had a much greater proportion of individuals reporting moderate to concerning alcohol use patterns (20.7%; see **Figure 3.4**) and high health concerns (36.3%; see **Figure 3.7**). Additionally, this psychologically concerning subgroup reported the highest percentage of individuals reporting high alienation/perceptual distortion (71.5%; see **Figure 3.8**), low emotional stability (49.8%; see **Figure 3.9**), low reasoning ability (36.7%; see **Figure 3.5**), and low self-control (31.2%; see **Figure 3.10**). Additionally, high self-reliance (43.7%; see **Figure 3.6**) and low interpersonal relations scores (48.9%; see **Figure 3.11**) were also more likely to be endorsed by individuals in this psychologically concerning subgroup relative to all other subgroups except for one.

Subgroup 6 ("Introverted") represented a subgroup of introverted individuals (6.5%, n = 257). All of the individuals from this subgroup reported low interpersonal relations (100.0%; see **Figure 3.11**), which represented by far the highest proportion of individuals endorsing these traits of all the subgroups. Additionally, nearly half of individuals from this introverted subgroups reported high self-reliance (47.5%; see **Figure 3.6**), which was greater than all other subgroups except the fifth psychologically concerning subgroup. The vast majority of individuals from this subgroup reported average levels of all other psychological characteristics.

Subgroup 7 ("Experienced/Average") represented a subset of older, militarily-experienced participants (18.3%, n = 722) who otherwise had relatively innocuous, indistinguishable psychological characteristics. The majority of individuals in this subgroup were older than 23 years old (52.7%), with very few being under 21 years old (0.6%; see **Figure 3.1**). Most participants had at least one combat deployment, with approximately 29.0% having at least two (see **Figure 3.2**). The vast majority of individuals in this older, militarily-experienced subgroup reported average intelligence (97.9%), reasoning ability (70.2%), self-reliance (92.1%), health concerns (95.5%), alienation/perceptual distortion (83.8%), emotional stability (89.3%), self-control (87.1%), and interpersonal relations (89.5%).

LPA Model Summary

The variables identified in Specific Aims 1 and 2 were used to develop three models to identify latent subgroups within the data. Each model iteration resulted in a unique solution. Models 1 and 2 did not converge precisely as indicated by the fit indices. However, Model 3 converged across fit indices with a 7-subgroup solution. This finding indicates that the data set was best characterized with seven latent, or unobserved, subgroups. This best-fitting, 7-subgroup model was then validated to evaluate its usefulness as a model.

VALIDATION OF LATENT MODEL

The best-converged model was validated by comparing the identified solutions (*i.e.*, 7-subgroup) using the training outcomes (*i.e.*, pass/fail, GPA, and instructor rating). A useful model will differentiate subgroups based on training performance, primarily on the attrition rate. For the LPA solutions, subgroup comparisons were made between the identified "Control/Average" subgroup from the solution and the other identified subgroups. Those analyses are detailed below.

7-Subgroup Solution (Model 3)

A Pearson's chi-square test revealed significant subgroup differences in training outcome, χ^2 (6, N = 3942) = 134.13, $p < .001$, $\varphi = 0.20$ (**Figure 4.1**). Additionally, **Table 3.4** depicts the summary of results from the binary logistic regression analysis examining differences in attrition rates across the eight subgroups. Relative to the older/average group (22.3% fail), individuals in the highly intelligent (11.7% fail; $p < .001$), well-regulated (10.5% fail; $p < .001$), and extroverted (14.7%; $p = .001$) subgroups were significantly less likely to fail out of the training program. By contrast, individuals in the psychologically concerning subgroup were approximately 3 times more likely to fail out of the training program in comparison to controls (53.8% fail; $p < .001$). The introverted (22.2% fail; p = .96) and younger/average (21.5% fail; $p = .66$) subgroups were not significantly different from the control group on overall training outcome.

Among training program completers (N = 2153), there were significant subgroup differences with regard to instructor rating, $F(6, 2146) = 5.47$, $p < .001$, $\eta^2 = 0.015$. The high intellectual functioning (M = 3.50, SD = 0.55), and well-regulated (M = 3.52, SD = 0.56) subgroups had the highest instructor ratings, but did not differ from each other. These two subgroups had significantly higher instructor ratings than the young/average (M = 3.34, SD = 0.54) and older/average (M = 3.41, SD = 0.61) subgroups (*p*s < .05), which did not differ from each other (*p*s > .05). The well-regulated subgroup also had significantly higher instructor ratings than the introverted subgroup (M = 3.39, SD = 0.59). The extroverted subgroup (M = 3.46, SD = 0.59) had significantly higher instructor ratings than the young/average subgroup ($p = .004$), but did not differ significantly from other subgroups. The psychologically concerning subgroup (M = 3.31, SD = 0.44) had the lowest instructor ratings, but did not significantly differ from any other subgroups.

Among training program completers (N = 3457), there were significant subgroup differences with regard to final GPA, $F(6, 3450) = 38.73$, $p < .001$, $\eta^2 = 0.063$. The high intellectual functioning subgroup (M = 94.02, SD = 2.84) had a significantly higher GPA than all other subgroups. By contrast, the psychologically concerning (M = 90.43, SD = 4.28) and older/average (M = 91.06, SD = 4.18) subgroups had the lowest GPAs, and did not significantly differ from each other. Relative to these subgroups, the well-regulated (M = 93.01, SD = 3.56), extroverted (M = 92.24, SD = 3.68), introverted (M = 92.54, SD = 4.22), and young/average (M = 91.77, SD = 3.91) subgroups had higher GPAs.

CHAPTER 5: Discussion

The purpose of this project was to improve the assessment and selection (A&S) efforts of an East Coast U.S. military training command. This effort at improvement was designed as a retrospective analysis of baseline, training outcome, and training performance data collected at the command. Baseline data included demographic, military, and psychological variables. Training outcome was assessed as either successfully completing or failing to complete the training program. Training performance for program completers included program grade point average (GPA) and instructor rating.

The following sections provide commentary on the specific aims, discuss the findings overall, explore some limitations of the project, and provide an outline of future directions for this project and other efforts at improved A&S.

COMMENTARY ON SPECIFIC AIMS

This study aimed to: (1) identify the relative contribution of demographic and psychological factors that enhance prediction of trainee attrition above and beyond current assessment and selection (A&S) methods; (2) identify the relative contribution of demographic and psychological factors that enhance prediction of trainee performance above and beyond current A&S methods; and (3) characterize subgroups within the sample population for whom training outcome risk was increased and/or decreased. These aims were accomplished using logistic regression, multiple regression, and latent profile analysis, respectively. The following sections discuss the findings associated with each specific aim.

Specific Aim 1

With regard to Specific Aim 1, identifying variables that optimally predict training attrition, this project established that **logistic regression can identify variables that predict passing versus failure better than chance**. Specifically, Model 8 from the results identified the following variables as statistically significant predictors of training completion: sex, rank, GT score, disciplinary counseling statements, physical fitness score, mental health treatment history, AUDIT total score, 16PF reasoning, 16PF social boldness, 16PF interpersonal relations, 16PF health concerns, and 16PF alienation/perceptual distortion.

Model 8, the most efficient model derived using logistic regression, correctly classified 83.5% of total responses. The model also identified 98.1% of the successful trainee outcomes. However, the model only was able to correctly identify 17.8% of the training failures. *A priori*, these classification accuracies indicate that the overall model was good, the identification of training successes was excellent, and the identification of training failures was inadequate.

Classification accuracy of training failures was considered the primary goal of this analysis, and the optimal model did not meet acceptable classification criteria for training failures. However, because the majority of training successes were correctly classified, the findings indicate that any case classified as a training failure should be flagged as very high risk for training failure. As such, **Model 8 may be of value to the training command as an additional check in trainee assessment and selection**.

Specific Aim 2

With regard to Specific Aim 2, identifying variables that optimally predict training performance, this project established that **multiple regression can identify variables that predict passing versus failure better than chance**. Specifically, Model 8a from Hypothesis 1 identified the following variables as statistically significant predictors of training performance for GPA: GT score, disciplinary counselings, physical fitness, combat deployments, emotional stability, liveliness, apprehension, emotional adjustment, and intellectual efficiency. Model 1b and 8b from Hypothesis 2 identified the following variables as statistically significant predictors of training performance for instructor ratings: rank, GT score, age, sex, disciplinary counselings, physical fitness, and combat deployments.

Among program completers, the most efficient multiple regression models were able to account for 29.7% of the variance for program GPA (Model 8a) and 27.0% of the variance for instructor ratings (Model 8b). Given that the participants included in these analyses all successfully completed the program, these models significantly predict relative training performance. However, it should be noted that participant characteristics (Model 1a) accounted for 27.3% of the variance for program GPA compared with the improvement to 29.7% of the variance in Model 8a with the inclusion of some psychological variables. Similarly, participant characteristics (Model 1b) were the only statistically significant predictors for instructor ratings.

Based on these findings, **baseline participant characteristics can be used to predict training program performance better than chance**. The identified variables are easily trackable, and so these variables can be used to predict training performance which may improve resource allocation for A&S at this command. However, the finding that baseline characteristics were the best predictors of instructor rating is a potentially concerning finding. It is posited that instructor ratings were of multiple, professional characteristics based on participant performance in the program. It is possible that no psychological variables predicted these ratings because the outcome variable was collapsed across those rated domains to create a single rating score. If the individual ratings had been analyzed separately, psychological characteristics may have been better predictors of those ratings than baseline characteristics alone.

Alternatively, this finding that baseline participant characteristics are the best predictor of instructor ratings may indicate that the instructors are not making accurate, explicit judgments of these professional characteristics. Rather, the instructor ratings are, at least to some extent, biased by explicit participant characteristics (e.g., sex) rather than participant performance.

Specific Aim 3

With regard to Specific Aim 3, the effort to identify, characterize, and validate subgroups within the sample population, it was found that latent profile analysis (LPA) can be used to achieve these efforts. Specifically, LPA revealed a best-fitting, 7-subgroup solution (*i.e.*, seven subgroups), indicating that **LPA can successfully be applied to training data to identify subgroups**. These seven subgroups were characterized based on their respective differentiating variable features. Finally, these seven subgroups were validated based on attrition risk, and they differed in risk ranging from a 46.3% pass rate to 89.5% pass rate. Therefore, **this specific aim was successful on all counts**.

General Discussion

The statistical approaches used in this study were able to predict training program attrition and performance to a statistically significant degree. Characterizing trainees into statistically-identified subgroups using Latent Profile Analysis (LPA) may be the most effective approach for predicting training program outcomes and improving upon current assessment and selection (A&S) methods.

The findings from this project will be used to inform A&S decisions for the training command from which the data were gathered. Further, the findings from this research can be validated and refined on subsequent groups of trainees at the command, resulting in continual process improvement.

Beyond the training command from which the data was acquired, the process developed in this study, particularly the application of LPA to A&S, can be adapted to other A&S programs. Although the results from the LPA do not assign individual risk for attrition like the logistic regression analyses, the LPA can assist organizations with broader risk stratification decisions when evaluating groups of trainees. The characteristics of these subgroups can be deemed desirable or undesirable based on the needs of the organization, and subsequent decisions can be made regarding efforts aimed at increasing retention of risky (*i.e.*, high attrition risk) but desirable subgroups as well as selecting out risky, undesirable candidates.

POTENTIAL LIMITATIONS

This project had limitations that may have adversely impacted the predictive utility of the results. First and foremost, there were missing data. Although the missing data was accounted for through statistical processes (*e.g.*, multiple imputation), ensuring that a dataset is complete would improve the certainty with which the study conclusions could be drawn. Secondly, some variables that may have increased predictive value (particularly the PCL-M) were dropped because of the amount of missing data. Fortunately, the PCL-M can and has been used as an independent check during assessment and selection (A&S). Specifically, individuals scoring in the clinical range on the PCL-M are most likely dropped because of unsuitability for special assignments. As such, this measure still can be used in conjunction with the approach developed in this study to improve A&S.

Another limitation of this study is that the latent profile analysis results did not converge perfectly (*i.e.*, the fit indices did not all indicate the best fitting solution). This good, though imperfect, model fit is a challenge inherent to latent profile analysis. Discretion on the part of the researcher is necessary to interpret the statistics resulting from the analyses and, in this case, the 7-subgroup solution was functionally interpretable. In other words, this solution made sense when validated against the training outcomes.

When considering the practical application of these results, a limitation that must be noted is that the outcome data are solely from the training command. The results from this research may not directly predict performance in the real-world mission settings that the successful graduates from the training program will encounter. While prediction of training attrition is an important step in improving A&S for this command, the next logical step would be to apply the methodology developed in this study to identify service members at risk for attrition following completion of the program.

Though the data set is large, it was acquired from a single training site. The findings may not generalize to other special training programs, military or civilian, because of the homogenous nature of training sample. In defense of the sample, it does contain service members from a variety of military occupational specialties, which suggests that it may be more generalizable than an analysis of training data from a single MOS training setting. However, the specific findings from these analyses (*i.e.*, statistically significant predictors) should be taken as a starting point for future program evaluations rather than as an ending.

RECOMMENDATIONS

One central purpose of this project was to provide direct, actionable feedback to the command from which the data was acquired. Based on this research, it is recommended that the Commanding Officer of this training command drop any student belonging to the "psychologically concerning" subgroup. The risk for attrition is far greater for individuals classified into this subgroup, and the combination of traits that characterize this subgroup is incompatible with success in a special military assignment.

A secondary finding is that instructor ratings should be viewed as supplemental information that may be less accurate than objective performance data. Instructor ratings do not appear to contribute to attrition risk evaluation better than baseline student characteristics.

Finally, the substantial attrition for women in the program should be explored. With full gender integration of military assignments as a pressing issue for the Department of Defense, the reasons underlying this higher proportion of attrition for women should be evaluated to ensure that there are no systemic biases that unfairly undermine the opportunity of women to serve in this command.

FUTURE DIRECTIONS

An important future direction for this research is to consider incorporating additional variables, particularly from past military performance, into the analyses that may improve predictive accuracy. An important performance variable that would be possible to obtain prior to entry into this and other special training programs is past military performance. Data such as past military evaluations, performance-related recognition or punishments, and performance in other service schools could provide valuable predictive improvements for attrition risk.

Another performance component that may improve attrition prediction could be peer evaluations. Peer evaluations are utilized in other training programs (*e.g.,* U.S. Army Ranger School), and this additional subjective evaluation may add an important interpersonal evaluative component that is not captured by instructor ratings or external performance, such as GPA (72).

Assessment and selection (A&S) improvements also might be made by assessing additional psychological factors. Specifically, motivation(s) to enter the training program (*e.g.*, intrinsic versus extrinsic), resilience or psychological hardiness, and locus of control (*e.g.*, internal versus external) are constructs that may potentially improve the A&S of this and other training programs. Motivations to enter the training program could be assessed using a measure such as the Work Extrinsic and Intrinsic Motivation Scale (64), as internal and instrumental motivations have been shown to impact attrition and performance in United States Military Academy Cadets (76). Resilience, psychological hardiness, and mental toughness are all related concepts correlated with performance in military training (1; 26; 36). Locus of control has been linked to success in a U.S. Army graduate nursing program and the Corps of Cadets at Texas A&M (16; 33).

A&S might be improved by including assessments for psychopathology. One of the most extensively researched clinical assessment tools is the Minnesota Multiphasic Personality Inventory (MMPI-2; MMPI-2-RF). This tool has been used in pre-employment screening, particularly with emergency responder candidates (60; 61). Including the MMPI-2 not only provides information on personality and psychopathology that could be disqualifying from a special assignment, but the instrument also includes validity scales that may be valuable in assessing response style (e.g., minimizing psychological concerns) for the candidates. However, this added information comes with a notable administrative burden of time (est. 35-50 minutes) and money as each administration of the MMPI costs money.

Another clinical construct that may improve A&S for special military populations is impulse control. It is possible that including a self-report measure of impulsiveness, such as the Barratt Impulsiveness Scale (BIS-11), could assist with identifying better candidates. Performance-based measures of impulse control, such as the Stroop Color-Word Test, may provide a relatively brief check on attention that could improve A&S.

When considering application of these findings, the relative attrition risk information and stratification may be valuable for the command to direct training resources toward the different risk groups. The command may determine that resources can best be spent supporting and observing candidates who fall into the higher attrition risk categories. Training staff and students could potentially be distributed in a manner to either comingle high and low risk students or to separate candidates into higher and lower attrition risk training units. Alternatively, commands may find that the increased probability of training attrition is not worth the time and effort when those resources could be directed toward students with a much higher probability of completing the program.

The methodology and findings from this project could lead to revisions in pre-accession screening for training programs throughout the military, and potentially within other organizations as well. The methodology utilized in this study (*i.e.*, (1) identify variables that predict success/failure; (2) enter these variables in latent profile analyses to identify subgroups within a sample) can be applied to any training command that collects and retains data on their student population. This methodology may provide insight into attrition patterns of the sample that are not readily apparent using current A&S methods.

The specific findings from this study (*i.e.*, the statistically significant variables that predict attrition) may be applicable to other programs within and outside of the military in limited circumstances. In the case of organizations that have established A&S programs, the variables identified in this study can serve as a benchmark from which to compare and contrast additional variables' relevance to attrition. For organizations looking to initiate or expand an A&S program, the significant variables identified in this project also can serve as a baseline from which to build a program tailored to that organization.

These changes in screening processes could save the U.S. military substantial financial and manpower assets by identifying high risk individuals or subgroups within a training population, and this identification could be used to reduce the attrition rate through increased resource allocation or increasingly selective entry standards. Further, a two- or multi-stage approach to A&S could be adopted in which a pool of candidates is stratified by risk, and the medium to high-risk candidates are subjected to a second round of screening. This multi-stage, iterative approach may improve identification of attrition over the single-stage model utilized in this study because the less frequent outcome (*i.e.*, attrition) becomes more common in the second-stage statistical analyses, improving predictive validity of the included variables. A multi-stage approach in this framework would select in candidates in the first round of analyses and select out candidates in the second round.

CONCLUSIONS

This study further validated the use of statistical methods to improve assessment and selection (A&S) in a special military training program. The novel application of latent profile analysis to training data provides a proof of concept to identify, characterize, and validate subgroups within a training command that can be stratified by relative training attrition risk. This same statistical approach may be applied to A&S at other military training settings as well as within civilian organizations. Although the specific results from this study may not be applicable to other settings, the methodology utilized is readily transferrable to any setting that collects baseline and outcome data and has an interest in improving the A&S of its potential members.

Appendix A – Tables

Table 1.1. *Summary of Logistic Regression Analysis for Demographic and Military Characteristics Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Sex	0.53	0.15	1.71 (1.28 – 2.28)	< .001
Rank				< .001
E-2/E-3	--	--	--	--
E-4	0.55	0.13	1.74 (1.34 – 2.25)	< .001
E-5	0.17	0.12	1.19 (0.94 – 1.49)	.15
MOS				.02
Support	--	--	--	--
Ground	0.26	0.12	1.30 (1.03 – 1.63)	.03
Aviation	0.03	0.13	1.03 (0.80 – 1.34)	.82
GT Score	-0.03	0.004	0.97 (0.96 – 0.98)	< .001
Mental Health Treatment	0.61	0.11	1.84 (1.47 – 2.30)	< .001
Counseling Statements	0.14	0.03	1.15 (1.09 – 1.20)	< .001
Physical Fitness Score	-0.01	0.002	0.992 (0.989 – 0.995)	< .001
AUDIT Total Score	0.18	0.02	1.20 (1.16 – 1.24)	< .001
Constant	2.47	0.64	--	< .001

Notes: MOS = military occupational specialty; GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test. Reference category for sex is female and for mental health treatment is none.

Table 1.2. *Summary of Logistic Regression Analysis for 16PF Response Style Scales Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Impression Management	-0.07	0.01	0.94 (0.92 – 0.95)	< .001
Infrequency	0.06	0.02	1.06 (1.02 – 1.10)	.006
Constant	-0.62	0.13	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 1.3. *Summary of Logistic Regression Analysis for 16PF Global Factor Scales Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Anxiety	0.19	0.03	1.21 (1.14 – 1.27)	< .001
Self-Control	-0.19	0.03	0.83 (0.78 – 0.88)	< .001
Constant	-1.23	0.27	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 1.4. *Summary of Logistic Regression Analysis for 16PF Primary Factor Scales Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Reasoning	-0.24	0.03	0.79 (0.75 – 0.84)	< .001
Emotional Stability	-0.15	0.04	0.86 (0.80 – 0.92)	< .001
Rule-Consciousness	-0.09	0.03	0.91 (0.86 – 0.97)	.002
Social Boldness	0.07	0.03	1.08 (1.02 – 1.13)	.008
Abstractedness	0.13	0.03	1.14 (1.06 – 1.21)	< .001
Self-Reliance	0.08	0.03	1.09 (1.03 – 1.15)	.004
Tension	0.05	0.03	1.05 (0.99 – 1.12)	.10
Constant	-0.53	0.49	--	.28

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 1.5. *Summary of Logistic Regression Analysis for 16PF Protective Services Dimensions Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Emotional Adjustment	-0.25	0.03	0.77 (0.74 – 0.82)	< .001
Intellectual Efficiency	-0.13	0.03	0.88 (0.83 – 0.93)	< .001
Constant	0.85	0.18	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 1.6. *Summary of Logistic Regression Analysis for 16PF Pathology-Oriented Scales Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Health Concerns	0.31	0.06	1.37 (1.22 – 1.53)	< .001
Suicidal Thinking	0.19	0.07	1.21 (1.06 – 1.38)	.006
Low Energy State	0.12	0.04	1.13 (1.04 – 1.23)	.005
Apathetic Withdrawal	-0.08	0.03	0.92 (0.87 – 0.98)	.01
Alienation/Perceptual Distortion	0.18	0.04	1.20 (1.11 – 1.29)	< .001
Threat Immunity	0.10	0.04	1.10 (1.02 – 1.18)	.01
Constant	-5.75	0.45	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 1.7. *Summary of Logistic Regression Analysis for Selected Demographic, Military, and Psychological Characteristics Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Sex	0.72	0.15	2.05 (1.52 – 2.76)	< .001
Rank				.001
E-2/E-3	--	--	--	--
E-4	0.47	0.14	1.61 (1.23 – 2.10)	.001
E-5	0.14	0.12	1.15 (0.91 – 1.46)	.24
GT Score	-0.02	0.01	0.98 (0.97 – 0.99)	.001
Mental Health Treatment	0.53	0.12	1.70 (1.35 – 2.14)	< .001
Counseling Statements	0.12	0.03	1.12 (1.07 – 1.18)	< .001
Physical Fitness Score	-0.007	0.002	0.993 (0.990 – 0.996)	< .001
AUDIT Total Score	0.16	0.02	1.17 (1.13 – 1.21)	< .001
16PF Self-Control	-0.07	0.04	0.93 (0.87 – 0.99)	.04
16PF Reasoning	-0.13	0.04	0.87 (0.81 – 0.94)	< .001
16PF Emotional Stability	-0.07	0.04	0.94 (0.87 – 1.01)	.08
16PF Social Boldness	0.07	0.03	1.07 (1.01 – 1.13)	.03
16PF Self-Reliance	0.10	0.03	1.11 (1.05 – 1.18)	.001
16PF Health Concerns	0.27	0.06	1.31 (1.17 – 1.46)	< .001
16PF Alienation/Perceptual Distortion	0.11	0.04	1.11 (1.03 – 1.20)	.002
16PF Threat Immunity	0.09	0.04	1.10 (1.01 – 1.19)	.02
Constant	-0.86	0.93	--	.36

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test; 16PF =

16 Personality Factor Questionnaire. Reference category for sex is female and for mental health treatment is none.

Table 1.8. *Summary of Logistic Regression Analysis for All Potential Demographic, Military, and Psychological Characteristics Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Age	0.06	0.02	1.06 (1.01 – 1.10)	.01
Sex	0.75	0.16	2.11 (1.54 – 2.88)	< .001
Rank				.004
E-2/E-3	--	--	--	--
E-4	0.46	0.16	1.59 (1.15 – 2.18)	< .001
E-5	0.14	0.14	1.15 (0.88 – 1.51)	.06
MOS	-0.12	0.06	0.88 (0.79 – 0.99)	.03
GT Score	-0.02	0.01	0.98 (0.97 – 0.99)	< .001
Mental Health Treatment	0.53	0.12	1.70 (1.34 – 2.15)	< .001
Combat Deployments	-0.18	0.07	0.83 (0.73 – 0.96)	.009
Counseling Statements	0.12	0.03	1.12 (1.07 – 1.18)	< .001
Physical Fitness Score	-0.007	0.002	0.993 (0.990 – 0.996)	< .001
AUDIT Total Score	0.16	0.02	1.17 (1.13 – 1.21)	< .001
16PF Anxiety	0.10	0.05	1.11 (1.00 – 1.22)	.05
16PF Self-Control	-0.11	0.04	0.89 (0.82 – 0.97)	.008
16PF Warmth	0.08	0.04	1.08 (1.00 – 1.18)	.07
16PF Reasoning	-0.12	0.04	0.89 (0.82 – 0.95)	.001
16PF Liveliness	0.10	0.05	1.10 (1.00 – 1.22)	.06
16PF Social Boldness	0.16	0.05	1.17 (1.07 – 1.29)	.001
16PF Sensitivity	-0.06	0.03	0.95 (0.88 – 1.01)	.10

16PF Apprehension	-0.11	0.05	0.89 (0.81 – 0.99)	.03
16PF Interpersonal Relations	-0.24	0.08	0.78 (0.67 – 0.91)	.002
16PF Health Concerns	0.26	0.06	1.29 (1.15 – 1.46)	< .001
16PF Suicidal Thinking	0.13	0.07	1.14 (1.00 – 1.31)	.06
16PF Alienation/Perceptual Distortion	0.11	0.04	1.12 (1.03 – 1.21)	.009
16PF Thrill Seeking	-0.06	0.03	0.94 (0.88 – 1.01)	.08
16PF Threat Immunity	0.09	0.04	1.10 (1.01 – 1.19)	.03
Constant	-1.73	1.17	--	.14

Notes: MOS = military occupational specialty; GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test; 16PF = 16 Personality Factor Questionnaire. Reference category for sex is female and for mental health treatment is none.

Table 1.9. *Summary of Logistic Regression Analysis for Empirically Selected Demographic, Military, and Psychological Characteristics Predicting Overall Training Outcome (Attrition).*

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Sex	0.64	0.15	1.90 (1.41 – 2.55)	< .001
Rank				.019
E-2/E-3	--	--	--	--
E-4	0.32	0.15	1.37 (1.02 – 1.85)	.036
E-5	0.03	0.13	1.03 (0.80 – 1.33)	.822
GT Score	-0.02	0.01	0.98 (0.97 – 0.99)	.001
Mental Health Treatment	0.52	0.12	1.68 (1.34 – 2.13)	< .001
Combat Deployments	-0.15	0.07	0.86 (0.75 – 0.98)	.028
Counseling Statements	0.116	0.03	1.12 (1.07 – 1.18)	< .001
Physical Fitness Score	-0.007	0.002	0.99 (0.99 – 1.00)	< .001
AUDIT Total Score	0.16	0.02	1.17 (1.13 – 1.22)	< .001
16PF Self-Control	-0.10	0.03	0.91 (0.85 – 0.97)	.003
16PF Reasoning	-0.14	0.04	0.87 (0.81 – 0.94)	< .001
16PF Social Boldness	0.07	0.03	1.08 (1.02 – 1.14)	.009
16PF Self-Reliance	0.11	0.03	1.12 (1.06 – 1.19)	< .001
16PF Health Concerns	0.28	0.05	1.32 (1.19 – 1.46)	< .001
16PF Alienation/Perceptual Distortion	0.12	0.04	1.12 (1.04 – 1.21)	.002
Constant	-0.31	0.86	--	.723

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test; 16PF = 16 Personality Factor Questionnaire. Reference category for sex is female and for mental health treatment is none.

Table 2A.1. *Summary of Multiple Regression Analysis for Demographic and Military Characteristics Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.06	0.02	0.05	.01
GT Score	0.09	0.01	0.29	< .001
Combat Deployments	-0.31	0.04	-0.15	< .001
Counseling Statements	-0.89	0.04	-0.35	< .001
Physical Fitness Score	0.01	0.002	0.08	< .001
AUDIT Total Score	-0.05	0.01	-0.06	< .001
Constant	81.00	0.92	--	< .001

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test.

Table 2A.2. *Summary of Multiple Regression Analysis for 16PF Response Style Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Impression Management	0.02	0.01	0.07	<.001
Constant	92.25	0.09	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2A.3. *Summary of Multiple Regression Analysis for 16PF Global Factor Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Anxiety	-0.03	0.02	-0.03	.09
Self-Control	0.07	0.02	0.06	.001
Constant	92.26	0.19	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2A.4. *Summary of Multiple Regression Analysis for 16PF Primary Factor Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Liveliness	-0.04	0.02	-0.05	.01
Rule-Consciousness	0.04	0.02	0.05	.01
Tension	-0.06	0.02	-0.06	.003
Constant	92.76	0.21	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2A.5. *Summary of Multiple Regression Analysis for 16PF Protective Services Dimensions Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Emotional Adjustment	0.04	0.02	0.04	.07
Integrity/Control	0.04	0.02	0.04	.07
Constant	92.03	0.13	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2A.6. *Summary of Multiple Regression Analysis for 16PF Pathology-Oriented Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Alienation/Perceptual Distortion	-0.05	0.02	-0.05	.01
Thrill Seeking	-0.04	0.02	-0.04	.03
Constant	93.03	0.12	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2A.7. *Summary of Multiple Regression Analysis for Selected Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.05	0.02	0.04	.05
GT Score	0.08	0.01	0.29	< .001
Combat Deployments	-0.37	0.04	-0.17	< .001
Counseling Statements	-0.88	0.04	-0.35	< .001
Physical Fitness Score	0.01	0.002	0.09	< .001
16PF Self-Control	0.12	0.02	0.12	< .001
16PF Liveliness	-0.05	0.02	-0.05	.003
16PF Tension	-0.04	0.02	-0.05	.02
Constant	80.51	0.92	--	< .001

Notes: GT = general technical (intelligence); 16PF = 16 Personality Factor Questionnaire.

Table 2A.8. *Summary of Multiple Regression Analysis for All Potential Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.06	0.02	0.04	.02
GT Score	0.09	0.01	0.29	< .001
Combat Deployments	-0.37	0.04	-0.17	< .001
Counseling Statements	-0.89	0.04	-0.35	< .001
Physical Fitness Score	0.01	0.002	0.09	< .001
AUDIT Total Score	-0.02	0.01	-0.03	.07
16PF Acquiescence	-0.01	0.003	-0.03	.07
16PF Emotional Stability	-0.10	0.03	-0.09	.001
16PF Liveliness	-0.06	0.02	-0.07	< .001
16PF Apprehension	0.09	0.02	0.08	< .001
16PF Emotional Adjustment	0.22	0.03	0.08	< .001
16PF Intellectual Efficiency	-0.05	0.02	-0.05	.002
16PF Self-Reproach	-0.04	0.02	-0.03	.06
Constant	80.40	0.95	--	< .001

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test; 16PF = 16 Personality Factor Questionnaire.

Table 2A.9. *Summary of Multiple Regression Analysis for Selected Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
GT Score	0.09	0.004	0.29	< .001
Combat Deployments	-0.33	0.03	-0.16	< .001
Counseling Statements	-0.89	0.04	-0.35	< .001
Physical Fitness Score	0.01	0.002	0.09	< .001
16PF Emotional Stability	-0.11	0.03	-0.10	< .001
16PF Liveliness	-0.07	0.01	-0.08	< .001
16PF Apprehension	0.07	0.02	0.06	.001
16PF Emotional Adjustment	0.24	0.03	0.24	< .001
16PF Intellectual Efficiency	-0.05	0.02	-0.06	.002
Constant	81.19	0.81	--	< .001

Notes: GT = general technical (intelligence); 16PF = 16 Personality Factor Questionnaire.

Table 2B.1. *Summary of Multiple Regression Analysis for Demographic and Military Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.16	< .001
Sex	-0.18	0.04	-0.08	< .001
GT Score	0.002	0.001	0.06	.001
Combat Deployments	0.03	0.01	0.11	< .001
Counseling Statements	-0.09	0.01	-0.25	< .001
Physical Fitness Score	0.004	0.001	0.21	< .001
AUDIT Total Score	0.004	0.002	0.03	.05
Constant	1.30	0.14	--	< .001

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test.

Reference category for sex is female.

Table 2B.4. *Summary of Multiple Regression Analysis for 16PF Primary Factor Scales Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Openness to Change	-0.01	0.003	-0.04	.01
Constant	3.51	0.02	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table 2B.7. *Summary of Multiple Regression Analysis for Selected Demographic and Military Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.17	< .001
Sex	-0.17	0.03	-0.08	< .001
GT Score	0.003	0.001	0.06	< .001
Combat Deployments	0.03	0.01	0.11	< .001
Counseling Statements	-0.09	0.01	-0.24	< .001
Physical Fitness Score	0.004	0.001	0.21	< .001
Constant	1.30	0.14	--	< .001

Notes: GT = general technical (intelligence). Reference category for sex is female.

Table 2B.8. *Summary of Multiple Regression Analysis for All Potential Personal, Military, and Psychological Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.16	< .001
Sex	-0.19	0.04	-0.09	< .001
GT Score	0.002	0.001	0.06	.001
Combat Deployments	0.03	0.01	0.11	< .001
Counseling Statements	-0.09	0.01	-0.24	< .001
Physical Fitness Score	0.004	0.001	0.21	< .001
AUDIT Total Score	0.004	0.002	0.03	.05
16PF Tough-Mindedness	-0.01	0.003	-0.04	.08
16PF Openness to Change	-0.01	0.003	-0.04	.03
Constant	1.38	0.14	--	< .001

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test; 16PF = 16 Personality Factor Questionnaire. Reference category for sex is female.

Table CVM 2A.1. *Summary of Multiple Regression Analysis for Demographic and Military Characteristics Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
GT Score	0.09	0.01	0.31	< .001
Combat Deployments	-0.25	0.04	-0.12	< .001
Counseling Statements	-0.90	0.04	-0.36	< .001
Physical Fitness Score	0.01	0.002	0.07	< .001
Mental Health Treatment	-0.26	0.15	-0.03	.08
AUDIT Total Score	-0.04	0.01	-0.05	.006
Constant	82.06	0.88	--	< .001

Notes: GT = general technical (intelligence); AUDIT = Alcohol Use Disorders Identification Test.

Reference category for mental health treatment history is "No prior treatment."

Table CVM 2A.2. *Summary of Multiple Regression Analysis for 16PF Response Style Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Impression Management	0.02	0.01	0.06	.004
Constant	92.32	0.09	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.3. *Summary of Multiple Regression Analysis for 16PF Global Factor Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Self-Control	0.08	0.02	0.07	.001
Constant	92.09	0.14	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.4. *Summary of Multiple Regression Analysis for 16PF Primary Factor Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Liveliness	0.05	0.02	0.05	.01
Rule-Consciousness	0.04	0.02	0.03	.08
Tension	-0.06	0.02	-0.06	.007
Constant	92.29	0.22	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.5. *Summary of Multiple Regression Analysis for 16PF Protective Services Dimensions Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Integrity/Control	0.06	0.02	0.07	.001
Constant	92.19	0.12	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.6. *Summary of Multiple Regression Analysis for 16PF Pathology-Oriented Scales Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Alienation/Perceptual Distortion	-0.06	0.02	-0.05	.007
Constant	93.03	0.12	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.7. *Summary of Multiple Regression Analysis for Selected Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
GT Score	0.08	0.01	0.30	< .001
Combat Deployments	-0.32	0.04	-0.15	< .001
Counseling Statements	-0.90	0.04	-0.36	< .001
Physical Fitness Score	0.01	0.002	0.09	< .001
16PF Self-Control	0.15	0.02	0.13	< .001
16PF Alienation/Perceptual Distortion	-0.03	0.02	-0.03	.10
Constant	80.94	0.90	--	< .001

Notes: GT = general technical (intelligence); 16PF = 16 Personality Factor Questionnaire.

Table CVM 2A.8. *Summary of Multiple Regression Analysis for All Potential Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
GT Score	0.09	0.01	0.30	< .001
Combat Deployments	-0.32	0.04	-0.15	< .001
Counseling Statements	-0.90	0.04	-0.36	< .001
Physical Fitness Score	0.01	0.002	0.08	< .001
Mental Health Treatment	-0.27	0.15	-0.03	.07
16PF Anxiety	0.20	0.05	0.19	< .001
16PF Liveliness	-0.06	0.02	-0.07	< .001
16PF Vigilance	-0.06	0.02	0.06	.01
16PF Tension	-0.10	0.03	-0.10	< .001
16PF Emotional Adjustment	0.21	0.04	0.21	< .001
16PF Intellectual Efficiency	-0.06	0.02	-0.06	.001
16PF Paranoid Ideation	0.04	0.02	0.04	.09
16PF Alienation/Perceptual Distortion	-0.05	0.02	-0.04	.05
Constant	80.93	0.96	--	< .001

Notes: GT = general technical (intelligence); 16PF = 16 Personality Factor Questionnaire. Reference category for mental health treatment history is "No Prior Treatment."

Table CVM 2A.9. *Summary of Multiple Regression Analysis for Selected Personal, Military, and Psychological Factors Predicting Grade Point Average in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
GT Score	0.09	0.01	0.30	< .001
Combat Deployments	-0.32	0.04	-0.15	< .001
Counseling Statements	-0.90	0.04	-0.36	< .001
Physical Fitness Score	0.01	0.002	0.08	< .001
16PF Anxiety	0.12	0.04	0.12	.002
16PF Liveliness	-0.06	0.02	-0.07	< .001
16PF Tension	-0.08	0.03	-0.08	.001
16PF Emotional Adjustment	0.18	0.03	0.19	< .001
16PF Intellectual Efficiency	-0.05	0.02	-0.05	.006
Constant	80.96	0.94	--	< .001

Notes: GT = general technical (intelligence); 16PF = 16 Personality Factor Questionnaire.

Table CVM 2B.1. *Summary of Multiple Regression Analysis for Demographic and Military Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.16	< .001
Sex	-0.17	0.04	-0.08	< .001
GT Score	0.003	0.001	0.07	< .001
Combat Deployments	0.04	0.01	0.12	< .001
Counseling Statements	-0.09	0.01	-0.23	< .001
Physical Fitness Score	0.004	0.001	0.22	< .001
Constant	1.16	0.15	--	< .001

Notes: GT = general technical (intelligence). Reference category for sex is female.

Table CVM 2B.2. *Summary of Multiple Regression Analysis for 16PF Response Style Scales Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Impression Management	-0.002	0.001	-0.04	.08
Constant	3.50	0.01	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2B.4. *Summary of Multiple Regression Analysis for 16PF Primary Factor Scales Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Reasoning	0.01	0.003	0.03	.09
Dominance	0.01	0.004	0.04	.06
Openness to Change	-0.01	0.003	-0.04	.03
Constant	3.45	0.03	--	< .001

Notes: 16PF = 16 Personality Factor Questionnaire.

Table CVM 2B.7. *Summary of Multiple Regression Analysis for Selected Demographic and Military Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.16	< .001
Sex	-0.17	0.04	-0.08	< .001
GT Score	0.003	0.001	0.07	< .001
Combat Deployments	0.04	0.01	0.12	< .001
Counseling Statements	-0.09	0.01	-0.23	< .001
Physical Fitness Score	0.004	0.001	0.22	< .001
Constant	1.16	0.15	--	< .001

Notes: GT = general technical (intelligence). Reference category for sex is female.

Table CVM 2B.8. *Summary of Multiple Regression Analysis for All Potential Personal, Military, and Psychological Characteristics Predicting Instructor Ratings in Successful Program Completers.*

Predictor	**B**	**SE**	**Beta**	***p* value**
Age	0.03	0.004	0.16	< .001
Sex	-0.18	0.04	-0.08	< .001
GT Score	0.003	0.001	0.07	< .001
Combat Deployments	0.04	0.01	0.12	< .001
Counseling Statements	-0.09	0.01	-0.23	< .001
Physical Fitness Score	0.004	0.001	0.22	< .001
Constant	1.16	0.15	--	< .001

Notes: GT = general technical (intelligence). Reference category for sex is female.

Table 3.1. *Fit Indices for the Latent Profile Analysis Determining Latent Subgroups of Service Members in a Specialized Training Program – Model 1.*

Clusters	**Parameters**	**BIC**	**aBIC**	**cAIC**	**LL**	**Entropy**
1	12	34400.05	34361.92	34412.05	-17150.341	1.00
2	25	33826.61	33747.17	33851.61	-16809.795	0.69
3	38	**33760.77**	33640.02	**33798.77**	-16723.052	0.77
4	51	33796.79	**33634.74**	33847.79	-16687.239	0.66
5	64	33862.74	33659.37	33926.74	-16666.386	0.63
6	77	33939.83	33695.16	34016.83	-16651.107	0.55
7	90	34021.69	33735.71	34111.69	-16638.213	0.63
8	103	34113.85	33786.56	34216.85	-16630.466	0.63
9	116	34200.87	33832.27	34316.87	-16620.152	0.63

Note: BIC = Bayesian Information Criterion; cAIC = Consistent Akaike Information Criterion; aBIC = sample size adjusted BIC; LL = Log-likelihood; Lower BIC, aBIC, and cAIC values suggest a better model fit, whereas higher entropy values suggest better classification accuracy. The lowest obtained value (*i.e.*, best model fit) for the classification statistics are shown in bold.

Table 3.2. *Fit Indices for the Latent Profile Analysis Determining Latent Subgroups of Service Members in a Specialized Training Program – Model 2.*

Clusters	Parameters	BIC	aBIC	cAIC	LL	Entropy
1	20	59401.55	59338.00	59421.55	-29617.97	1.00
2	41	57596.14	57465.86	57637.14	-28628.32	0.89
3	62	56582.15	56385.14	56644.15	-28034.37	0.92
4	83	56267.10	56003.36	56350.10	-27789.90	0.85
5	104	56160.57	55830.11	56264.57	-27649.69	0.76
6	125	**56153.26**	55756.07	**56278.26**	-27559.09	0.81
7	146	56219.24	**55755.32**	56365.24	-27505.13	0.80
8	167	56309.41	55778.76	56476.41	-27463.27	0.80
9	188	56412.53	55815.15	56600.53	-27427.88	0.75

Note: BIC = Bayesian Information Criterion; cAIC = Consistent Akaike Information Criterion; aBIC = sample size adjusted BIC; LL = Log-likelihood; Lower BIC, aBIC, and cAIC values suggest a better model fit, whereas higher entropy values suggest better classification accuracy. The lowest obtained value (*i.e.*, best model fit) for the classification statistics are shown in bold.

Table 3.3. *Fit Indices for the Latent Profile Analysis Determining Latent Subgroups of Service Members in a Specialized Training Program – Model 3.*

Clusters	Parameters	BIC	aBIC	cAIC	LL	Entropy
1	19	56125.37	56064.99	56144.37	-27984.02	1.00
2	39	54801.95	54678.02	54840.95	-27239.50	0.53
3	59	54226.45	54038.97	54285.45	-26868.94	0.76
4	79	53828.62	53577.60	53907.62	-26587.22	0.76
5	99	53570.12	53255.55	53669.12	-26375.17	0.73
6	119	53485.21	53107.09	53604.21	-26249.91	0.71
7	139	**53440.82**	**52999.14**	**53579.82**	-26144.90	0.75
8	159	53509.89	53004.66	53668.89	-26096.63	0.74
9	179	53590.47	53021.69	53769.47	-26054.11	0.70
10	199	53678.02	53045.69	53877.02	-26015.08	0.74

Note: BIC = Bayesian Information Criterion; cAIC = Consistent Akaike Information Criterion; aBIC = sample size adjusted BIC; LL = Log-likelihood; Lower BIC, aBIC, and cAIC values suggest a better model fit, whereas higher entropy values suggest better classification accuracy. The lowest obtained value (*i.e.*, best model fit) for the classification statistics are shown in bold.

Table 3. 2. Summary of Binary Logistic Regression Analysis for Latent Subgroups Predicting Training Failure Outcomes.

Predictor	**B**	**SE**	**OR (95% CI)**	***p* value**
Older/Average	--	--	--	< .001
Psychologically Concerning	1.40	0.24	4.04 (2.52 – 6.49)	< .001
Intelligent	-0.77	0.16	0.46 (0.34 – 0.63)	< .001
Well-Regulated	-0.89	0.15	0.41 (0.30 – 0.55)	< .001
Extroverted	-0.51	0.15	0.60 (0.45 – 0.81)	.001
Introverted	-0.01	0.18	0.99 (0.70 – 1.40)	.96
Young/Average	-0.05	0.12	0.95 (0.76 – 1.19)	.66
Constant	-1.25	0.09		< .001

Notes: The final overall model was statistically significant, χ^2 (6, N = 3946) = 121.75, $p < .001$. The overall model explained between 3.0% (Cox and Snell R^2) and 5.0% (Nagelkerke R^2) of the variance in training outcome.

Appendix B – Figures

Figure 1.1. *Percentage of Males Across Latent Subgroups.*

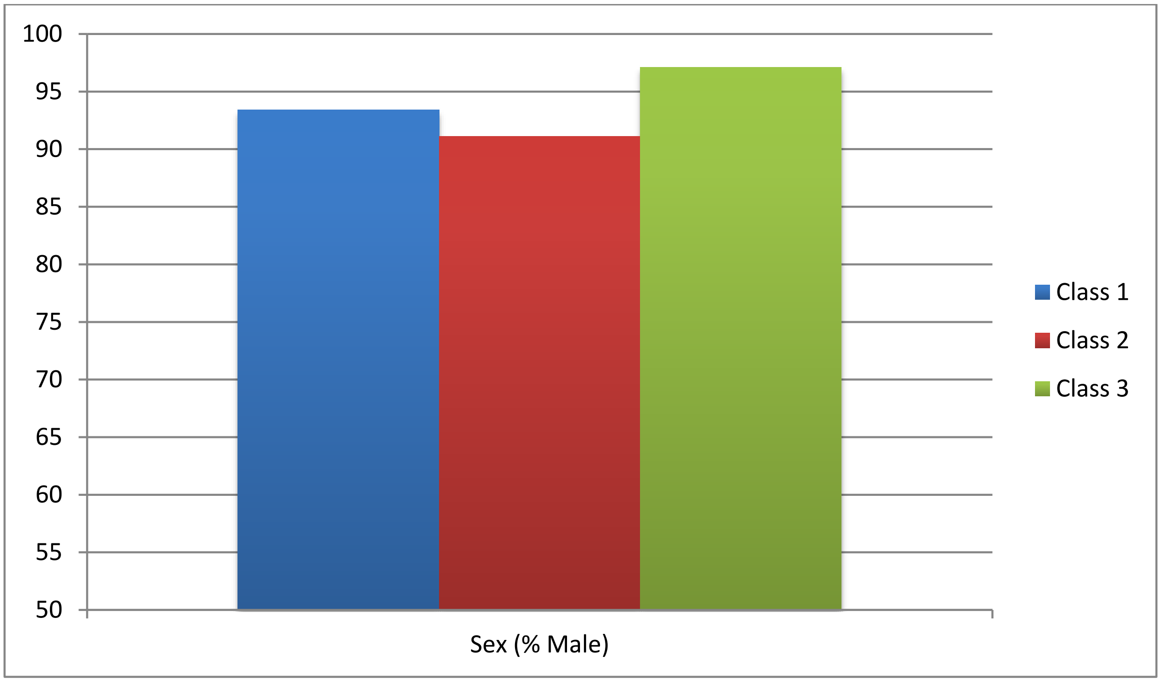


Figure 1.2. *Percentage of Individuals with Above Average Intelligence Across Latent Subgroups.*

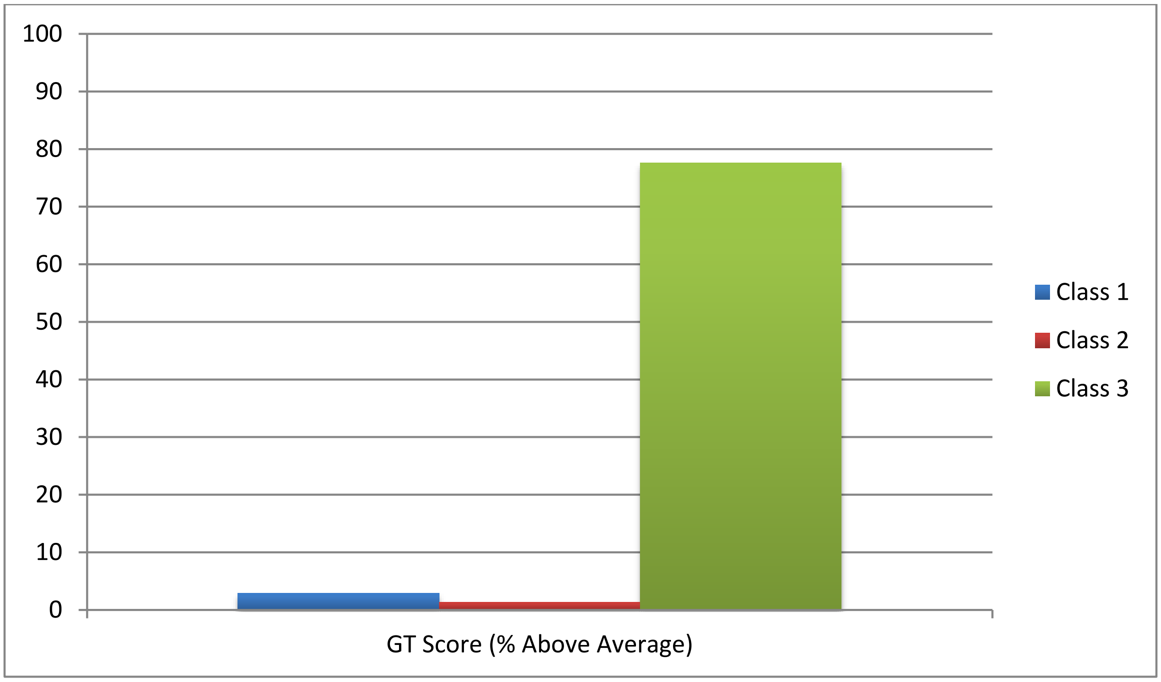


Figure 1.3. *Percentage of Individuals with First Class Physical Fitness Scores Across Latent Subgroups.*

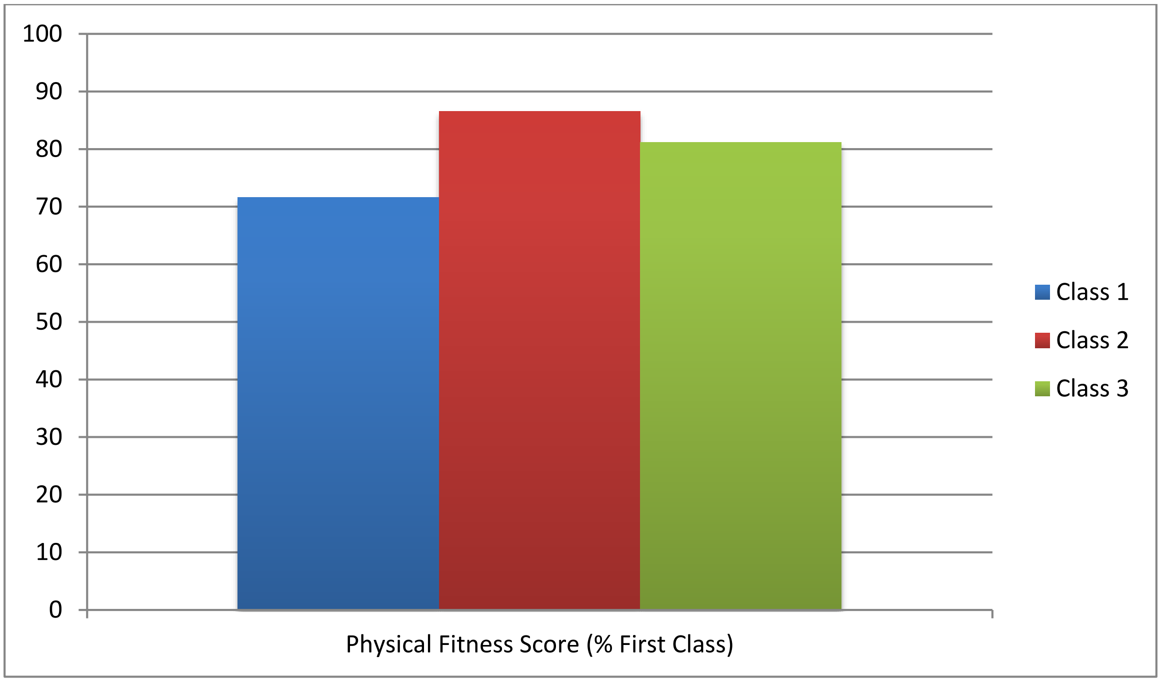


Figure 1.4. *Breakdown of Prior Disciplinary Counseling Statements Across Latent Subgroups.*

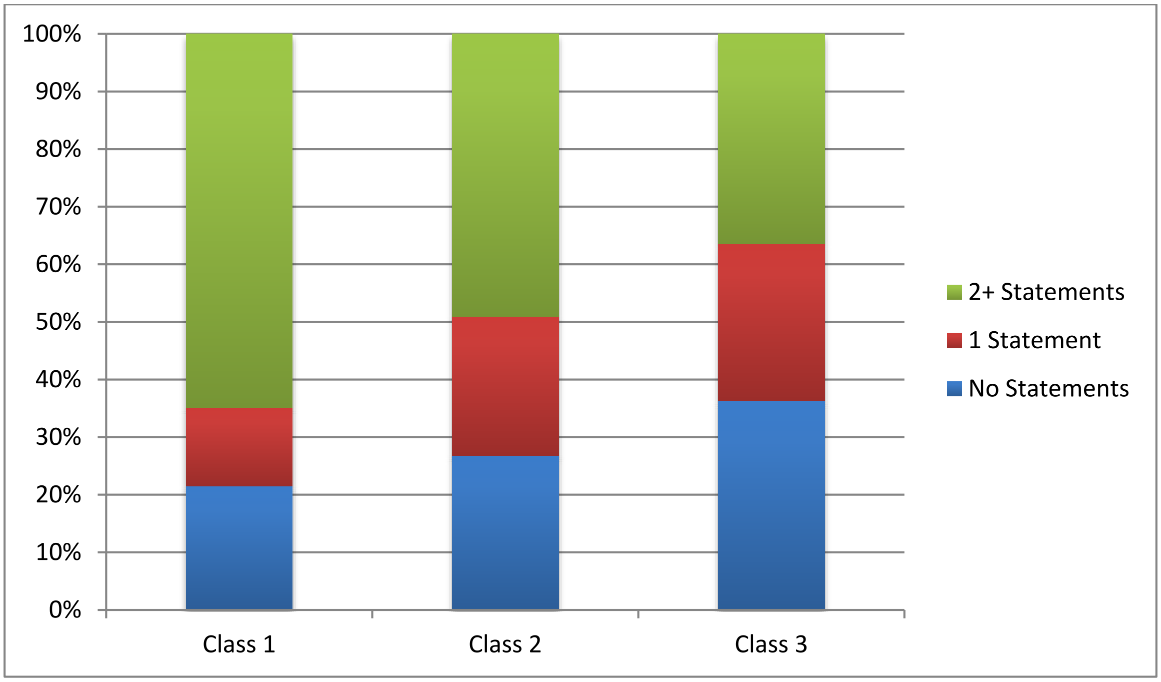


Figure 1.5. *Percentage of Individuals with Prior Mental Health Treatment Across Latent Subgroups.*

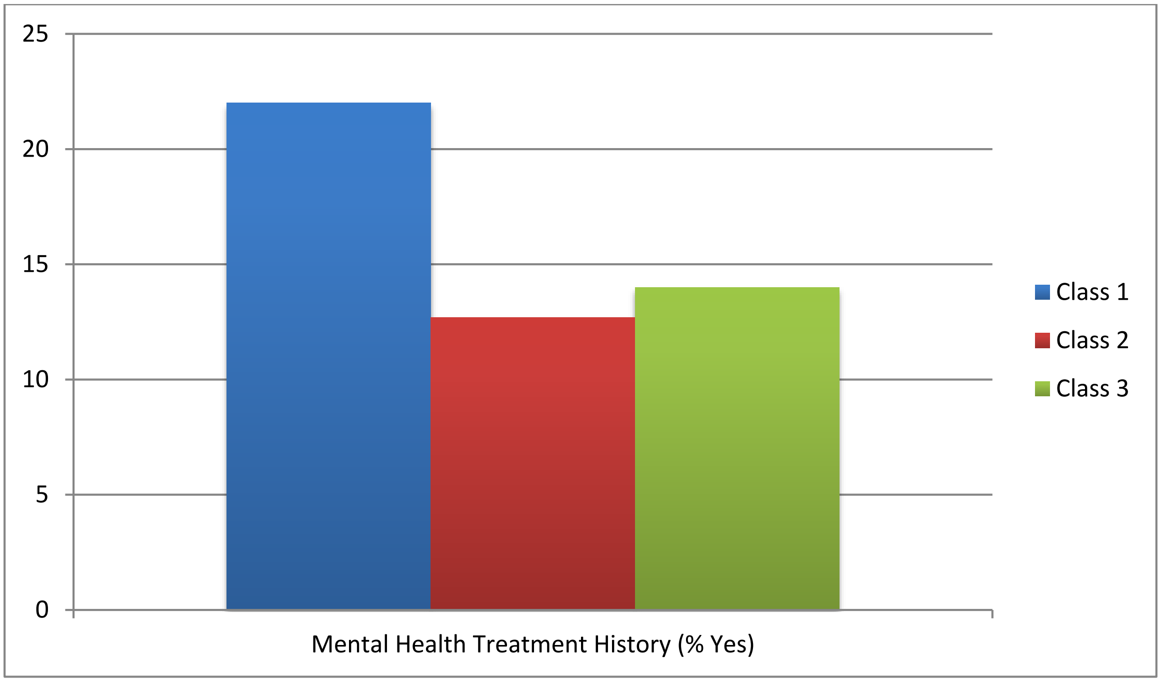
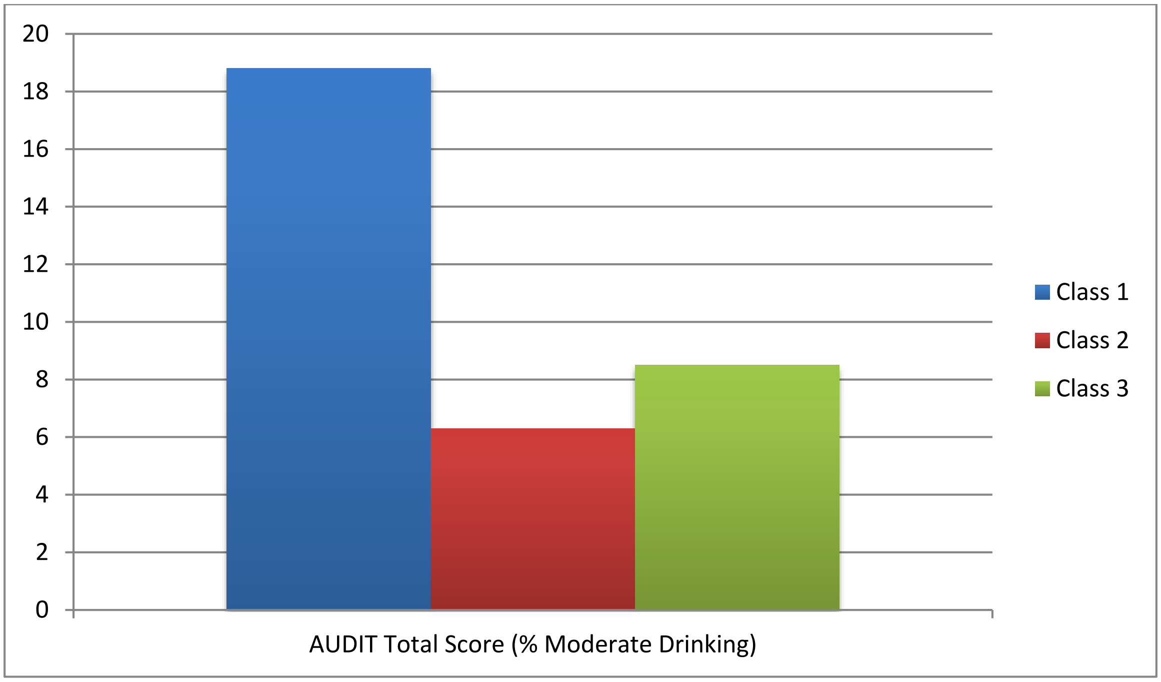
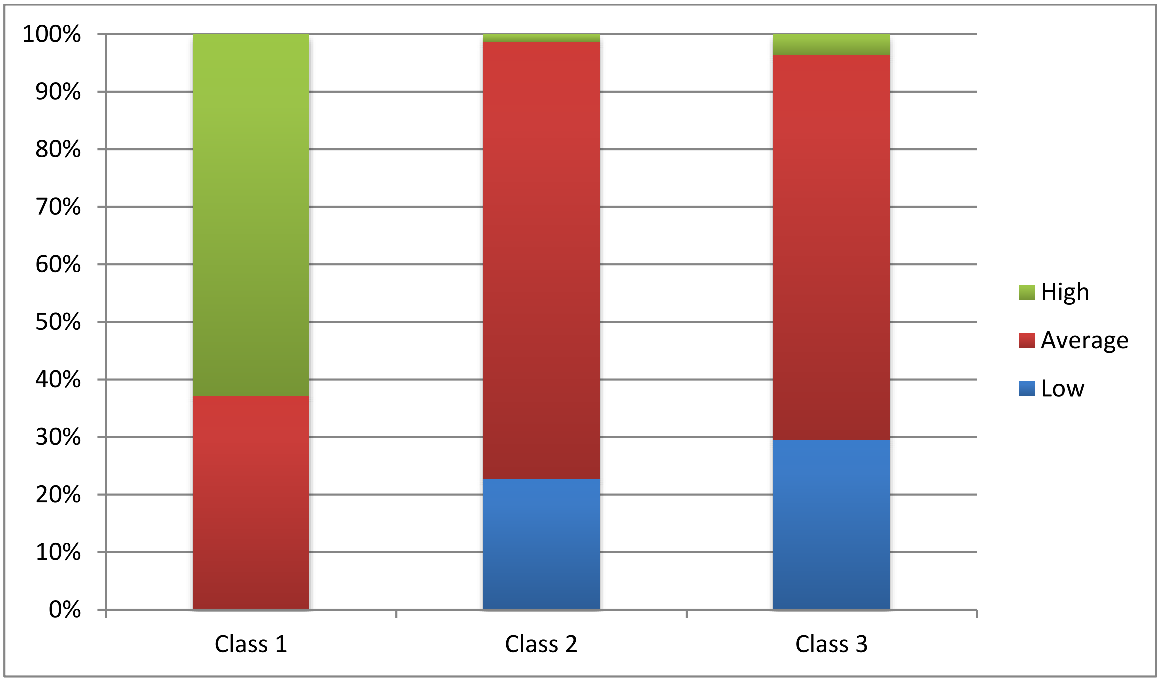


Figure 1.6. *Percentage of Individuals Reporting at Least Moderate Alcohol Use Across Latent Subgroups.*



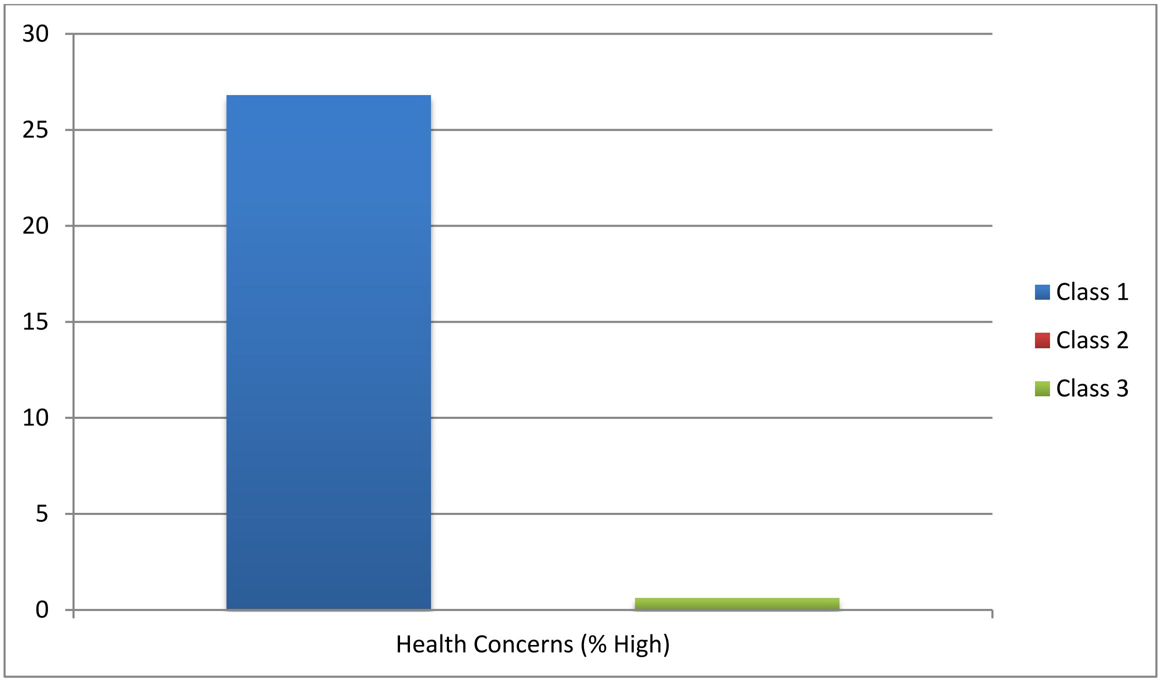
Note: AUDIT = Alcohol Use Disorders Identification Test.

Figure 1.7. *Breakdown of 16PF Alienation/Perceptual Distortion Scale Scores Across Latent Subgroups.*



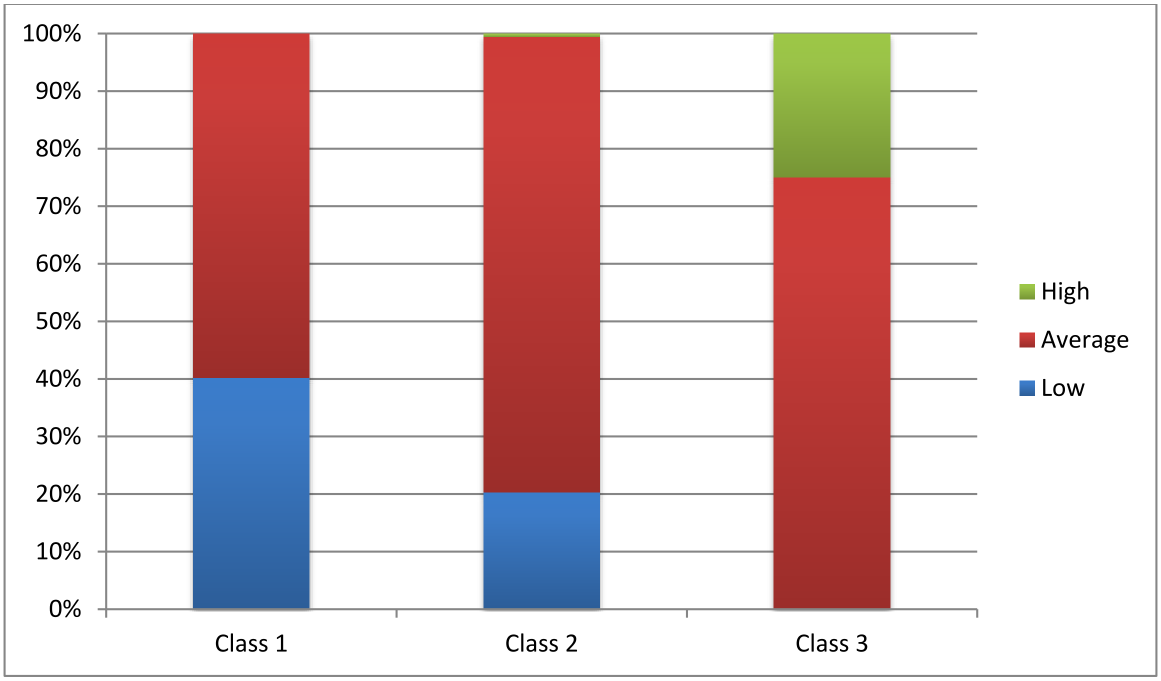
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 1.8. *Percentage of Individuals with High 16PF Health Concerns Scale Scores Across Latent Subgroups.*



Note: 16PF = 16 Personality Factor Questionnaire.

Figure 1.9. *Breakdown of 16PF Reasoning Scale Scores Across Latent Subgroups.*



Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.1. *Breakdown of Military Rank Across Latent Subgroups.*

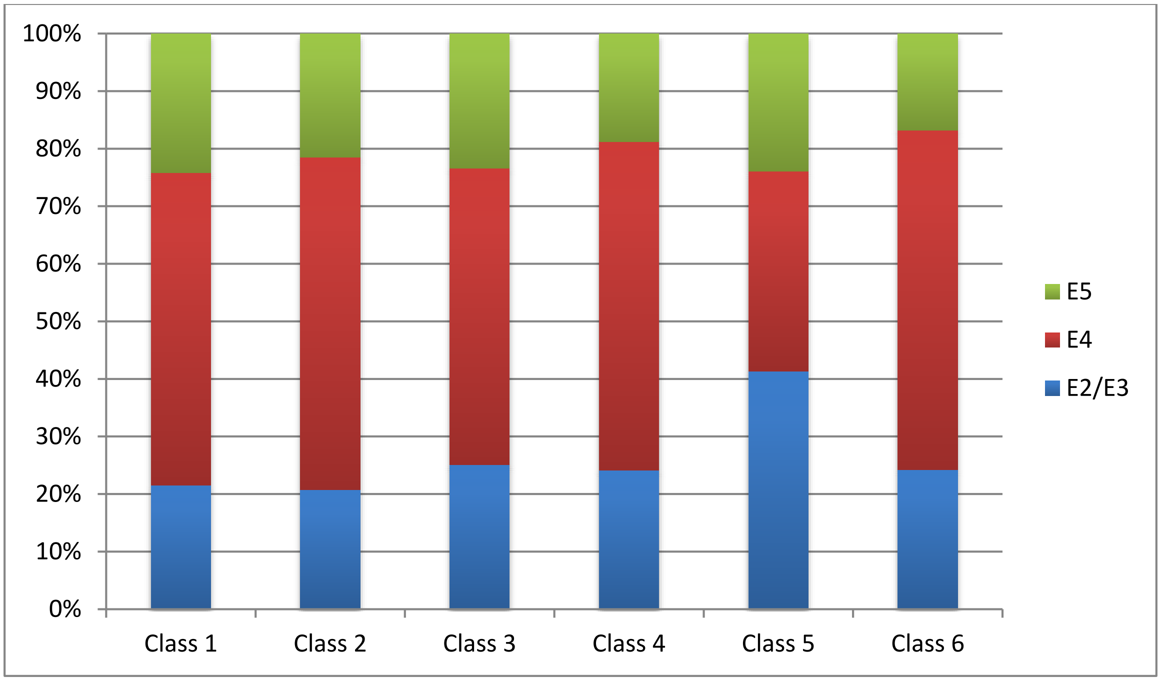


Figure 2.2. *Percentage of Individuals with Above Average Intelligence Across Latent Subgroups.*

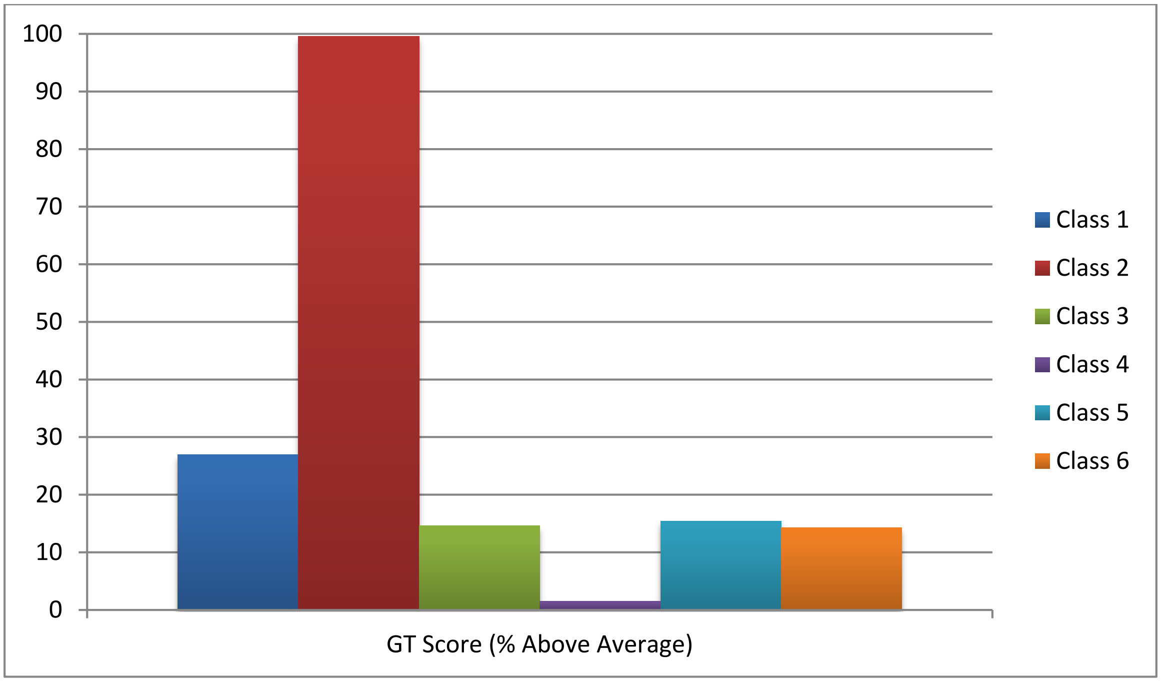
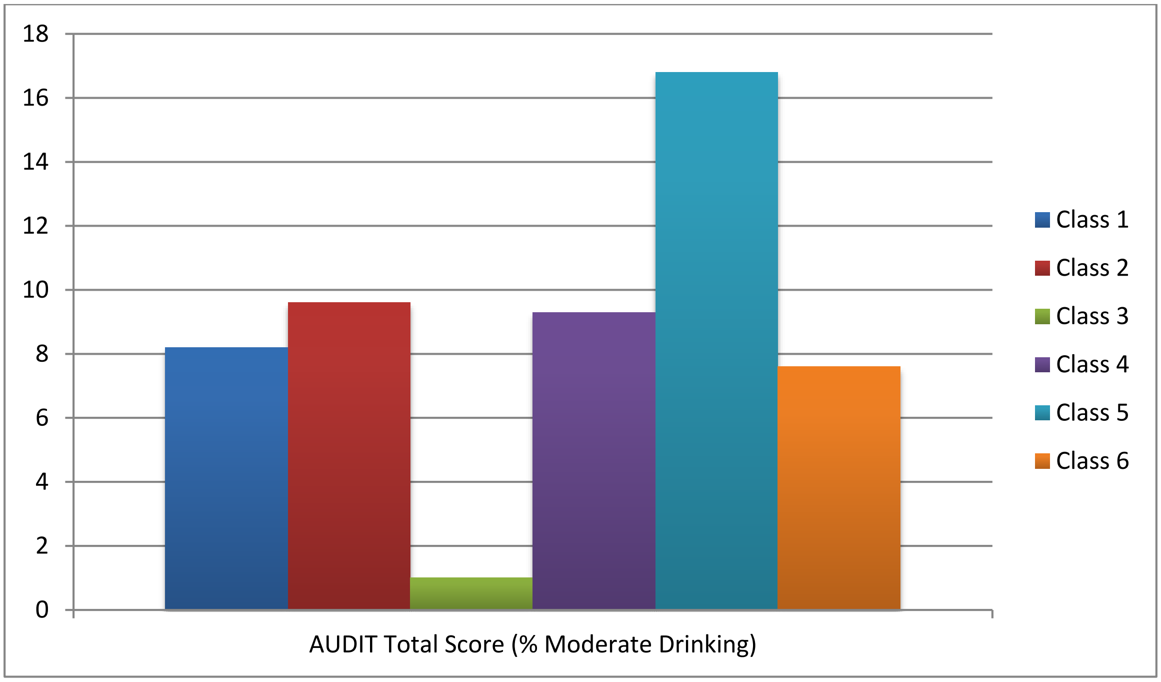
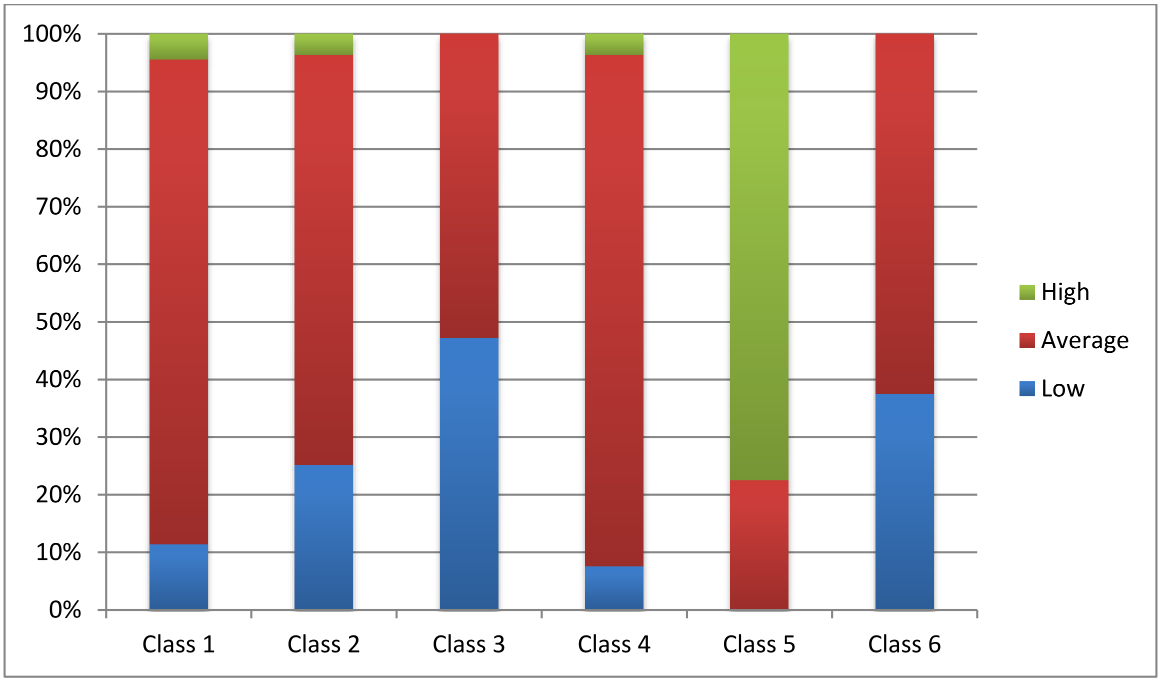


Figure 2.3. *Percentage of Individuals Reporting Moderate to Concerning Alcohol Use Patterns Across Latent Subgroups.*



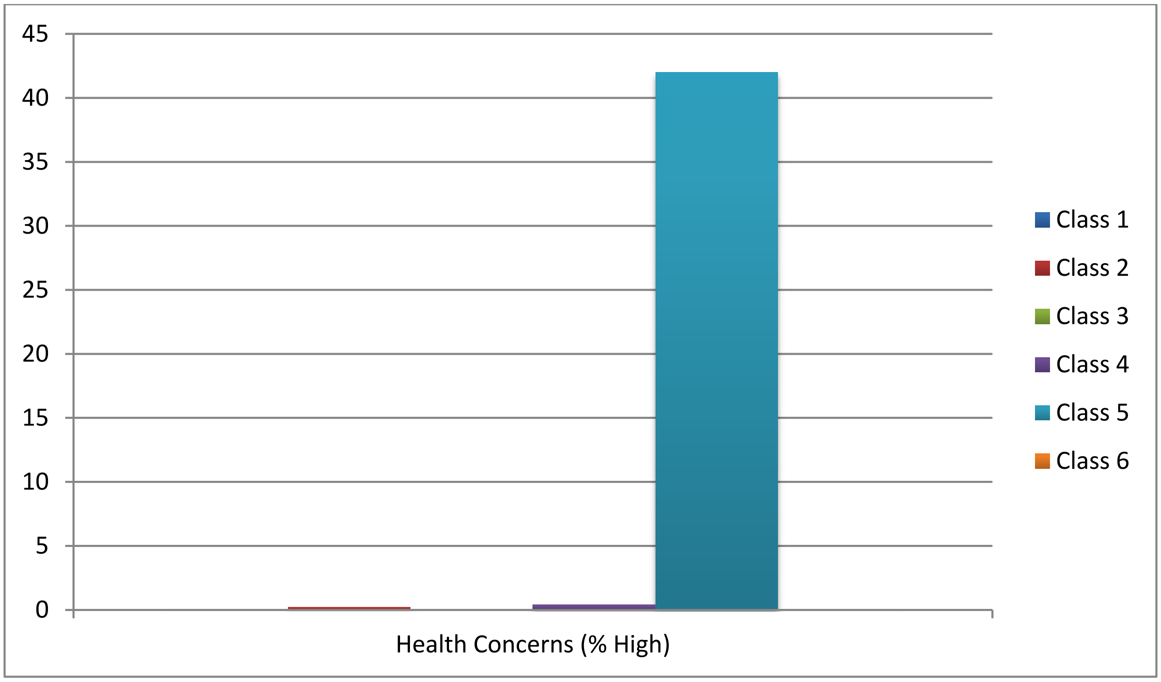
Note: AUDIT = Alcohol Use Disorders Identification Test.

Figure 2.4. *Breakdown of 16PF Alienation/Perceptual Distortion Scale Scores Across Latent Subgroups.*



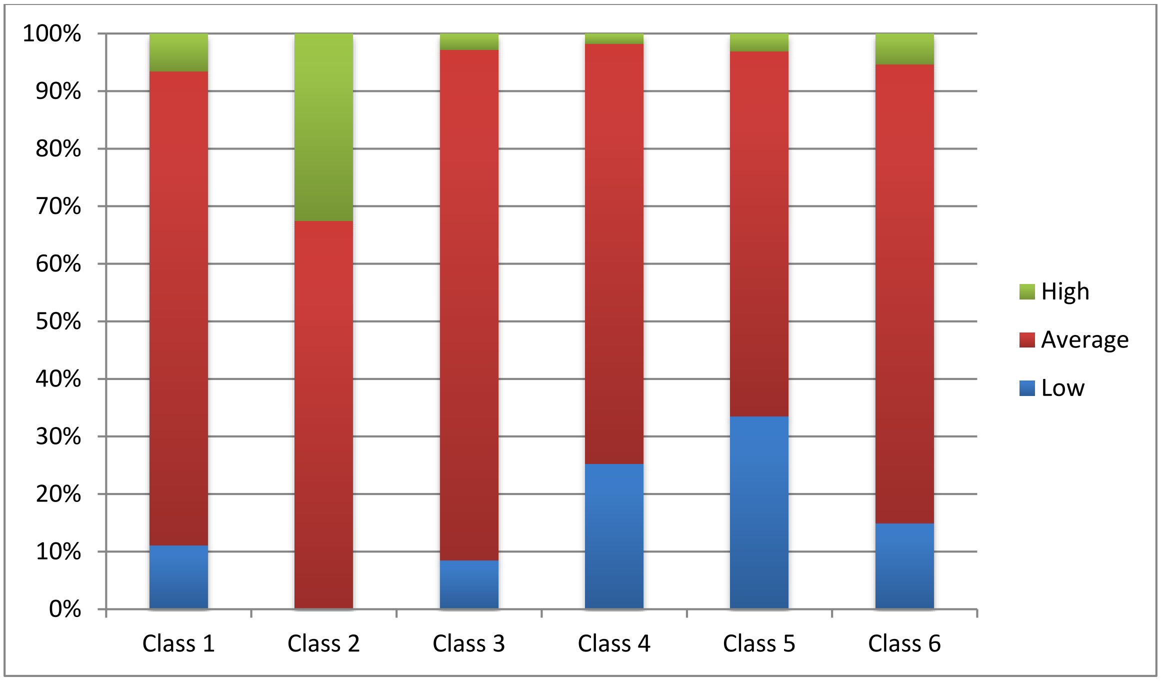
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.5. *Percentage of Individuals with High 16PF Health Concerns Scale Scores Across Latent Subgroups.*



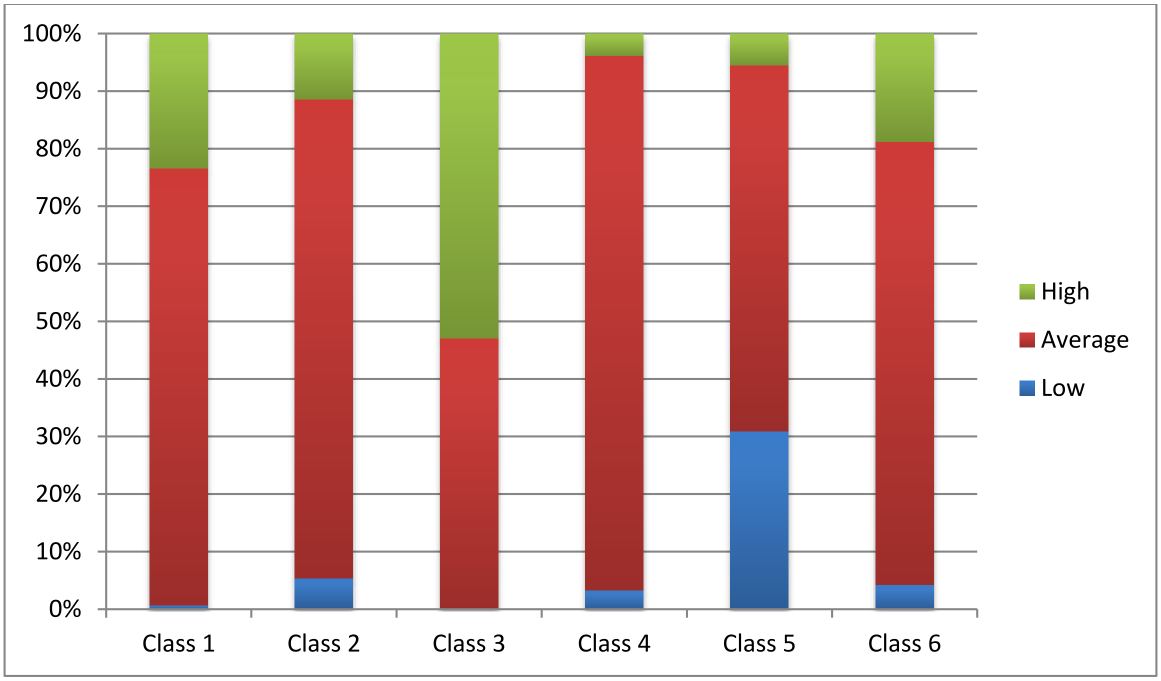
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.6. *Breakdown of 16PF Reasoning Scale Scores Across Latent Subgroups.*



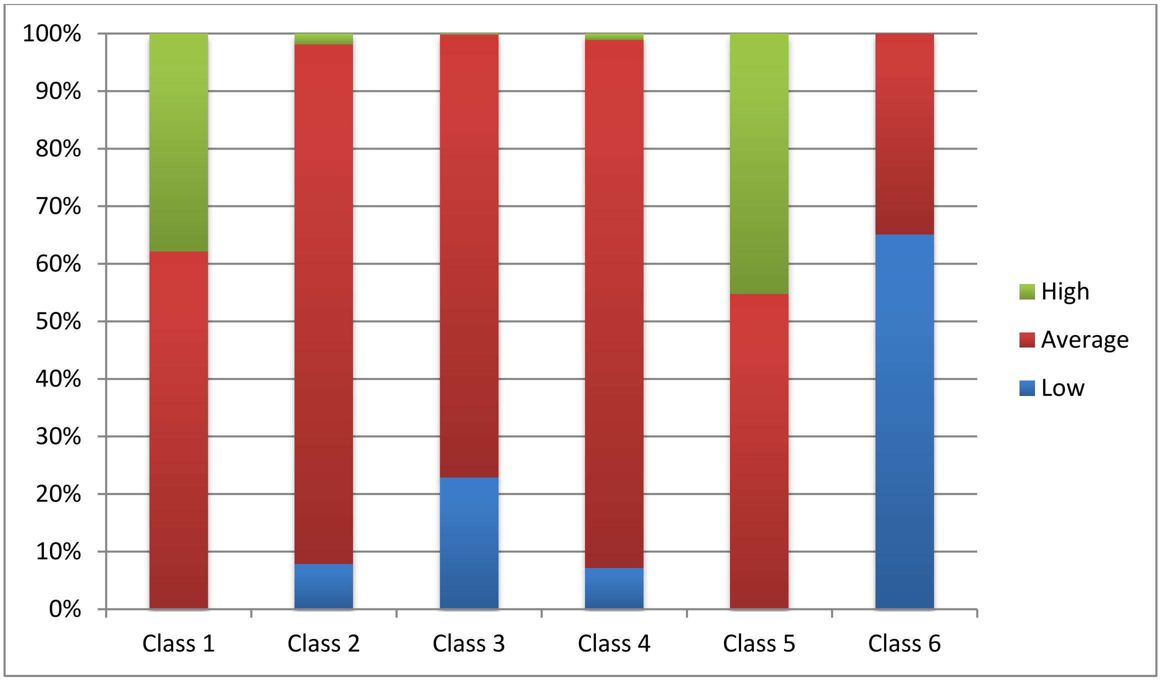
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.7. *Breakdown of 16PF Self-Control Scale Scores Across Latent Subgroups.*



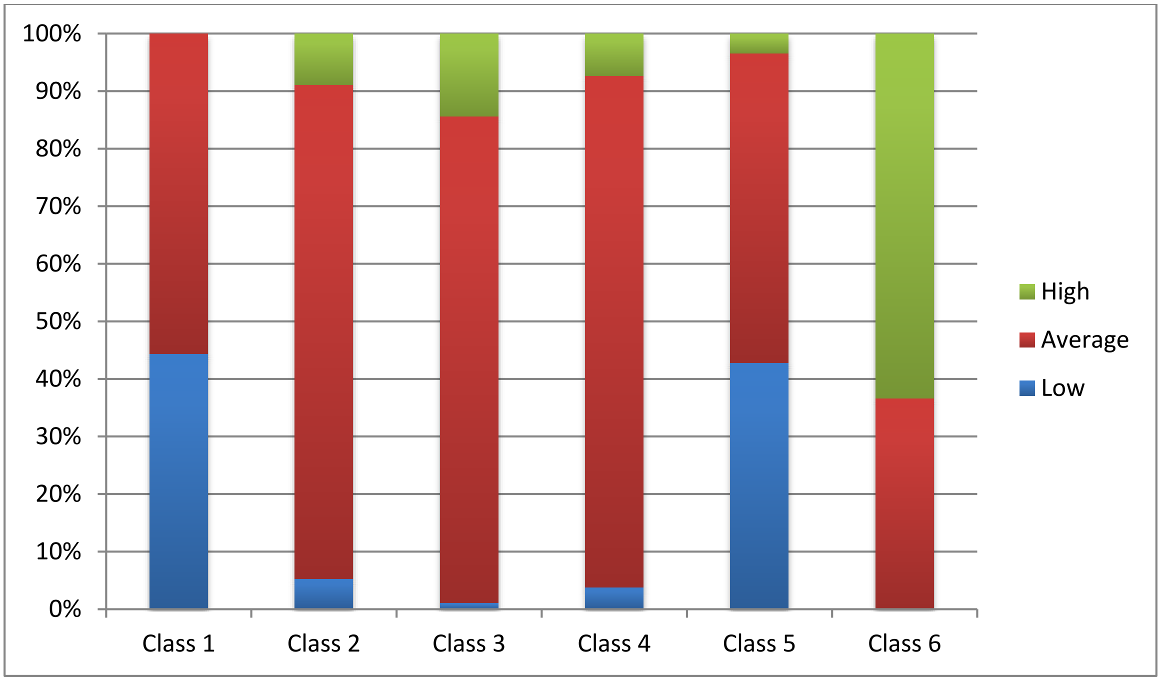
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.8. *Breakdown of 16PF Self-Reliance Scale Scores Across Latent Subgroups.*



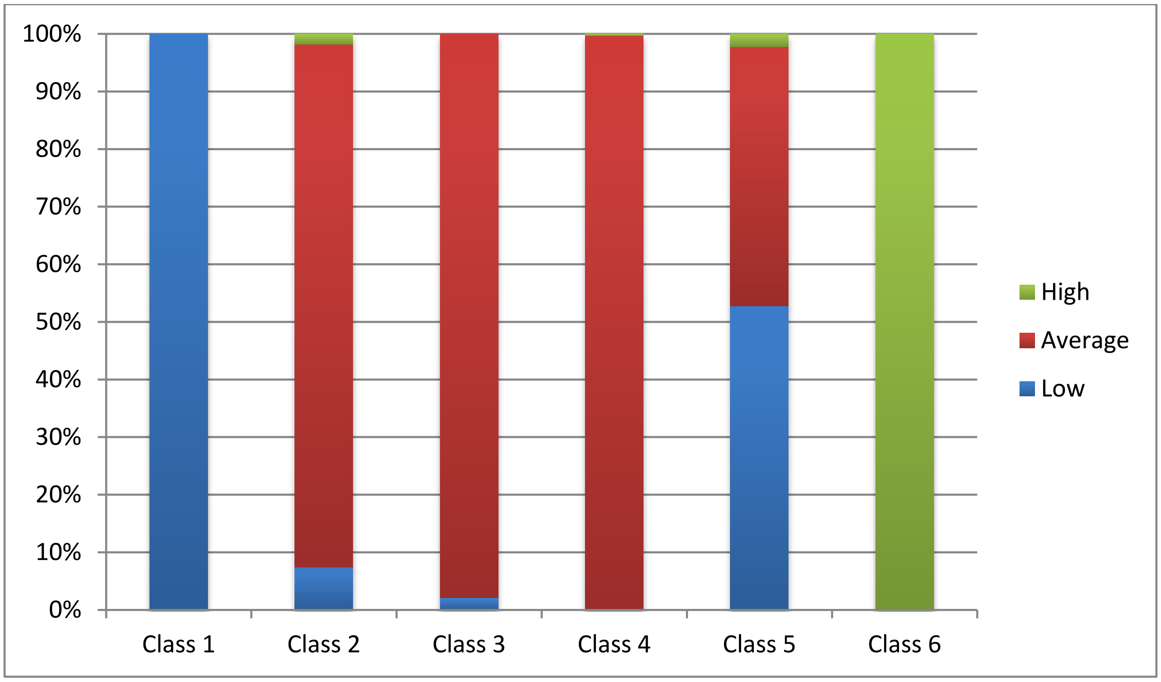
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.9. *Breakdown of 16PF Social Boldness Scale Scores Across Latent Subgroups.*



Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.10. *Breakdown of 16PF Interpersonal Relations Scale Scores Across Latent Subgroups.*



Note: 16PF = 16 Personality Factor Questionnaire.

Figure 2.11. *Percentage of Individuals with First Class Physical Fitness Scores Across Latent Subgroups.*

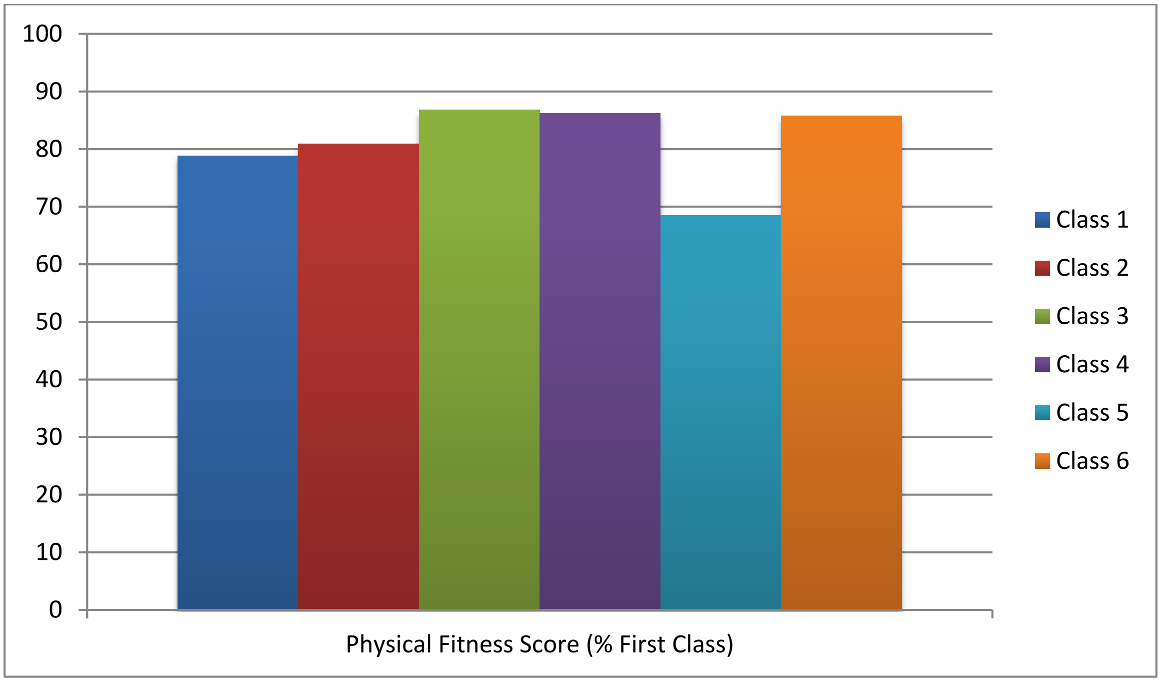


Figure 2.12. *Breakdown of Prior Disciplinary Counseling Statements Across Latent Subgroups.*

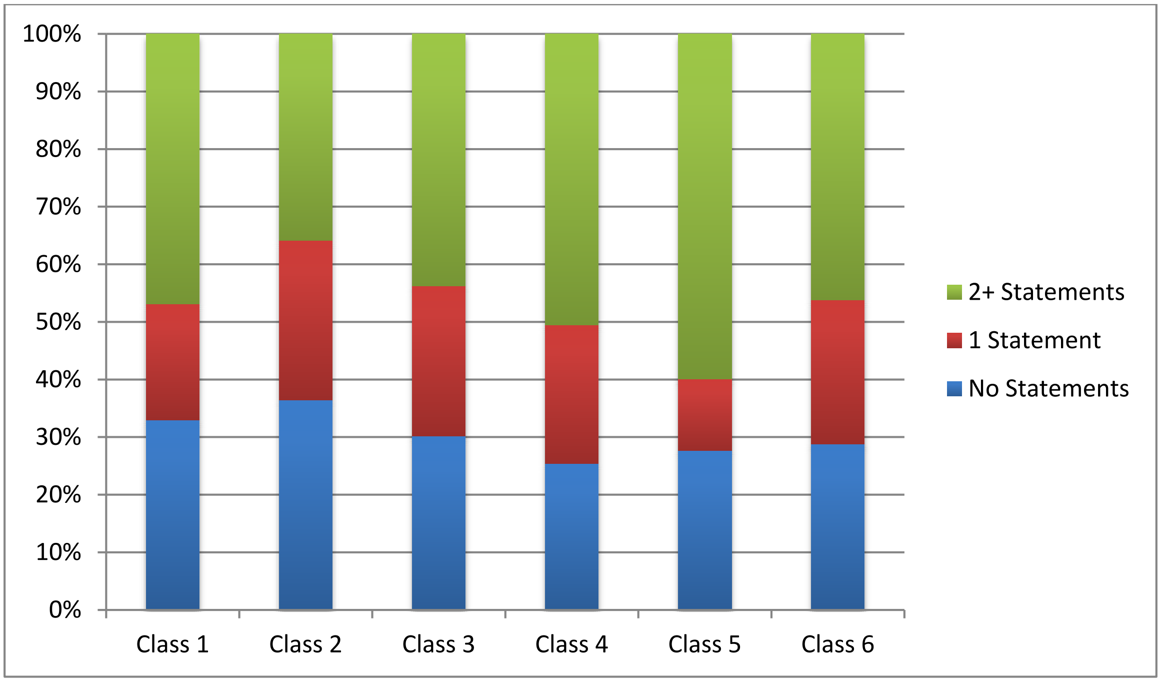


Figure 3.1. *Breakdown of Age Across Latent Subgroups*.

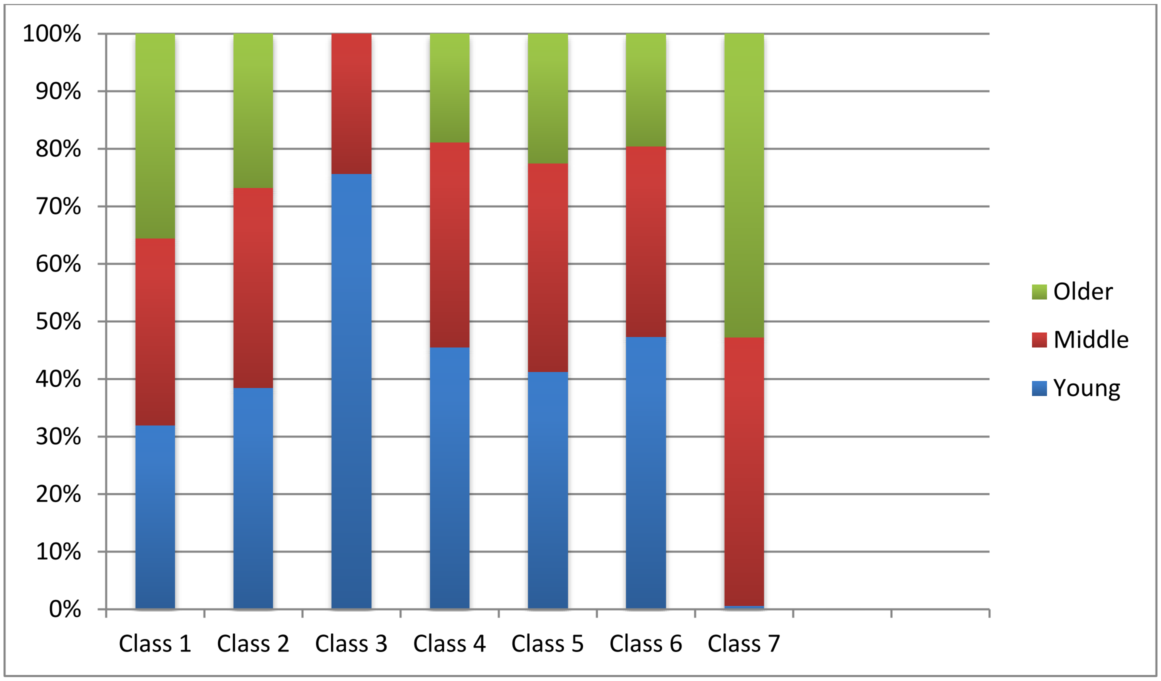


Figure 3.2. *Breakdown of Prior Combat Deployments Across Latent Subgroups*.

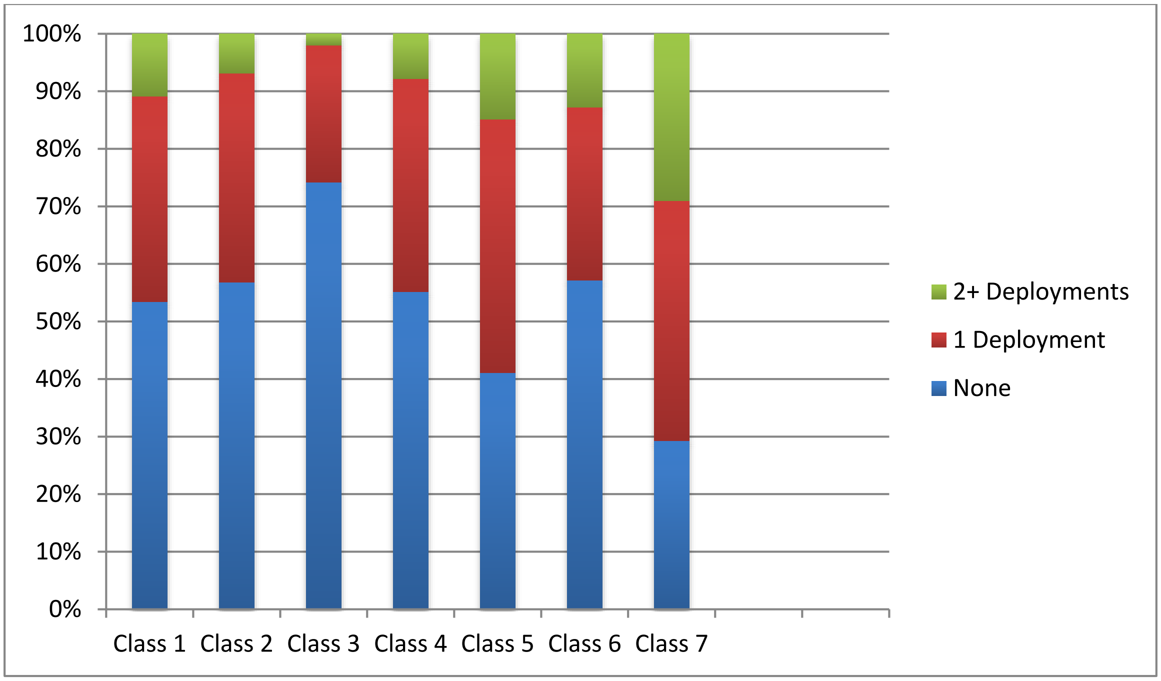


Figure 3.3. *Percentage of Individuals with Above Average Intelligence Across Latent Subgroups.*

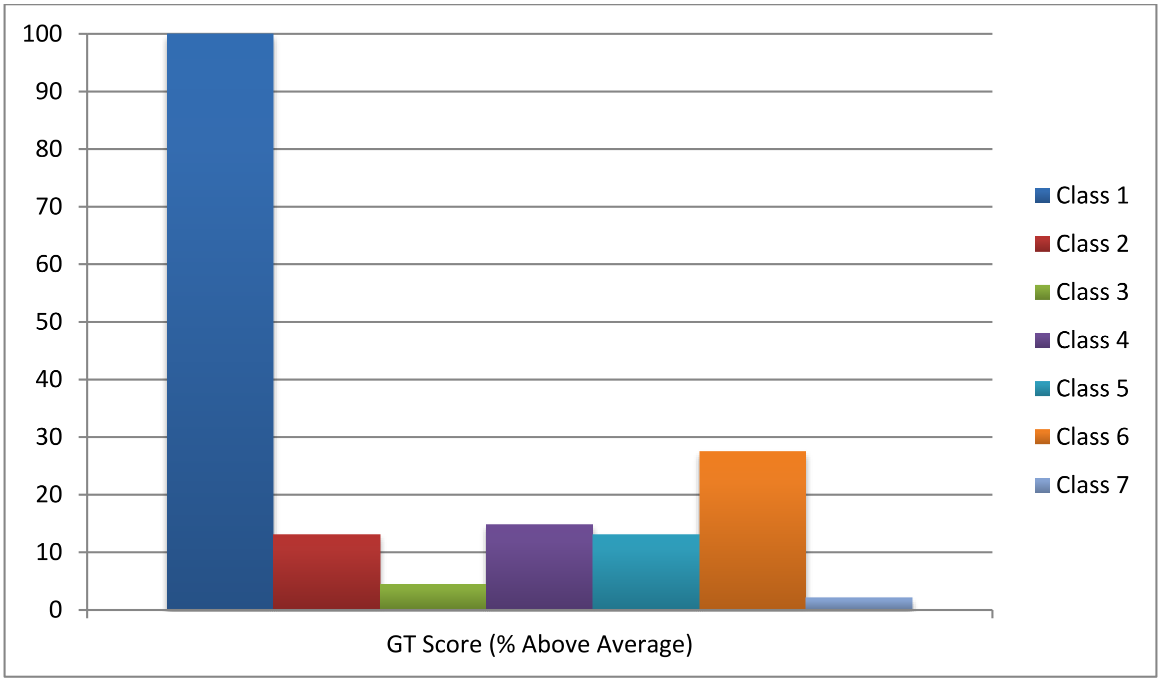
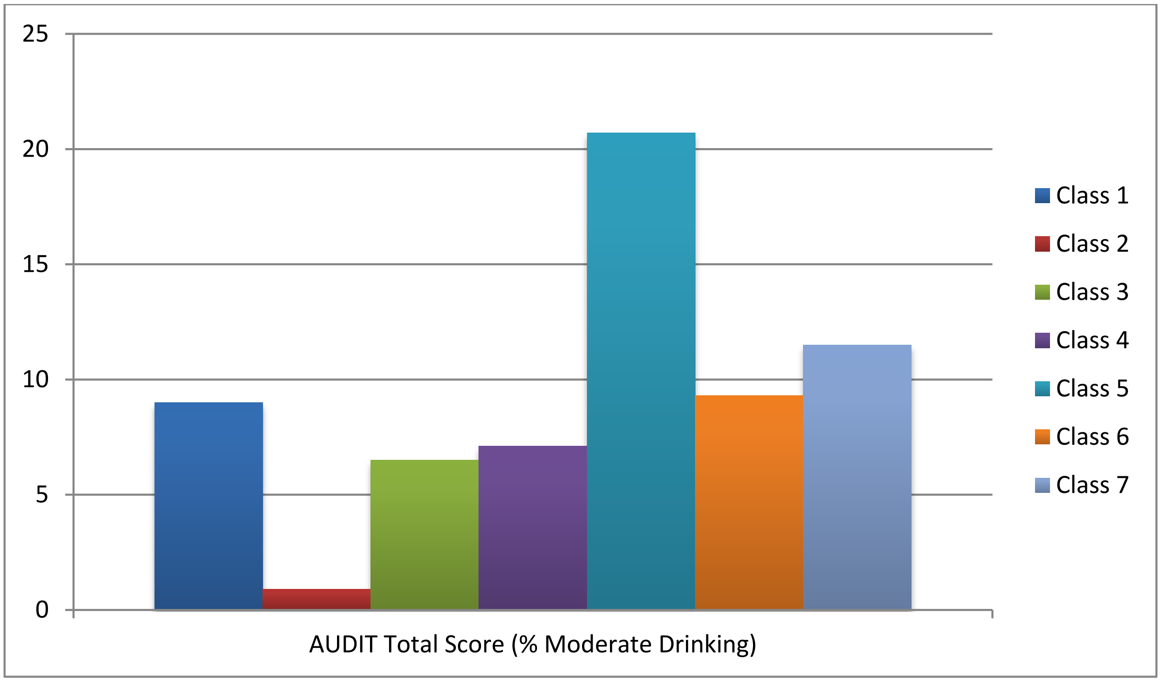
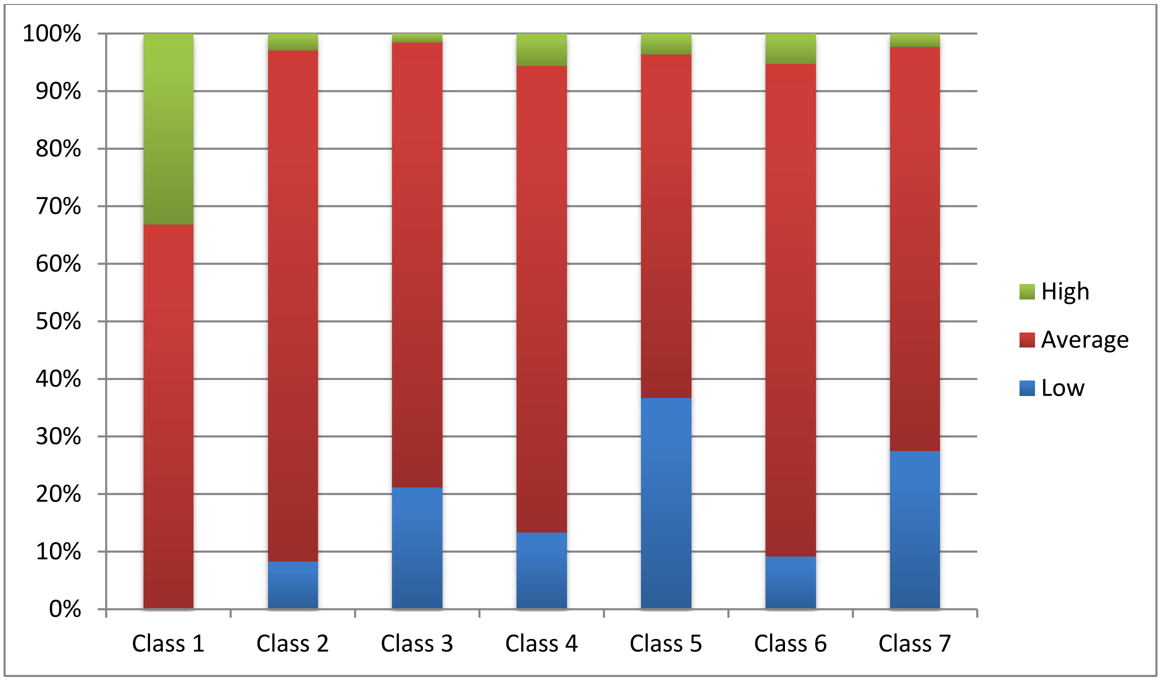


Figure 3.4. *Percentage of Individuals Reporting Moderate to Concerning Alcohol Use Patterns Across Latent Subgroups.*



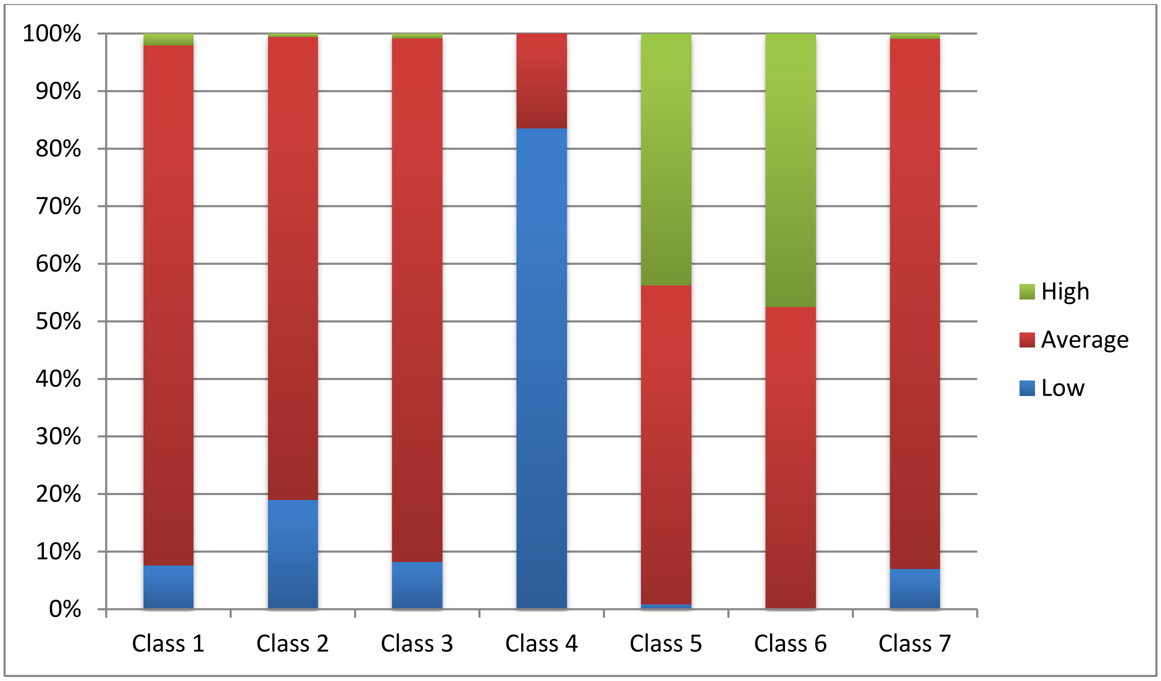
Note: AUDIT = Alcohol Use Disorders Identification Test.

Figure 3.5 *Breakdown of 16PF Reasoning Scale Scores Across Latent Subgroups.*



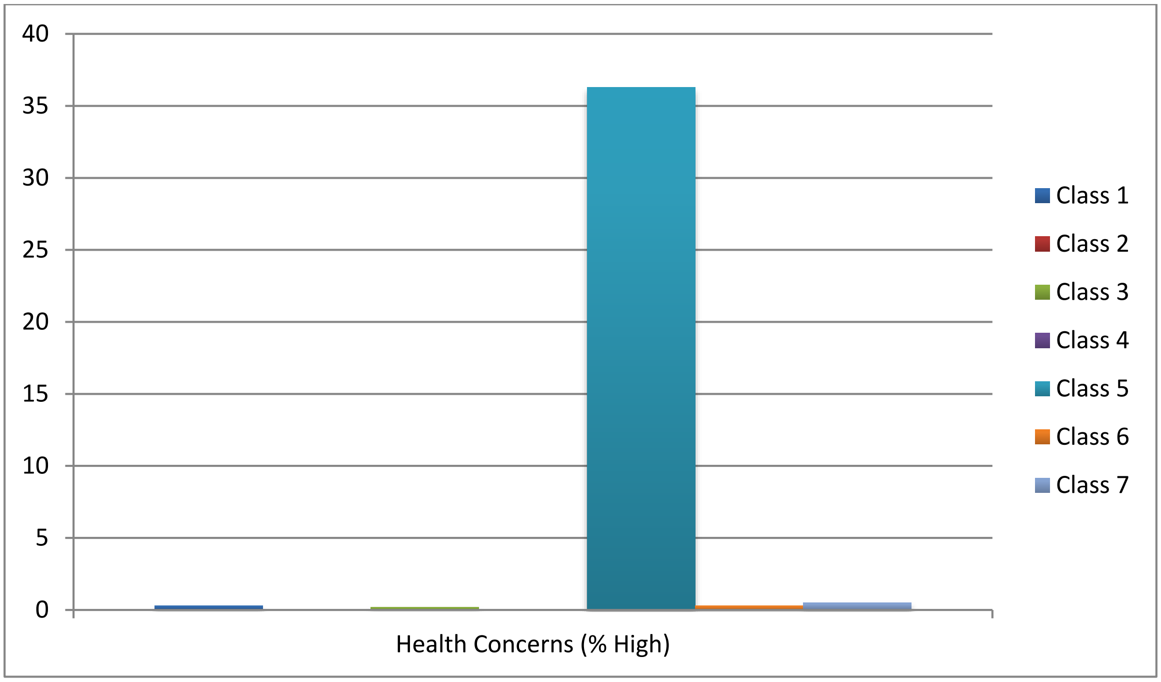
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.6 *Breakdown of 16PF Self-Reliance Scale Scores Across Latent Subgroups.*



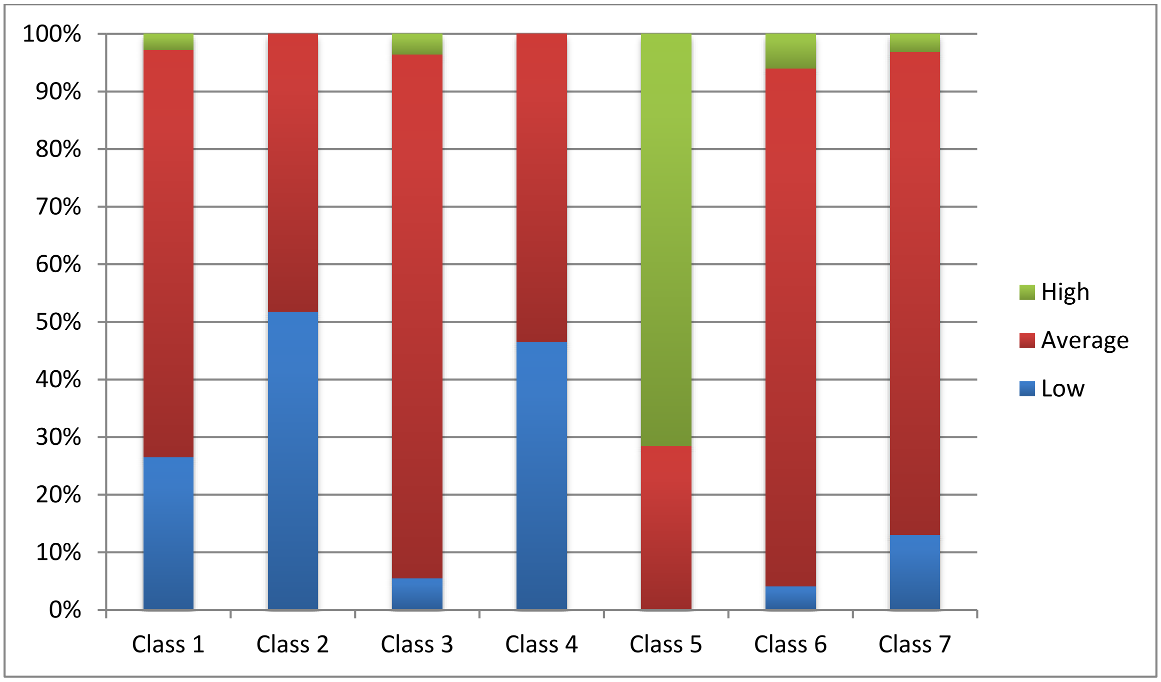
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.7. *Percentage of Individuals with High 16PF Health Concerns Scale Scores Across Latent Subgroups.*



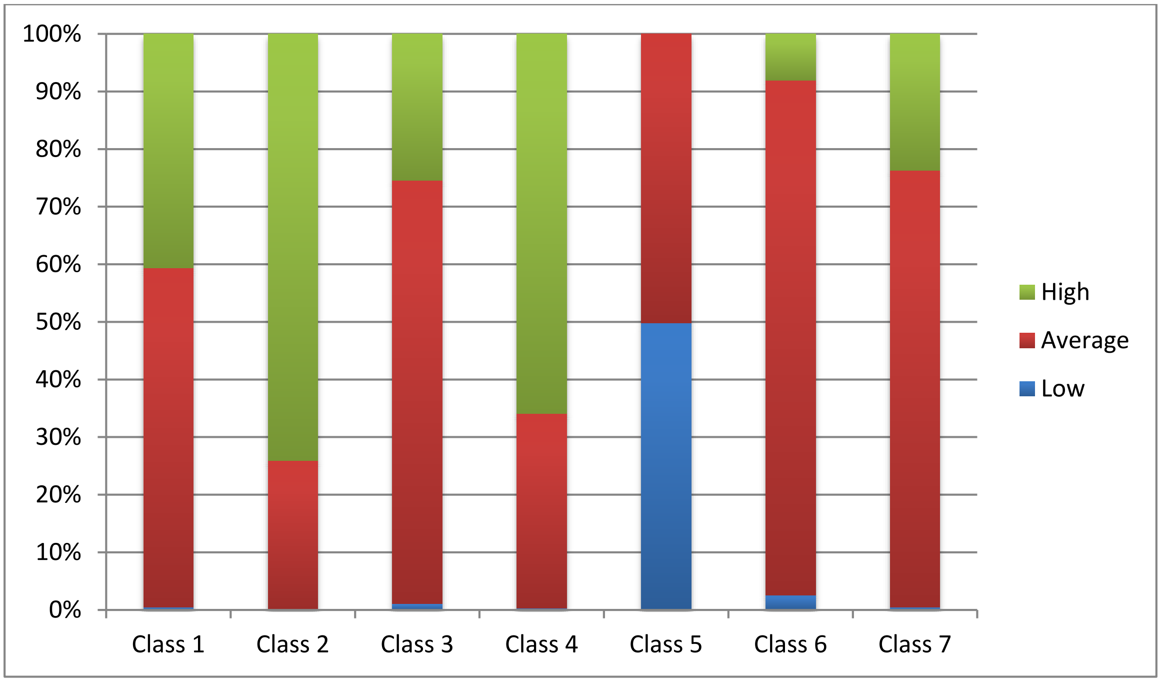
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.8 *Breakdown of 16PF Alienation/Perceptual Distortion Scale Scores Across Latent Subgroups.*



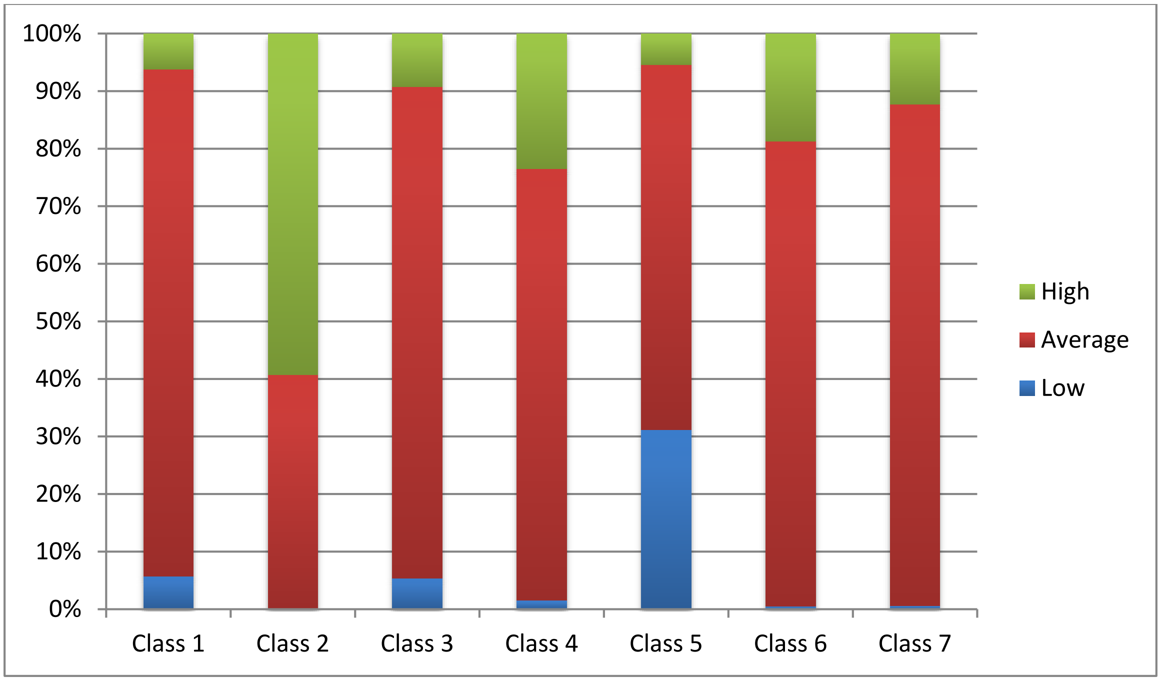
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.9 *Breakdown of 16PF Emotional Stability Scale Scores Across Latent Subgroups.*



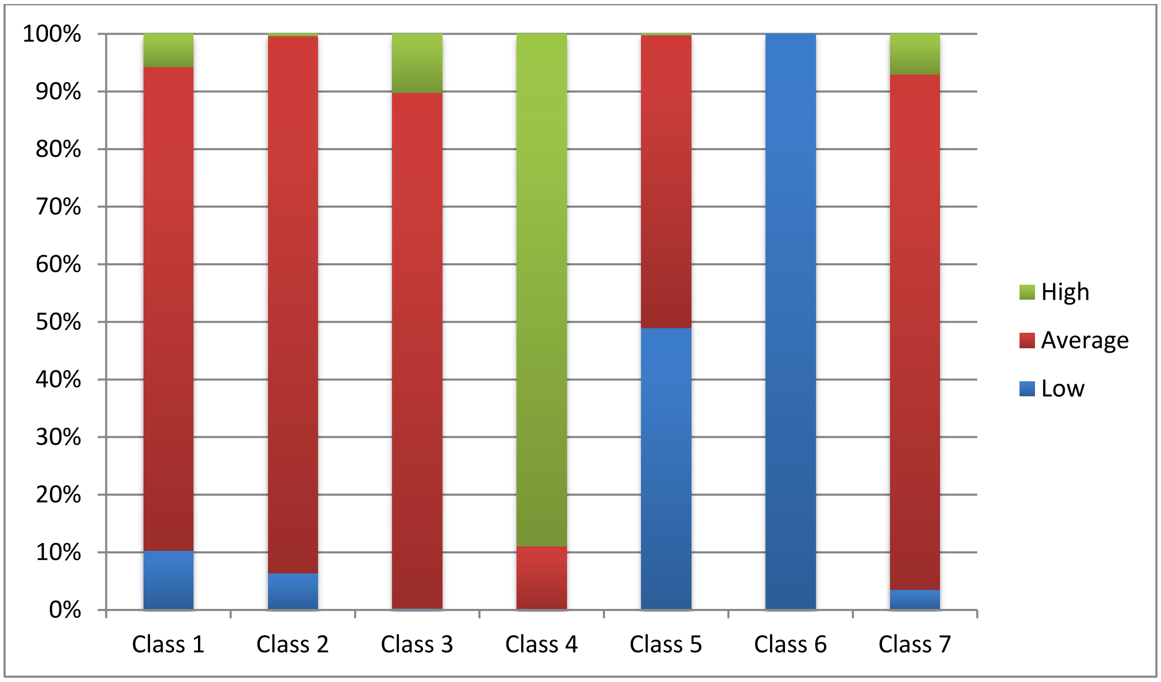
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.10 *Breakdown of 16PF Self-Control Scale Scores Across Latent Subgroups.*



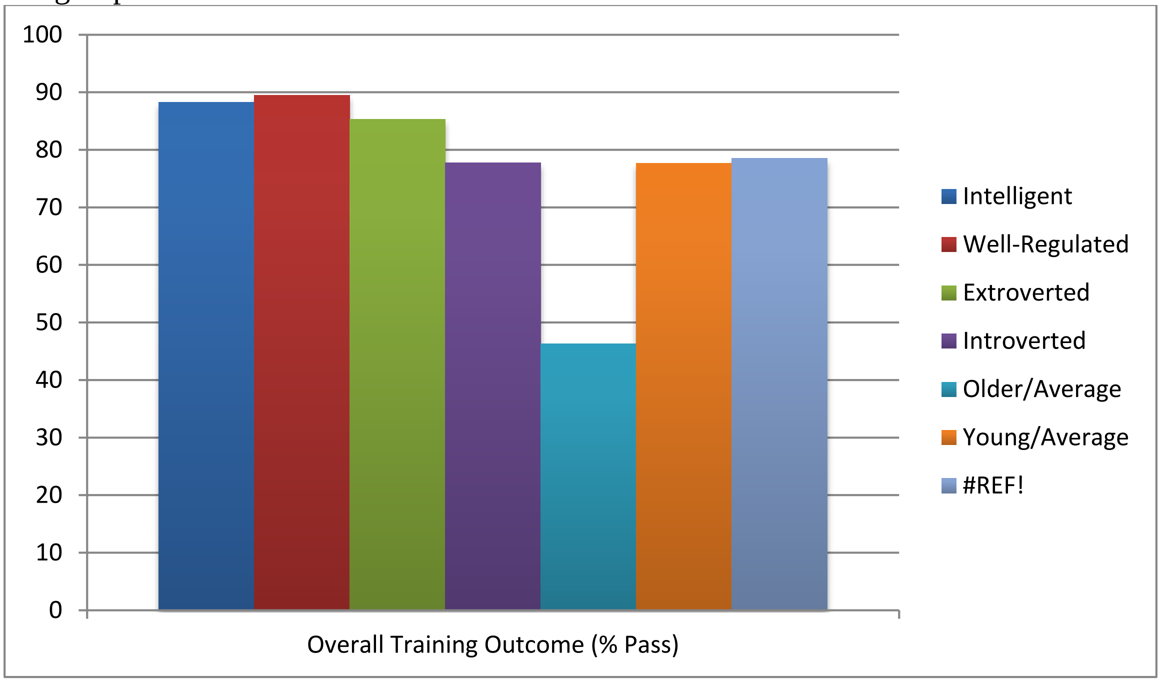
Note: 16PF = 16 Personality Factor Questionnaire.

Figure 3.11 *Breakdown of 16PF Interpersonal Relations Scale Scores Across Latent Subgroups.*



Note: 16PF = 16 Personality Factor Questionnaire.

Figure 4.2. Comparison of Training Program Completion Rates Across Latent Subgroups.



Appendix C – IRB Exemption

UNIFORMED SERVICES UNIVERSITY
of the Health Sciences

OFFICE OF RESEARCH
4301 JONES BRIDGE ROAD
BETHESDA, MAYLAND 20814
PHONE: (301) 295-3303; FAX: (301) 295-6771

NOTICE OF PROJECT APPROVAL

Change Number: Original

VPR Site Number:	T0-MPS-72-3551-01
Principal Investigator:	Finton, Brendan (MPS-72)
Department:	Medical and Clinical Psychiatry
Project Type:	Student
Project Title:	Predicting Attrition in a Military Special Program Training Command
Project Period:	5/26/2015 to 12/31/2015

Assurance and Progress Report Information:

Name	Sup	Approval Type	Status	Approved On	Forms Received
Progress Report	0			To be Submitted	N/A

Remarks:
This Notice of Project Approval has been reviewed and approved. Please remember that you must submit a final Progress Report (Form 3210) upon completion of this project.

Questions regarding this approval should be directed to the following person in the Office of Research:
Ronda Dudley, (301) 295-9818.

for 5/27/15

Toya V. Randolph, Ph.D, MSPH — Date
Acting Vice President for Research
Uniformed Services University of the Health Sciences

cc: Finton, Brendan (MPS-72)
Vernell Shaw
File
Neil Grunberg
Andrew Waters

UNIFORMED SERVICES UNIVERSITY OF THE HEALTH SCIENCES
4301 JONES BRIDGE ROAD
BETHESDA, MARYLAND 20814-4712
http://www.usuhs.mil

MAY 26, 2015

MEMORANDUM FOR LT BRENDAN FINTON, MS, USA, MEDICAL AND CLINICAL PSYCHOLOGY

SUBJECT: Uniformed Services University Institutional Review Board (FWA 0001628; DoD Assurance P60001) Determination of Non-Human Subjects Research

Research protocol TO-72-3551 entitled "*Predicting Attrition in a Military Special Program Training Command*" has been reviewed by the Uniformed Services University's Human Research Protections Program Office and determined not to meet the criteria defining human subjects research at 32 CFR 219.102, and applicable DoD policy. As such, this protocol does not require Institutional Review Board (IRB) review.

This protocol involves the receipt and analysis of previously collected, de-identified training performance and outcome data from the

The goal of this program evaluation project is to assist the

in evaluating their assessment and selection program by identifying those factors related to and/or predictive of successful Marine Security Guards.

Should your project data sources or methodology change, please contact this office before you begin any new phase of your work so that we may review it with you. Otherwise we cannot ensure that you will be in compliance with all applicable human subject research regulations. The IRB staff is a key resource that is available to assist you to ensure that you are in compliance with applicable human research regulations.

If you have questions regarding this IRB action, or questions of a more general nature concerning human participation in research, please contact the undersigned at micah.stretch@usuhs.edu or (301) 295-9534.

Micah R. Stretch, M.A., J.D.
Exemption Determination Official

cc: Chair, MPS
VPR
File

Learning to Care for Those in Harm's Way

Appendix D – Selected Measures

PTSD Checklist – Military (PCL-M)

PCL-M

INSTRUCTIONS: Below is a list of problems and complaints that veterans sometimes have in response to stressful military experiences. Please read each one carefully, then circle one of the numbers to the right to indicate how much you have been bothered by that problem in the past month.

		Not at all	A little bit	Moderately	Quite a bit	Extremely
1.	Repeated, disturbing *memories, thoughts,* or *images* of a stressful military experience?	1	2	3	4	5
2.	Repeated, disturbing *dreams* of a stressful military experience?	1	2	3	4	5
3.	Suddenly *acting* or *feeling* as if a stressful military experience *were happening again* (as if you were reliving it)?	1	2	3	4	5
4.	Feeling *very upset* when *something reminded you* of a stressful military experience?	1	2	3	4	5
5.	Having *physical reactions* (e.g., heart pounding, trouble breathing, sweating) when *something reminded you* of a stressful military experience?	1	2	3	4	5
6.	Avoiding *thinking about* or *talking about* a stressful military experience or avoiding *having feelings* related to it?	1	2	3	4	5
7.	Avoiding *activities* or *situations* because *they reminded you* of a stressful military experience?	1	2	3	4	5
8.	Trouble *remembering important parts* of a stressful military experience?	1	2	3	4	5
9.	*Loss of interest* in activities that you used to enjoy?	1	2	3	4	5
10.	Feeling *distant* or *cut off* from other people?	1	2	3	4	5
11.	Feeling *emotionally numb* or being unable to have loving feelings for those close to you?	1	2	3	4	5
12.	Feeling as if your *future* will somehow be *cut short*?	1	2	3	4	5
13.	Trouble *falling* or *staying asleep*?	1	2	3	4	5
14.	Feeling *irritable* or having *angry outbursts*?	1	2	3	4	5
15.	Having *difficulty concentrating*?	1	2	3	4	5
16.	Being "*super-alert*" or watchful or on guard?	1	2	3	4	5
17.	Feeling *jumpy* or easily startled?	1	2	3	4	5

PCL-M for DSM-IV (11/1/94) Weathers, Litz, Huska, & Keane National Center for PTSD - Behavioral Science Division

Alcohol Use Disorders Identification Test (AUDIT)

The Alcohol Use Disorders Identification Test (AUDIT), developed in 1982 by the World Health Organization, is a simple way to screen and identify people at risk of alcohol problems.

1. How often do you have a drink containing alcohol?

(0) Never (Skip to Questions 9-10)
(1) Monthly or less
(2) 2 to 4 times a month
(3) 2 to 3 times a week
(4) 4 or more times a week

2. How many drinks containing alcohol do you have on a typical day when you are drinking?

(0) 1 or 2
(1) 3 or 4
(2) 5 or 6
(3) 7, 8, or 9
(4) 10 or more

3. How often do you have six or more drinks on one occasion?

(0) Never
(1) Less than monthly
(2) Monthly
(3) Weekly
(4) Daily or almost daily

4. How often during the last year have you found that you were not able to stop drinking once you had started?

(0) Never
(1) Less than monthly
(2) Monthly
(3) Weekly
(4) Daily or almost daily

5. How often during the last year have you failed to do what was normally expected from you because of drinking?

(0) Never
(1) Less than monthly
(2) Monthly
(3) Weekly
(4) Daily or almost daily

REFERENCES

1. Arthur CA, Fitzwater J, Hardy L, Beattie SJ, Bell J. 2015. Development and validation of a military training mental toughness inventory. *Military Psychology* 27:232-41
2. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. 2001. The alcohol use disorders identification test. In *Guidelines for Use in Primary Care*: World Health Organization
3. Banks LM. 2006. The history of special operations psychological selection. In *Psychology in the Service of National Security*, ed. DA Mangelsdorff, pp. 83-96. Washington DC: American Psychological Association
4. Bell NS, Amoroso PJ, Yore MM, Smith GS, Jones BH. 2000. Self-reported risk-taking behaviors and hospitalization for motor vehicle injury among active duty army personnel. *American journal of preventive medicine* 18:85-95
5. Bing MN, Eisenberg KL. Development and validation of the SubMarine Attrition Risk Test (SMART). *Proc. The American Psychological Association, 112th Annual Conference, Honolulu, HI, 2004*:
6. Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. 1996. Psychometric properties of the PTSD Checklist (PCL). *Behaviour research and therapy* 34:669-73
7. Bliese PD, Wright KM, Adler AB, Cabrera O, Castro CA, Hoge CW. 2008. Validating the primary care posttraumatic stress disorder screen and the posttraumatic stress disorder checklist with soldiers returning from combat. *Journal of consulting and clinical psychology* 76:272

8. Boake C. 2002. From the Binet–Simon to the Wechsler–Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology* 24:383-405

9. Bonanno GA, Mancini AD, Horton JL, Powell TM, LeardMann CA, et al. 2012. Trajectories of trauma symptoms and resilience in deployed US military service members: prospective cohort study. *The British Journal of Psychiatry* 200:317-23

10. Britt TW, Dickinson JM, Moore D, Castro CA, Adler AB. 2007. Correlates and consequences of morale versus depression under stressful conditions. *Journal of Occupational Health Psychology* 12:34-47

11. Britt TW, Moore MA, Adler AB, Bartone PT. 1995. Responsibility, stress, and health: Testing the triangle model of responsibility, Walter Reed Army Institute of Research, Washinton DC

12. Campbell JP, Knapp DJ, eds. 2001. *Exploring the limits in personnel selection and classification*. Mahwah, NJ: Lawrence Erlbaum Associates.

13. Cattell HP, Mead AD. 2008. The sixteen personality factor questionnaire (16PF). *The SAGE handbook of personality theory and assessment* 2:135-78

14. Cattell RB, Cattell HP. 1995. Personality structure and the new fifth edition of the 16PF. *Educational and Psychological Measurement* 55:926-37

15. Center for Disease Control and Prevention. 2010. Web-based Injury Statistics Query and Reporting System (WISQARS). National Center for Injury Prevention and Control. Available at: www.cdc.gov/injury/wisqars/index.html

16. Chalupa SB. 2014. *Personality Predictors of First-Year Attrition From A Military Training Program*. Doctoral Dissertation. Texas A&M University

17. Collins DL. 2002. Psychological issues relevant to astronaut selection for long-duration space flight: a review of the literature. *Human performance in extreme environments: the journal of the Society for Human Performance in Extreme Environments* 7:43-67

18. Cox BD, Schmidt LL, Slack KJ, Foster TC. 2013. Assessment and selection of military aviators and astronauts. In *Aeromedical Psychology*, ed. CH Kennedy, GG Kay:20. Burlington, VT: Ashgate Publishing Company. Number of 20 pp.

19. Department of the Army. 1987. Alternate flight aptitude selection test (AFAST) information pamphlet. Washington, DC: Department of the Army

20. Distifeno C. 2008. Effects of Moral Conduct Waivers on First-Term Attrition of US Army Soldiers. *Rep. NPS-HR-08-008*, Naval Postgraduate School, Monterey, CA

21. Etcho LL. 1996. The Effect of Moral Waivers on First-Term, Unsuitability Attrition in the Marine Corps, Naval Postgraduate School, Monterey, CA

22. Feeley SP. 1998. *Special forces assessment and selection*. Naval Postgraduate School Monterey, CA

23. Flyer ES. 1959. Factors relating to discharge for unsuitability among 1956 Airman accessions to the Air Force. Lackland Air Force Base, TX: Personnel Laboratory Wright Air Development Center

24. Gebicke ME, Farrell BS. 1999. Military Personnel: Actions Needed to Better Define Pilot Requirements and Promote Retention. Washington, DC: United States General Accounting Office

25. Gilmore GJ. 2001. *Recruit attrition rates fall across the services*. http://www.defense.gov/news/newsarticle.aspx?id=44782

26. Gucciardi DF, Hanton S, Gordon S, Mallett CJ, Temby P. 2015. The concept of mental toughness: tests of dimensionality, nomological network, and traitness. *Journal of personality* 83:26-44

27. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. 2006. *Multivariate data analysis*. Upper Saddle River, NJ: Pearson Prentice Hall

28. Handler L. 2001. Assessment of men: Personality assessment goes to war by the Office of Strategic Services assessment staff. *Journal of personality assessment* 76:556-78

29. Harkins G. 2013. *Marines' embassy guard boost will be tough, experts say*. http://www.usatoday.com/story/news/nation/2012/12/24/marines-embassy-security/1789739/

30. Hoge CW, Auchterlonie JL, Milliken CS. 2006. Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *Jama* 295:1023-32

31. Hoge CW, Lesikar SE, Guevara R, Lange J, Brundage JF, et al. 2002. Mental disorders among US military personnel in the 1990s: association with high levels of health care utilization and early military attrition. *American Journal of Psychiatry* 159:1576-83

32. Hoge CW, Toboni HE, Messer SC, Bell N, Amoroso P, Orman DT. 2005. The occupational burden of mental disorders in the US military: psychiatric hospitalizations, involuntary separations, and disability. *American Journal of Psychiatry* 162:585-91

33. Hulse JA, Chenowith T, Lebedovych L, Dickinson P, Cavanaugh G, Garrett N. 2007. Predictors of student success in the US Army Graduate Program. *AANA Journal* 75:339-46

34. Jensen M, Rotter J. 1945. The validity of the Multiple Choice Rorschach Test in officer candidate selection. *Psychological Bulletin* 42:182-5

35. Jeppe AL. 2008. Determining the relationship between moral waivers and Marine Corps unsuitability attrition, Naval Postgraduate School, Monterey, CA

36. Johnsen BH, Bartone P, Sandvik AM, Gjeldnes R, Morken AM, et al. 2013. Psychological hardiness predicts success in a Norwegian armed forces border patrol selection course. *International Journal of Selection and Assessment* 21:368-75

37. Johnson VE. 2013. Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences* 110:19313-7

38. Kempthorne D, Byrd RC. 1997. Military Attrition: DOD could save millions by better screening enlisted personnel. Washington DC: United States General Accounting Office

39. Kennedy CH, Kay GG. 2013. *Aeromedical Psychology*. Burlington, VT: Ashgate Publishing, Ltd.

40. Knapik JJ, Jones BH, Hauret K, Darakjy S, Piskator E. 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

41. Kubisiak UC, Lentz E, Horgen KE, Bryant RH, Connell PW, et al. 2009. Review of interventions for reducing enlisted attrition in the US military: An update, U.S. Army Research Institute for Behavioral and Social Sciences, Arlington, VA

42. Larson GE, Highfill-McRoy RM, Booth-Kewley S. 2008. Psychiatric diagnoses in historic and contemporary military cohorts: combat deployment and the healthy warrior effect. *American Journal of Epidemiology* 167:1269-76

43. Luxton DD, Skopp NA, Maguen S. 2010. Gender differences in depression and PTSD symptoms following combat exposure. *Depression and anxiety* 27:1027-33

44. Means B. 1983. Moral standards for military enlistment: Screening procedures and impact, Human Resources Research Organization, Alexandria, VA

45. Merrill LL, Stander VA, Thomsen CJ, Crouch JL, Milner JS. 2004. Childhood exposure to family violence and attrition in the Navy. *Military Medicine* 169

46. Morgan III CA, Southwick SM, Hazliett G, Dial-Ward M. 2008. Baseline dissociation and prospective success in special forces assessment and selection. *Psychiatry (Edgmont)* 5:53-8

47. Muthen LK, Muthen BO. 1998-2012. *Mplus user's guide*. Los Angeles, CA: Muthen & Muthen

48. National Center for PTSD. 2014. Using the PTSD Checklist for DSM-IV (PCL). Washington, DC: U.S. Department of Veterans Affairs

49. Navy Medicine Operational Training Center. 2015. *ASTB-E Overview*. http://www.med.navy.mil/sites/nmotc/nami/Pages/ASTBOverview.aspx

50. Picano JJ, Williams TJ, Roland RR. 2012. Assessment and selection of high-risk operational personnel. In *Military Psychology: Clinical and Operational Applications*, ed. CH Kennedy, EA Zillmer. New York, NY: Guilford Press. Number of.

51. Putka DJ, Noble C, Becker D, Ramsberger P. 2004. Evaluating moral character waiver policy against service member attrition and in-service deviance through the first 18 months of service. Alexandria, VA: Human Resources Research Organization

52. Reis JP, Trone DW, Macera CA, Rauh MJ. 2007. Factors associated with discharge during marine corps basic training. *Military medicine* 172:936-41

53. Resnick RJ. 1997. A brief history of practice - expanded. *American Psychologist* 52:463-8

54. Rezvan PH, Lee KJ, Simpson JA. 2015. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC medical research methodology* 15:30

55. Seal KH, Metzler TJ, Gima KS, Bertenthal D, Maguen S, Marmar CR. 2009. Trends and risk factors for mental health diagnoses among Iraq and Afghanistan veterans using Department of Veterans Affairs health care, 2002–2008. *American Journal of Public Health* 99:1651-8

56. Seifert AE, Polusny MA, Murdoch M. 2011. The association between childhood physical and sexual abuse and functioning and psychiatric symptoms in a sample of US Army soldiers. *Military medicine* 176:176-81

57. Stowers MR, Thompson JL. 2011. Personality as a Predictor of Military Suitability. *Journal of Organizational Psychology* 11:24-34

58. Suedfeld P, Steel GD. 2000. The environmental psychology of capsule habitats. *Annual Review of Psychology* 51:227-53

59. Tabachnick B, Fidell L. 2007. *Using Multivariate Statistics*. Boston, MA: Pearson

60. Tarescavage AM, Brewster J, Corey DM, Ben-Porath YS. 2014. Use of Prehire Minnesota Multiphasic Personality Inventory-2–Restructured Form (MMPI-2-RF) Police Candidate Scores to Predict Supervisor Ratings of Posthire Performance. *Assessment*:1073191114548445

61. Tarescavage AM, Fischler GL, Cappo BM, Hill DO, Corey DM, Ben-Porath YS. 2015. Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) predictors of police officer problem behavior and collateral self-report test scores. *Psychological assessment* 27:125

62. Taylor MK, Markham AE, Reis JP, Padilla GA, Potterat EG, et al. 2007. Physical fitness influences stress reactions to extreme military training. *Military Medicine* 173

63. Tein J-Y, Coxe S, Cham H. 2013. Statistical power to detect the correct number of classes in latent profile analysis. *Structural equation modeling: a multidisciplinary journal* 20:640-57

64. Tremblay MA, Blanchard CM, Taylor S, Pelletier LG, Villeneuve M. 2009. Work Extrinsic and Intrinsic Motivation Scale: Its value for organizational psychology research. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement* 41:213

65. U.S. Army. 2014. *Airborne and Ranger Training Brigade: Ranger school statistics*. http://www.benning.army.mil/infantry/rtb/

66. U.S. Navy. 2014. *SEAL + SWCC FAQ*. http://www.sealswcc.com/navy-seals-frequently-asked-questions-faq.html#.VKnYCIrF8mU

67. USAF. 2015. *Pilot Candidate Selection Method FAQ*. http://access.afpc.af.mil/pcsmdmz/faq.html

68. Welsh Jr JR, Kucinkas SK, Curran LT. 1990. Armed services vocational battery (ASVAB): Integrative review of validity studies, Air Force Human Resources Laboratory, Brooks Air Force Base, TX

69. Whanger JC, Bing MN, America A, Lamb JC. 2009. Psychometric assessment in the US submarine force. In *Human Performance Enhancement in High-Risk Environments: Insights, Developments, and Future Directions from Military Research*, ed. PE O'Connor, JV Cohn. Santa Barbara, CA: ABC-CLIO, LLC. Number of.

70. White IR, Royston P, Wood AM. 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in medicine* 30:377-99

71. 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

72. Williams C. 2015. *Peer evaluations can mean success, failure in Ranger School*. http://www.military.com/daily-news/2015/05/04/peer-evaluations-can-mean-success-failure-in-ranger-school.html

73. Williams JO, Bell NS, Amoroso PJ. 2002. Drinking and other risk taking behaviors of enlisted male soldiers in the US Army. *Work: A Journal of Prevention, Assessment and Rehabilitation* 18:141-50

74. Wittson C, Hunt W, Older H. 1944. The use of the Multiple Choice Group Rorschach Test in military screening. *The Journal of Psychology* 17:91-4

75. Wright C, Grodin DM, Harig PT. 1990. Occupational outcome after military treatment for alcoholism. *Journal of Occupational and Environmental Medicine* 32:24-32

76. Wrzesniewski A, Schwartz B, Cong X, Kane M, Omar A, Kolditz T. 2014. Multiple types of motives don't multiply the motivation of West Point cadets. *Proceedings of the National Academy of Sciences* 111:10990-5