UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE ROLE OF SEX, MILITARY SCIENCE CLASS, AND ANTHROPOMETRICS ON PERFORMANCE OF THE ARMY COMBAT FITNESS TEST (ACFT) IN RESERVE OFFICERS' TRAINING CORPS (ROTC) CADETS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

MEGAN NICOLE SAX VAN DER WEYDEN

Norman, Oklahoma

2021

THE ROLE OF SEX, MILITARY SCIENCE CLASS, AND ANTHROPOMETRICS ON PERFORMANCE OF THE ARMY COMBAT FITNESS TEST (ACFT) IN RESERVE OFFICERS' TRAINING CORPS (ROTC) CADETS

A THESIS APPROVED FOR THE
DEPARTMENT OF HEALTH AND EXERCISE SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Jason Campbell, Chair

Dr. Christopher Black

Dr. Daniel Larson

Dr. Rebecca Larson

Acknowledgements

First and foremost, I would like to thank God for blessing me with the talents and opportunities, without which I would not be where I am today. I would also like to thank the United States Army and Cadet Command for allowing me to pursue advanced degrees within the exercise physiology field and for supporting my research.

I am extremely grateful to my committee chair, Dr. Jason Campbell, for accepting me into his lab, assisting me through this rigorous research process, and advising me on all sorts of subjects. Likewise, I would like to thank my thesis committee of Dr. Chris Black, Dr. Rebecca Larson, and Dr. Daniel Larson for the time, energy, and expertise that they have dedicated to this project.

Thank you to my family for supporting me in all of my endeavors, both big and small, and for raising me with an everlasting hunger for knowledge. Lastly, thank you Paige for always being a voice of reason and my calm during the storm. I hope to make you all proud, always.

Table of Contents

List of Tables

List of Figures

Abstract

The Army Combat Fitness Test (ACFT) was created to better assess all realms of physical fitness on an age and sex-neutral scale. Despite an increasing number of females joining the more physically rigorous military occupational specialties, to date, no published research has analyzed how female soldiers score on the sex-neutral ACFT compared to male soldiers. Similarly, no research has analyzed how military experience plays a role in ACFT performance nor how body anthropometrics may affect ACFT outcomes. Lastly, no research has been presented analyzing how effective a semester of military physical training is at eliciting performance gains on the ACFT. **PURPOSE:** The main purpose of this study was to identify if and where sex-based and experience-based differences exist and to identify possible relationships between anthropometric variables and ACFT outcomes in ROTC cadets. An additional aim of this study was to analyze if a semester of physical training is an adequate stimulus to elicit ACFT performance improvements and to determine significant predictors that put a cadet at more or less risk to fail or score in a particular qualification category. **METHODS:** Retrospective ACFT scores, military experience, and anthropometric data were used from 20 universities within Cadet Command's 5th BDE. Each university's ROTC program conducted the ACFT at the end of the semester per published protocol and reported the de-identified results. Analysis was conducted using SPSS (SPSS Inc., Chicago, IL, version 26.0) and statistical significance was set at $p \leq 0.05$. **RESULTS:** Male cadets performed significantly better than female cadets on all ACFT events, in total ACFT score, and in average qualification category. Female cadets' worst event was the leg tuck (LT) with 40% of females failing the event. MS-3 and MS-4 (upperclassmen) cadets tended to outperform their underclassmen counterparts and scored statistically better than MS-1 (freshmen) cadets on the standing power throw (SPT), sprint-drag-carry (SDC), two-mile run (2MR), total points scored, and average qualification level. Height, mass, and BMI were significantly correlated with ACFT outcomes. Moderate correlations (> 0.5) were observed between height and the three-repetition maximum dead lift (MDL), SPT, SDC, and total points. Moderate correlations were also observed between mass and MDL, SPT, and SDC. All other correlations were significant but weak. When separated by sex, anthropometric values were significantly, but less strongly correlated to ACFT outcomes. A semester of physical training was able to elicit performance gains with male cadets significantly improving their MDL, SDC, 2MR, total points, and qualification category outcomes and female cadets significantly improving their SDC, LT, SPT, total points, and qualification category outcomes. Lastly, multinomial logistic regression determined sex, height, body mass, and military experience to be significant predictors in ACFT qualification categories for individual events and total outcomes. **CONCLUSION:** In summary, females performed significantly worse on the ACFT than males and cadets tend to increase their overall fitness each year they are enrolled in the ROTC program. Based on these, pre- and post-semester, and anthropometric results, additional resistance training to increase strength, power, anaerobic capacity, and muscle mass may be beneficial for female and MS-1 cadets.

Key Words: Tactical Fitness, Military Physical Testing, Female Soldiers, Strength and Conditioning

Chapter I – Introduction

Introduction

The nature of conflicts in the 21st century necessitates all United States Army soldiers, regardless of sex, must possess the basic soldier functions required to deploy, fight, and win the Nation's wars. Since the ban on women serving in combat roles was lifted in 2015, increasing amounts of females have joined the physically demanding combat arms branches. Additionally, females have now entered and passed prestigious schools like Army Ranger School and Army Special Forces Qualification. Females are now being held to the same physical standards as males. Because of this, a new sex-neutral physical fitness test was recently implemented.

Previously, the Army Physical Fitness Test (APFT), consisting of push-up, sit-up, and 2-mile run events, was the official fitness test of the United States Army from 1980 until 2020 (Knapik & East, 2014). It has long been argued the APFT was not a good indicator of soldiers' overall fitness levels, ability to succeed at job duties, and perform in combat, because it did not test strength, power, anaerobic capacity, but rather muscle endurance and aerobic capacity only (Batchelor, 2008). Recently, the Army announced the newly established Army Combat Fitness Test (ACFT), will replace the APFT in 2020 as the official test of record (Headquarters Center for Initial Military Training, 2019). The ACFT incorporates events that test muscular strength, power, agility, and anaerobic endurance as well as the traditional muscular endurance and aerobic endurance (Bedard et al., 2019). The ACFT is scored on a sex and age neutral scale and minimum requirements are determined by a soldier's military occupational specialty (MOS).

Despite females having successfully served in combat roles and passing intense military schools, there are still inherent physiological differences between males and females. On average, females possess lower cardiorespiratory fitness, less absolute muscular strength and power, and lower anaerobic capacity compared to males (Courtright et al., 2013). Additionally, the average female is shorter, weighs less, and has more relative fat mass and less fat free mass than males (Miller et al., 1993; Rogol et al., 2002; Wells & Plowman, 1983). Mechanisms behind these physiological disparities stem from differences in growth during puberty, inherent differences in heart size and function, muscle morphology and the amount of muscle mass to produce force (Cain et al., 2009; Jeon et al., 2019; Miller et al., 1993). These innate differences between males and females may pose difficulties for performance on the ACFT, where multiple realms of fitness are tested. While some characteristics, such as height and heart size cannot be corrected, others such as muscle mass, fat mass, and cardiorespiratory fitness can be altered through specific training. If the differences between males and females are identified, future preparation can be modeled to attempt to decrease the gap.

The Army Body Composition Program has established guidelines for minimum and maximum allowable body mass index (BMI) depending on age and sex of soldiers (Headquarters Department of the Army, 2019). Failure to pass BMI standards results in the assessment of body fat percent through circumference measurements where further failure may result in disciplinary action. The applicability of BMI as an adequate measure of body composition in active populations has been debated in the literature as it can incorrectly label individuals with high muscle mass and low body fat percent as overweight or obese (Grier, Canham-Chervak, Anderson, et al., 2015). While higher BMIs have been negatively correlated with performance on the APFT and tests of cardiovascular endurance, they have also been positively correlated with performance on tests of power and strength (J. Pierce et al., 2018). The ACFT incorporates tests of cardiovascular endurance, muscular strength and endurance, power, and anaerobic capacity and therefore, optimal BMI for successful performance may differ across the varying events.

Active Duty, National Guard, and Reserve soldiers, as well as Army ROTC (Reserve Officers' Training Corps) programs, are all adopting the ACFT as their official physical fitness test. ROTC, the largest commissioning source for the Army's officers, produces the future leaders of the Army. Thus, cadets are expected to maintain physical fitness to the same standards as regular soldiers. No formal basic training is required for a cadet to initially sign up for ROTC. Therefore, military education and tri-weekly physical training sessions are used to enhance physical fitness and military occupational awareness. Thus, upperclassmen may differ from underclassmen since they have had significantly more experience and exposure to both. To date, no published studies have analyzed the physical differences between cadets in different military science classes nor on the basis of sex related to ACFT performance. Therefore, understanding the physical fitness capabilities across sexes and military science classes is integral to recognizing where the greatest changes in fitness may occur during a cadet's career and how to better determine the training strategies that may facilitate the most equitable performance on the ACFT.

Purpose

The ACFT was created to mirror the physical requirements needed to perform job duties, Warrior Tasks/Battle Drills, and Common Soldier Tasks but will require widespread adjustment across the entire force. The purpose of this study was to identify if and where sex-based and experience-based differences exist in ACFT scores within a cohort that regularly conducts military-based physical training. The purpose of this study was also to identify possible relationships between anthropometric variables and ACFT outcomes. Additionally, the purpose of this study was the analyze if a semester of physical training is an adequate stimuli to elicit ACFT performance improvements. Lastly, the purpose of this study was to determine significant predictors that put a cadet at more or less odds to fail or score in a particular qualification category.

Research Questions

1. Do male and female ROTC cadets achieve similar score totals on the ACFT?

   a. Do male and female ROTC cadets achieve similar scores on the individual events of the ACFT?

2. Do ROTC cadets of different military science classes achieve similar score totals on the ACFT?

   a. Do cadets of different military science classes score similarly on the individual events of the ACFT?

3. Are there correlations between BMI and ACFT outcomes?

   a. Which event(s) is BMI positively correlated with?

   b. Which event(s) is BMI negatively correlated with?

4. Do male and female ROTC cadets significantly improve their total ACFT scores over the course of a semester?

   a. Do male and female ROTC cadets significantly improve in individual ACFT events over the course of a semester?

5. Do sex, MS-class, and anthropometrics play a role in ROTC cadets' overall qualification category?

Hypotheses

1. There will be no significant difference between male and female ROTC cadets' total ACFT scores.

   a. There will be no significant difference between male and female ROTC cadets' individual ACFT event scores.
2. ROTC cadets of different military science classes will not have significantly different total ACFT scores.
   a. ROTC cadets of different military science classes will not have significantly different individual event scores on the ACFT.
3. There will be no correlations between BMI and ACFT outcomes.
   a. No events are positively correlated with BMI.
   b. No events are negatively correlated with BMI.
4. Male and female cadets will not significantly improve their total ACFT scores.
   a. Male and female cadets will not significantly improve their individual ACFT event scores.
5. Sex, MS-class, and anthropometrics will not play a significant role in cadets' overall qualification category.

Alternate Hypotheses

1. Males and female ROTC cadets will have significantly different score totals on the ACFT.
   a. There will be significantly different individual ACFT event scores between male and female ROTC cadets.
2. ROTC cadets in different military science classes will have significantly different total ACFT scores
   a. ROTC cadets of different military science classes will have significantly different individual event scores on the ACFT.

3. There will be significant correlations between BMI and ACFT outcomes.
    a. BMI will be positively correlated with the three-repetition maximum deadlift and the standing power throw.
    b. BMI will be negatively correlated with the two-mile run.
4. Male and female cadets will significantly improve their total ACFT score over the course of a semester.
    a. Male and female cadets will significantly improve their individual ACFT event scores over the course of a semester.
5. Sex, MS-class, and anthropometrics will play a significant role in cadets' overall qualification category.

Significance

Army Training and Doctrine Command (TRADOC) has identified physical tasks required of all soldiers such as running, changing direction, pushing, pulling and/or climbing, lifting and jumping and landing (Knapik et al., 2009; Lowman, 2010). These are key elements in the successful completion of Warrior Tasks and Battle Drills and Common Soldier Tasks; basic tasks every soldier must be proficient at given the nature of today's conflicts. A survey of Army Majors, who are combat veterans, revealed the most important physical tasks in combat to be the following: move from one covered and concealed position to another, lift a weight from the ground, drag a casualty to safety, conduct a "Fireman's" carry, engage in continuous movement under combat load, and climb a wall (Batchelor, 2008). These, along with tasks identified by TRADOC (Training and Doctrine Command, U.S. Army), are critical in ensuring not only the success of the Army's mission, but also the well-being and safety of each and every soldier.

ROTC cadets, during field training exercises, are often required to complete the Warrior Task and Battle Drills. Thus, physical tasks required of soldiers apply to cadets as well.

The ACFT was created to better assess the physical fitness required for soldiers to complete these tasks. Because females now serve in combat roles (key for the assumption of leadership positions within the military), the importance of females meeting the same ACFT standards as males is integral to mission safety, mission success, and for women to have access to all Army military occupational specialties (MOS). No studies to date have been published regarding the ACFT and differences between the sexes. Therefore, identifying where these differences may exist has provided crucial insight on how to adjust training to decrease, where possible, the disparity between sexes. Since ROTC cadets, upon graduation, will serve as officers in the National Guard, Reserve, or Active Duty, preparing them to meet the same standards as the soldiers they will someday lead is of utmost importance. Because the ACFT was only recently established, there are currently no studies that have been published incorporating the test. Therefore, this study helps fill the gap in literature and provides baseline knowledge that can be included in future studies.

Delimitations

The delimitations of this study are as follows:

1. Cadets currently participating in an Army ROTC Program.
2. Cadets that have taken at least one prior ACFT before the test included in the data set.
3. Retrospective ACFT scores obtained at the end of a semester of physical fitness training.
4. Cadets will be 18-35 years old.
5. Cadets have no musculoskeletal injuries and cardiac or metabolic abnormalities at the time of testing.

6. Participating ROTC programs within Cadet Command's 5^{th} Brigade.

Limitations

The limitations of this study are as follows:

1. Researchers were not present when the ACFT was conducted and could not ensure all protocols were followed correctly.
2. Cadets may train and/or perform to meet the standard of their desired MOS or program mandated minimum and not to achieve the best score possible.
3. Researchers could not control for training that cadets conduct outside of structured, mandated ROTC physical training sessions.
4. Researchers could not account for differences in nutritional intake and sleep habits of cadets.
5. Cadets may join ROTC with differing levels of baseline fitness and/or training experience.

Assumptions

The assumptions of this study are as follows:

1. All Army Combat Fitness Tests were conducted according to published Army protocol and under similar conditions (Headquarters Center for Initial Military Training, 2019).
2. Cadets gave their best effort during the ACFT and scores are indicative of their abilities.
3. Cadets were free of injury or illness at the time of testing and observed normal patterns for nutritional intake and sleep.
4. Cadets join ROTC with similar baseline fitness and training experience.

Operational Definitions

1. **Army Physical Fitness Test (APFT):** Test compromised of two minutes of push-ups, two minutes of sit-ups, and a two-mile run. (Knapik & East, 2014)
2. **Army Combat Fitness Test 2.0 (ACFT):** Test compromised of six events; three repetition maximum deadlift (MDL), hand release pushup (HRP), standing power throw (SPT), sprint-drag-carry (SDC), leg tuck (LT), and a two-mile run (2MR) (Headquarters Center for Initial Military Training, 2019).
3. **Qualification Categories:** Four different qualification categories a soldier/cadet can achieve on individual events and the ACFT 2.0 as a whole; “FAIL”, “MODERATE”, “SIGNIFICANT”, and “HEAVY”. Minimum qualifications that must be met by a soldier are determined by their MOS. See Appendix A for an ACFT score chart.
4. **Reserve Officer Training Corps (ROTC):** ROTC is the primary commissioning source for United States Army Officers. ROTC programs require cadets to complete military-centric coursework in addition to their respective major. Cadets attend military science class at least once per week, complete structured physical fitness training at least three times per week, and attend one leadership laboratory session per week. Leadership laboratory teaches cadets the skills and tactics necessary to conduct field operations such as attacks, raids, ambushes, reconnaissance, and defense.
5. **Military Science Class (MS-n):** Cadets attend different military science classes depending on how long they have been enrolled in ROTC. Freshmen are considered MS-1, sophomores MS-2, juniors MS-3, and seniors MS-4.

6. **Military Occupational Specialty (MOS):** The specific job a soldier holds usually referred to as a letter and a number. Example: 11B, 13A, 68W, etc. Some MOS have higher physical fitness requirements than others on the ACFT 2.0 scale.
7. **Body Mass Index (BMI):** Body mass index is an anthropometric measured calculated with the equation; kg/m^2. Body mass is measured in kilograms and height in meters.

Chapter II – Literature Review

Introduction

The wide variety of physical requirements soldiers face on the battlefield, in training, and in garrison has caused debate about how the United States Army needs to test their soldiers. For four decades, the APFT, was the official fitness test of the Army, despite testing only local muscular and cardiovascular endurance on a sex and age based scale. As a better indicatory of overall fitness, the ACFT was created and implemented in 2020. The ACFT tests muscular strength, muscular endurance, power, anaerobic capacity, and aerobic capacity through six events and is scored on a sex and age neutral scale. The sex-neutral scoring system for the ACFT poses potential problems as there are inherent physiological differences between males and females. However, because of the recent repeal of the ban on females serving in combat roles, the importance of females meeting the same standards on physical assessments as males with the same jobs is integral to successful completion of the mission. Therefore, identifying where physiological differences may exist, as seen through ACFT scores, and decreasing the gap is of utmost importance to mission success. Because of new muscular strength, power, and anaerobic components of the ACFT, body composition variables, such as amount of muscle mass, may play an important role in the passing or failing of this assessment. Consequently, body mass index (BMI) and its applications to tactical populations' and their physical fitness is a hotly debated topic that will be assessed in this study.

This chapter aims to review the current literature surrounding these topics. A search was conducted of the Medline database for the keywords; 1) Army Combat Fitness Test OR ACFT, 3) Army Physical Fitness Test OR APFT, 4) Occupational Physical Assessment Test OR OPAT, 5) Military Fitness, 6) Reserve Officer Training Corps OR ROTC, 7) Body Mass Index OR BMI,

8) Sex differences, and combinations of those. Inclusion criteria for research articles was to be published between 2000-present and to utilize non-sedentary test subjects. Articles in the citations of those found through the systematic search were also included. Exceptions to these criteria were made for research that first established and reported sex differences in normal populations before 2000.

This review will first discuss the established physiological differences between sexes in anthropometrics and body composition, cardiovascular fitness, and muscular characteristics, and the mechanisms behind them. Then, it will discuss previously researched sex differences in tactical populations quantified through military fitness tests and laboratory based measurements. Finally, this article will review BMI, its correlations with tactical fitness, and the adequacy of BMI as a body composition measure.

Physiological Differences

Anthropometric and Body Composition Differences

Height

The stages of puberty and physiological differences between males and females during the first and second decade of life were first reported by J. Tanner and colleague, W. Marshall, in the 1960's and 1970's. Tanner and Marshall were able to analyze various aspects of anthropometric measures and how they change during pre-puberty. The average height difference between grown men and women is 13cm (Rogol et al., 2002). According to Tanner and Marshall, this is due to males having an additional two-year period of pre-pubertal growth compared to girls. Additionally, during that time, males reach a greater peak height velocity, or grow at a greater rate per year, than females do (Marshall & Tanner, 1969, 1970). The increased time and velocity of growth leads males to, on average, surpass females in height.

Fat Free Mass

Fat free mass (FFM) is any portion of the body that is not fat. This includes organs, bones, muscle, and connective tissue. This review is most focused on the skeletal muscle mass aspect of FFM and will not be reviewing the other contributors to FFM. Men in their 20's have approximately 50% more upper body and 30% more lower body muscle mass than females(Miller et al., 1993). According to Miller, et al., this corresponds to 14% more and 35% larger muscle fibers (Miller et al., 1993). The primary reason for this disparity in muscle mass between the sexes is the androgen stimulated skeletal muscle hypertrophy in males (Sharp, 1993). Overall, males have greater amounts of testosterone than females. During puberty, with the initial increase of testosterone in developing males, males develop larger and stronger muscles than females (Tanner, 1965). It is well known that males begin puberty later than females. Because of that, most of adult FFM is gained by the ages of 15-16 in females but not until the age of 19-20 in males (Malina & Bouchard, 1991). This is especially important to note in the context of this study because the average age of enlisted personal is 20 years old and that is when men are leaner and have more muscle match than women of the same age (Nindl et al., 2016).

Fat Mass and Fat Percent

The average female has less FFM, and more fat mass (FM) compared to the average male. Healthy females, typically, have 20-25% body fat compared to only 13-16% in males (Wells & Plowman, 1983). In males and females entering Army basic training, there have been significant differences in body fat reported. This untrained population averaged 16.7 % ± 6.3 in males and 28.8 % ± 6.5 in females (Knapik et al., 2001). With the rise of overweight and obese persons in the 21st century, the average American male and female now has 28% and 40% body fat, respectively, and most would not qualify for military service at that body composition (St-Onge, 2010). Nevertheless, females have greater percent of fat mass compared to males.

The mechanisms behind this, again, center around differences in puberty. During puberty, percent body fat in females increases and fat mass is accumulated at ~ 1.14 kg/year. In contrast, males decrease their percent body fat and their fat mass by 1.15 kg/year (Rogol et al., 2002). One possible explanation for this is plasma leptin. Leptin is thought to help regulate body fat content by sending signals to the brain to alter food intake and energy expenditure. In females, plasma leptin levels will increase after Tanner stage 2 of puberty. However, in males, plasma leptin levels decrease. Females may become more resistive to the effects of leptin and thus accumulate more levels of fat (Saad et al., 1997; Wong et al., 2004). A simpler explanation to the disparity in body fat levels may be the higher essential fat required by females, 12%, compared to males, 3% (Nindl et al., 2016; Sharp, 1993). The additional 9% of essential body fat in females is located mostly in the breasts and surrounding the reproductive system (Wells, 1991).

Cardiovascular Differences

VO_{2max}

The maximum amount of oxygen than a person can be consumed and used during exhaustive exercise is commonly referred to as the VO_{2max}. Testing a person's VO_{2max} is the gold-standard method of assessing cardiorespiratory fitness. On average, females' VO_{2max} is anywhere between 15-30% lower than males' (Drinkwater, 1973; Yanovich et al., 2008). Before physical training, significant differences have been reported between absolute and relative VO_{2max} in males and females (Knapik et al., 2001). Likewise, before physical training in another study, males have 22% higher aerobic fitness than females (Yanovich et al., 2008). Even after completing 10 weeks of physical training, the female to male VO_{2max} ratio was still 0.63 in absolute terms and 0.75 in relative terms (Sharp, 1993).

The expression of VO_{2max} in relative terms is meant to correct the absolute value to a person's body weight and is reported as ml/kg/min. Relative VO_{2max} is greatly influenced by excess fat. The more fat a person has, the more metabolically inactive tissue, or dead weight, a person has. Females, on average, have greater FM and thus, more dead weight that does not contribute to aerobic metabolism. Though ratios between males and females may increase after physical training, thereby decreasing the difference between the two, there is still a significant difference in VO_{2max} in absolute and relative terms. Physical training can only improve a females VO_{2max} to a certain extent before other physiological variables hinder obtaining the same VO_{2max} as an average male.

Heart Characteristics

A possible mechanism for the disparity in VO_{2max} between males and females is that females have smaller hearts. Cardiac output is the volume of blood being pumped by the heart per unit of time and is calculated as Heart Rate x Stroke Volume. The stroke volume component of cardiac output is influenced by left ventricular volume, blood volume, and heart muscle mass. Regardless of body size, females have thinner left ventricular walls, less heart mass, and smaller cavity sizes (Cain et al., 2009; Epstein et al., 2013). Therefore, females have lower cardiac output compared to males and thus, less blood being sent to the body each contraction during aerobic exercise. It is calculated that this difference between cardiac output in males and females is as high as 30% (Wells, 1991). Therefore, to maintain intense aerobic exercise, females rely on increased heart rate while males rely on preload and the Frank Starling mechanism (Fleg et al., 1995). Females, however, have the same age-predicted maximum heart rate as males and there is only so high one's heart rate can increase.

Hemoglobin

Hemoglobin is an iron-containing oxygen transport protein in red blood cells that is responsible for carrying oxygen from the lungs to the body. Females, on average, have 6% fewer red blood cells and 10-16% less hemoglobin than males (Charkoudian & Joyner, 2004; Wells, 1991; Yanovich et al., 2008). Less hemoglobin means decreased oxygen carrying capacity. If less oxygen is getting to the skeletal muscles during maximal aerobic exercise, females will fatigue quicker than males and have a lower $VO_{2max.}$

Musculoskeletal Differences

Muscular Strength

As discussed in previous sections, males, on average, have more fat free mass than females. More muscle mass equates to greater strength due to increased amount of muscle fibers that can create force. Therefore, males are generally stronger than females and the differences in strength are largely due to differences in skeletal muscle mass. Reports on absolute differences in muscle strength vary but come to a general consensus of approximately 40% difference in upper and lower body strength between sexes (Epstein et al., 2013; Knapik et al., 2001; Yanovich et al., 2008). Recruits who were tested before basic combat training yielded significantly different results on tests of muscular strength. Males and females were significant different in upper and lower body static strength, dynamic lifts, and vertical jump (Knapik et al., 2001). It is thought that the difference between males and females for lower body strength is less than that for upper body strength. This can be seen in that female to male lower body strength ratios are higher than that of upper body strength ratios (Sharp, 1993). Less disparity in lower body strength may be due to similar habitual activities between the sexes such as walking, running, and climbing stairs and/or greater quantity of muscle mass in the lower body regions than upper body regions in females.

Another possible mechanism behind increased strength in males is muscle fiber type distribution. Fast twitch, or Type II muscle fibers, are characterized by the ability to generate larger amounts of force than slow twitch, or Type II, fibers. Men have greater cross-sectional area of type II fibers and a higher ratio of type II fibers to type I fibers than in women (Jeon et al., 2019; Miller et al., 1993). Having more type II fibers than type I and having greater cross-sectional area of those type II fibers is advantageous for producing force. However, when measures of strength are corrected for differences in muscle mass, most of the sex difference disappears (Miller et al., 1993; Sharp, 1993). That is, the ability of muscle or produce force between the sexes does not differ; just the muscle mass does. Therefore, differences in fiber type CSA and distribution may not play a key role in strength disparities between the sexes.

Muscular Endurance

Muscular endurance is the ability of a muscle to sustain repeated contractions against a submaximal resistance for an extended period of time (Epstein et al., 2013). It has been established that females exhibit muscular endurance equal to or greater than males when exercising at a percent of maximal strength between 50-70% of one-repetition maximum (Maughan et al., 1986). Muscular endurance is equal between sexes when it is a percent of one's own maximal strength. However, when lifting an absolute load, females become fatigue faster than males because they are working at a greater percent of their maximal strength (Sharp, 1993). The advantage females have in muscular endurance declines as the contraction intensity increases.

A possible mechanism behind why females exhibit such greater muscular endurance compared to strength when compared to males is their proportion of muscle fibers. There is speculation that the proportion of type I to type II muscle fibers is higher in females than in males. These type I muscle fibers are fatigue resistant and contribute greatly to muscular endurance. Another possible mechanism is substrate utilization differences between the sexes. Males have greater glycolytic capacity and rely on those pathways more than females. In contrast, females rely more on fat sources, obtained through beta oxidation, for energy during submaximal endurance exercise. Fat is an energy rich substrate that has more stores in the body than glycogen (Hicks et al., 2001).

Anaerobic Capacity

Anaerobic exercise is characterized by a lack of oxygen and thus occurs for short duration, high intensity activities. Anaerobic capacity can be determined through tests such as the Wingate test; an all-out 30 second sprint on a cycle ergometer. Before taking part in basic training, males had 28.6% higher anaerobic fitness, as determined by Wingate test, than females (Yanovich et al., 2008). Likewise, in a similar study utilizing the Wingate test, males had absolute anaerobic power 35% and 40% higher than females for peak power and mean power, respectively. When expressed relative to FFM, the differences between males and females dropped to 10% and 17% for peak and mean power (Murphy et al., 1986). Similarly, a study utilizing the Wingate test in male and female college students found a 48% difference in peak power between the sexes that dropped 15.2% when expressed relative to total body weight. Both were still significant differences. However, when expressed relative to FFM, the percent difference between the sexes for peak power dropped to an insignificant 1.6% (Maud & Shultz, 1986). Therefore, absolute anaerobic power differences between the sexes but the disparity diminished when it is expressed relative to FFM. Unfortunately, in the context of this study, adjusting for FFM is not possible in a military operational environment when two soldiers of differing sexes and FFM have to complete the same task to the same standards.

The mechanisms behind the anaerobic differences between males and females may stem from fiber type distribution and/or contribution of anaerobic energy systems. Females possess less muscle mass than males and thus, have fewer available muscle fibers to contribute to the power production necessary for a test such as the Wingate test. Similarly, if the theory that females have a greater ratio of type I to type II muscle fibers hold true, then females have less of the fast twitch fiber types that are capable of producing the high intensity, short duration force associated with anaerobic exercise. As seen in recreational weight lifters, at higher work intensities, there is a greater anaerobic contribution in men than in women (Morgan et al., 2003).

Military Applications

Physical Fitness Test Sex Differences

The United States Army has cycled through various physical fitness tests throughout the past century with one of the most prominent being the Army Physical Fitness Test (APFT). The APFT remained the physical fitness test of the U.S. Army from 1980 until 2020 and had the Occupational Physical Assessment Test (OPAT) join in 2017 as a requirement for new recruits' job selection (Knapik & East, 2014). The APFT has sex and age graded scoring system while the OPAT is sex and age neutral. Announced in 2018 and officially implemented in 2020, the Army has officially transitioned to a new fitness test named the Army Combat Fitness test, or ACFT. The ACFT has a sex and age neutral scoring system and centers its minimum standards around a soldier's job requirements. Data on sex differences in the APFT and OPAT have been published but no studies regarding the ACFT, in any capacity, have been published to date.

APFT

The scoring system for the APFT assumes there will be sex and/or age-based differences for the three events: push-ups, sit-ups, and two-mile run. When administered at the beginning of basic combat training (BCT), males and females showed significantly difference absolute values for push-ups, sit-ups, and two-mile run (Knapik et al., 2001). Likewise, trainees measured before BCT showed a 52% difference in pushups, 10% in sit-ups, and 19% in the two-mile between males and females. After BCT, fitness was improved in females but there still remained a 33%, 5% and 18% difference, respectively (J. Pierce et al., 2018). Even in Army ROTC cadets, push-ups and two-mile run time were significantly difference between males and females, but sit-ups were not (Draicchio et al., 2020). In a meta-analysis analyzing the sex difference in physical abilities within persons that had physically strenuous occupations, the authors determined that core-strength is not significantly different between the sexes (Courtright et al., 2013). Therefore, the findings that ROTC cadets did not differ in sit-ups and soldiers after BCT only had a 5% difference in sit-ups is not surprising.

OPAT

The OPAT consists of a deadlift, an interval aerobic run, a standing long jump, and a medicine ball power throw and is graded on a sex and age neutral scoring system. The reasoning behind this is because the OPAT is used to determine eligibility for specific military occupational specialties and males and females vying for the same job must meet the same standards. In ROTC cadets conducting the OPAT, there were significant differences between males and females for all events with females scoring lower (Draicchio et al., 2020). Likewise, in basic trainees, males and females performed significantly different on all events. Females performed 41% fewer shuttles on the interval aerobic run, 34% shorter power throw distance, and 25% shorter standing long jump. Additionally, a lower portion of females were able to lift the maximum deadlift weight, 220lbs, than males; 15% vs 88% (J. Pierce et al., 2018). Only ROTC Cadets and basic trainees take the OPAT and therefore no data is available on Active Duty, Reserve, or National Guard soldiers.

ACFT

Not data has been published to date regarding the ACFT. The ACFT shares only one event with any previous military fitness test: the two-mile run in the APFT. While the ACFT does also incorporate a deadlift, like the OPAT, the event is a three-repetition maximum deadlift with the highest maximum score at 350lbs (Headquarters Center for Initial Military Training, 2019). The OPAT maximum is 220lbs and only requires one-repetition.

Lab-Based Fitness Test Sex Differences

Laboratory based fitness tests, or those outside of typical military physical fitness tests, also conclude that there are significant differences in physical abilities between males and females. After testing 406 101st Airborne soldiers in a variety of musculoskeletal, anaerobic, and aerobic tests, researchers determined that males and females were significantly different on many variables. Males displayed significantly greater strength, when assessed on shoulder and knee strength, and anaerobic power and capacity, according to the Wingate test, compared to females (Allison et al., 2015). Additionally, males had significantly greater aerobic capacity (47.5 ±7.6 ml/kg/min vs 40.3±5.4 ml/kg/min) as determined through a treadmill maximal graded exercise test (Allison et al., 2015). Similarly, an experiment comparing a shuttle run aerobic test to the typical two-mile run seen in military fitness tests determined that females scored significantly different, and worse, than males on both the shuttles and the two-mile run (Canino et al., 2018).

A meta-analysis of studies utilizing a population of people with physically demanding occupations, such as military, police officers, emergency medical technicians, firefighters, and the like, assessed sex based musculoskeletal differences as reported in the literature. In a meta-analysis, Cohen's d is used to report an effect size for the difference between two means. Muscle tension, defined by authors as exerting force against objects in pushing, pulling, lifting, etc., had the largest difference between males and females across all the studies analyzed and had a d value of d =2.13 (Courtright et al., 2013). This is a very large effect because 0.2 is considered a small effect and 0.8 a large one. Likewise, the meta-analysis determined that differences between the sexes for muscular endurance and power were significant and with d-values of d =1.47 and 1.11, respectively. When analyzing muscular strength for total body and body regions, there were d values of 2.22, 1.88, 1.60, 0.25 for total body, upper body, lower body, and core (Courtright et al., 2013). This is congruent to previous reports that upper body differences between the sexes are greater than that for lower body (Sharp, 1993). Lastly, it was determined there were significant differences between males and females for cardiovascular endurance with a reported effect size of d = 1.81. Therefore, even in an active population, the differences in fitness measures of muscular strength, muscular endurance, power, and cardiovascular fitness persist.

Body Mass Index (BMI)

Sex Differences

BMI, calculated as kg/m^2, has been reported as significantly different between soldiers. In the 101st Airborne, there were significant differences in both BMI and body fat percent between male and females soldiers (Allison et al., 2015). Females BMI and body fat perfect were greater than in males. In contrast, another investigation reported significant differences between sexes for height, body mass, and BMI. However, in this cohort of active duty enlisted soldiers and officers, males BMI was greater than in females (Canino et al., 2018).

Correlations in Tactical Populations

Correlations between BMI and fitness measures have been made in tactical populations based on military fitness tests, tactical obstacle course, and laboratory-based measurement outcomes. The APFT, a test of muscular endurance and cardiovascular endurance, has been found to be negatively correlated with BMI. Predictive factors for APFT failure in female soldiers was high BMI and researchers have concluded that BMI greater than 25 kg/m^2 has a negative influence on performance of cardiovascular and muscular endurance for both males and females (Anderson et al., 2017; Kazman et al., 2015). In female basic trainees, high BMI was correlated with increased APFT two-mile run time (J. Pierce et al., 2018). In contrast, however, in ROTC cadets, total APFT scores were not associated with body composition determined by BMI and two-mile run times were only correlated with body composition assessed with the BodPod and BIA, but not BMI (Steed et al., 2016). In addition to the APFT, high BMI was correlated to decreased OPAT interval aerobic run and standing long jump scores in male and female basic trainees (J. Pierce et al., 2018). Similarly, higher BMI has been associated with reduced speed and agility in male and female soldiers (J. R. Pierce et al., 2017).

In other tests of fitness, however, there are positive correlations between BMI and fitness variables. In basic trainees, high BMI was correlated with increased power throw distance as well as displaying a significant positive trend between BMI and weight lifted during the one-repetition deadlift (J. Pierce et al., 2018). Higher BMI was also associated with significant improvements in muscular strength and power among male and female soldiers and had no effect on performance in a 1600m loaded march or a warrior task and battle drill obstacle course (J. R.

Pierce et al., 2017). Interesting to note is that BMI was not related to performance on the most occupationally relevant tests. These authors reported that not meeting the Army Body Composition Program BMI standard only affect a minimal number, about 6%, of soldiers' ability to pass the APFT and therefore may require allowances on physical tasks with high military relevance (J. R. Pierce et al., 2017).

BMI Applicability

It is, therefore, questioned as to whether BMI is an adequate measure of body composition in soldiers, often referred to as tactical athletes. In male infantryman, the correlation between BMI and body fat percent measured by DEXA was significant at r =0.86 ($p<0.01$) (Grier, Canham-Chervak, Sharp, et al., 2015). In male firefighters, obesity estimated by BMI was similar to the of percent body fat and therefore was concluded as a good indicator of obesity (BMI > 30) (Porto et al., 2016). BMI has been cautioned for use in military personal, however, as it can overestimate body fatness (Grier, Canham-Chervak, Sharp, et al., 2015). Two soldiers of the same height and weight would be classified as the same BMI when, in fact, there FM and FFM levels could vary greatly. When body composition methods were tested among ROTC cadets, BMI was not correlated with body composition determined by BodPod or BIA, two generally more valid tools (Steed et al., 2016). Army Body Composition Program (ABCP) maximum standard BMIs for soldiers younger than 21 years old is 25.9 kg/m^2 and 25.0 kg/m^2 for males and females, respectively (Headquarters Department of the Army, 2019). While BMI may be a good indicator of obesity at values greater than 30 kg/m^2, for soldiers who must possess more muscle mass compared to the normal population, BMI may misclassify them as overweight or obese. This will cause them to fail ABCP standards even if they successfully pass all their fitness tests.

Conclusion

While the inherent physiological differences between males and females have been addressed in numerous research investigations over the years, no research published to date has analyzed these differences related to the ACFT. Military service requires both male and female soldiers to perform the same tasks that require enhanced levels of physical fitness. Furthermore, the nature of today's conflicts and women's new roles in combat require all soldiers to meet a minimum fitness standard, regardless of sex. Thus, for the purpose of identifying and minimizing disparities, the absolute and relative differences between males and females should be identified within a military population, so they might be addressed accordingly. Additionally, reconsidering how BMI relates to performance in regard to the ACFT may provide new information on anthropometric measures and their relationship to tests of strength, power, endurance, and anaerobic and aerobic capacity. These analyses may result in practical recommendations for improving performance across the sexes through training regimes and/or adjusted body composition attitudes.

Chapter III - Methodology

Introduction

Multiple studies have analyzed the differences between male and female performance outcomes on military fitness tests (Draicchio et al., 2020; Knapik et al., 2001; J. Pierce et al., 2018). Additionally, the applicability of body mass index (BMI) as an indicator of body composition and its relation to fitness in active populations has been discussed at length (Grier, Canham-Chervak, Sharp, et al., 2015; J. Pierce et al., 2018; J. R. Pierce et al., 2017). However, no studies published to date have assessed how males and females differ on the new Army Combat Fitness Test (ACFT) nor the role BMI may play in relation to performance on this test. Therefore, the purpose of this cross-sectional study was to identify if and where differences exist in ACFT scores between sexes and across military science classes as well as identify possible correlations between anthropometric variables and ACFT outcomes. Additionally, the purpose was to analyze if a semester of physical training significantly improves ACFT outcomes and what variables put a cadet at greater odds to fail or score in a particular qualification category. This chapter will discuss the sampling method and population used, the specifics of conducting an ACFT, the overall research design, and data collection and analysis procedures.

Sampling

Thirty-five ROTC programs from Cadet Command's 5th Brigade were invited to participate in this study. A convenience sample from Army ROTC cadets has been used in previous research when analyzing sex-based differences and body composition correlations with the APFT and OPAT (Draicchio et al., 2020; Steed et al., 2016). Only ACFT scores of cadets currently enrolled in an ROTC program will be used. Cadets with injuries were not included in the data set. Based on pilot study data provided to the researchers by the local University ROTC cadre, an a-priori power analysis (G*Power 3, Düsseldorf, Germany) to compute minimum sample size with a 0.85 effect size, a 1-β of 0.93, and an 80:20 male to female ratio required a minimum of 100 males and 20 females being included in our analysis (Faul et al., 2009).

Instrumentation and Measurement

Army Combat Fitness Test

The ACFT consists of a three-repetition maximum deadlift, standing power throw, hand-release t-pushups, sprint-drag-carry, leg tuck, and two-mile run (Headquarters Center for Initial Military Training, 2019). The ACFT was administered per published Army standards and events are always conducted in the aforementioned order (Headquarters Center for Initial Military Training, 2019). Equipment needed to conduct the ACFT includes a hexagon bar, bumper plates, barbell collars, 4.5-kg medicine ball, 25m measuring tape, 1m stick, 12 field cones, hand towel, stopwatch, nylon sled with pull straps, two 18-kg kettlebells, pull-up bar, and a paved, generally flat course, to conduct the two-mile run. The protocol for conducting each event has been published elsewhere (Headquarters Center for Initial Military Training, 2019) but is described below.

Three-Repetition Maximum Deadlift (MDL)

The MDL is conducted using a hex bar and bumper plates. Soldiers step inside the hex-bar and bend at the knees and hips to grasp the handles. With arms fully extended, back flat, head neutral with spine, and feet flat on the ground, the soldier lifts the bar by straightening hips and knees to reach the Straddle Stance. Back must remain straight, not rounded or flexed, and feet must remain in the same position. On downward movement, the soldier must remain in control of the bar and must not drop the weight to the ground. Three consecutive repetitions must be completed with the same weight and proper form. A ten-minute warm up period is allotted before this event.

Standing Power Throw (SPT)

The SPT is a two-attempt test that requires soldiers to perform an over the head, backwards medicine ball throw for distance. A preparatory countermovement is allowed but soldiers cannot step over the foul line. The furthest throw counted for record. The medicine ball is 10lbs.

Hand Release Pushup (HRP)

The HRP is a timed event that requires soldiers to complete as many HRPs as possible in two-minutes. To properly execute a HRP, a soldier began in the prone position, pushed entire body up as a single unit until elbows are fully extended, lowered oneself back to the ground as a single unit, released hands from ground and move them outwards in a T-position, and replaced hands back to shoulder width apart for one complete repetition.

Sprint-Drag-Carry (SDC)

The SDC is a time, 250m shuttle event consisting of 5x50m shuttles of various requirements. On the command “GO”, soldiers sprint 25m, touch the line, and sprint back 25m. Second, soldiers grasp the strap handles of a 90lb sled and drag the sled backwards down 25m and back 25m. Next, soldiers perform a lateral shuffle down 25m, switch lead legs, and shuffle back 25m. During the lateral shuffle, feet must not cross one another, must remain parallel to one another and perpendicular to the direction of travel. Then, soldiers grasp the handles of two 40lb kettlebells, sprint 25m down and 25m back, and place the kettlebells on the ground without dropping them. Finally, soldiers sprint down 25m and back 25m to the start line to finish the test.

If soldiers failed to touch the start or 25m line at any point, the grader would call them back to do so. Total time to complete the shuttles was recorded.

Leg Tuck

The leg tuck test required soldiers to do as many leg tucks as possible and the test will end when the soldier voluntarily stops. To complete a leg tuck, a soldier must straight arm hang from a pull-up bar with an alternating grip and flex at the elbows, knees, hips and waist to bring the knees or thighs up to touch both elbows. Then, they must return to a straight arm hang with controlled movement to complete one full repetition.

Two-Mile Run (2MR)

The 2MR required soldiers to run a two-mile course as fast as possible without any physical help (being picked up, pushed, or pulled by someone else). Pacing and verbal encouragement was permitted. Leaving the designated course would have resulted in termination and failure. The course was a solid, improved surface that is generally flat with no more than 3% uphill or downhill grade. Start and finish were at approximately the same altitude.

Rest Periods

The approximate overall work to rest is 34-37 minutes of work and 17 minutes of rest. Periods of rest were scheduled between each event with a minimum of 2 minutes between the MDL and SPT, 3 minutes between SPT and HRP, 3 minutes between HRP and SDC, 4 minutes between SDC and leg tuck, and 5 minutes between leg tuck and 2MR. This does not include MDL warm-up or event instruction.

Research Design

This is a retrospective data analysis utilizing previously recorded ACFT scores and anthropometric variables within Army ROTC Cadets. Army ROTC programs used in this study reported de-identified data. A possible threat to internet validity is instrumentation. Because this study sought to use retrospective data, the time, date, and manner in which the tests were given may vary. However, individual ROTC programs were expected to adhere to the published ACFT guidelines on how to conduct and score the test. Adhering to Center for Initial Military Training published protocols should have mitigated differences due to improper testing techniques (Headquarters Center for Initial Military Training, 2019).

Another threat to internal validity was environment. Because the data came from ROTC programs across various states within the United States, certain extraneous variables, such as COVID-19 restrictions or weather, may have affected one program's ACFT scores more than another's. A possible threat to external validity was the non-random sampling used in this study. The data and conclusions drawn from this study can only be able to be generalized to an Army ROTC population as no Active Duty, Reserve, or National Guard soldiers are included.

Data Collection Procedures

Thirty-five schools within Army Cadet Command's 5th Brigade were eligible to be included in this study. Army Combat Fitness Tests were conducted at the end of the semester in accordance with published ACFT protocol (Headquarters Center for Initial Military Training, 2019). The ROTC programs shared their ACFT scores with the researcher remotely. The researcher was not be present for the test. Data was de-identified by the programs prior to being shared and data is stored in a password protected computer.

Data Analysis Procedures

All data was analyzed using Statistical Package for the Social Sciences (SPSS, Version 26 (IBM, New York, USA). All data are reported as means ± SD unless otherwise noted. Statistical significance was set at $p \leq 0.05$. The data's normality of distribution was analyzed using the Shapiro-Wilk test. The data did not meet the assumption of normality. The data was Log-10 transformed but still failed to meet the assumption of normality. Thus, non-parametric tests were used for all analyses in this study. A Mann-Whitney U analysis was conducted to analyze differences between males and females on total ACFT scores, individual event scaled scores, individual event raw scores, and anthropometric variables of height, weight, and BMI. A Kruskal-Wallis H-test was used to analyze differences between the four military-science classes. Individual Mann-Whitney U analyses were used for post-hoc analysis. Manual Bonferroni calculated levels of significance were set at < 0.00833. Due to the non-parametric nature of the data, a 2x5 factorial ANOVA with Bonferroni post-hoc analysis could not be used to analyze a sex by military science class interaction.

A Spearman's rank correlation coefficient was conducted to analyze the relationship between anthropometric variables and ACFT outcomes for the combined population. A Spearman's rank correlation coefficient was also conducted to analyze the relationship between anthropometric variables and ACFT outcomes for each sex. The pre-semester vs. post-semester data also failed the assumption of normality. Therefore, Wilcoxon-signed ranks tests were used to determine if male and female cadets' post-semester ACFT scores were significantly different than their pre-semester scores. Stepwise multinomial logistic regression analyses were conducted to determine the significant predictors in an equation to determine likelihood of failing (category 0), or scoring in "MODERATE" (category 1), "SIGNIFICANT" (category 2), or "HEAVY" (category 3) for individual ACFT events and overall ACFT qualification category. The reference category was set to "HEAVY" (category 3) and the reference MS-class was set to MS-1 (freshmen). The predictors entered into the stepwise multinomial logistic regression analyses were sex, height, mass, and MS-class.

Chapter IV – Results and Discussion

Results

Population Characteristics

In total, complete ACFT scores were received for 1061 cadets from 20 different universities. Complete ACFT scores and sex of each cadet was received for 1013 cadets, consisting of 757 males and 256 females. Age was reported for 694 males and 232 females, height for 502 males and 178 females, and mass for 480 males and 174 females. Age, along with, all anthropometric measures are shown in Table 1. Data are reported as mean ± standard deviation.

Table 1: Age and anthropometric measures for male and female cadets

	Males	Females	P-value	Effect Size
N=	757	256		
Age	21.46 ± 3.27	20.99 ± 3.34	0.003*	0.14
Height(m)	1.77 ± 0.073	1.63 ± 0.072	0.000*	1.93
Mass (kg)	77.73 ± 11.65	63.41 ± 9.97	0.000*	1.32
BMI	24.73 ± 3.17	23.72 ± 3.26	0.000*	0.31

Data are presented as Mean ± SD
* = denotes significant difference ($p \leq 0.05$) from male cadets

ACFT Scores and Sex Differences

Mann-Whitney U analyses showed males and females to be significantly different on all ACFT events for both raw and scaled scores, deadlift expressed relative to body mass, total scores, and qualification categories ($p < 0.001$). Comparisons between sexes for individual event scores and scaled total scores are located in Figure 1 and Figure 2, respectively. Large effect sizes were observed for all differences except two-mile run points which exhibited a medium effect size. Male vs. female means and comparisons are located in Table 2 and Table 3. Overall, 124 females failed the ACFT, representing 48.4% of the total population whereas only 59 males failed the ACFT, representing 7.8% of the population. Only 3 females (1.1% of all females), qualified for the heaviest category, while 379 males (50% of all males), qualified for the heaviest category (Figure 3).

The leg tuck was the most failed event by females, with a total of 102 failures (39.8%). In contrast, the two-mile run was the most failed event by males, with a total of 32 failures (4.2%). The least failed event for both sexes was the hand release push-up with 2 female and 0 male failures. Percent of each sex that failed per event can be found in Table 4. Percent of males and females that scored in the "HEAVY" category per event can be found in Table 5. When the leg tuck was eliminated from the test, only 75 females failed the ACFT (29.29% of population vs. 48.4% when the leg tuck was included) which increased the number of those qualifying for the moderate category by 41, and those qualifying for the significant category by 8.

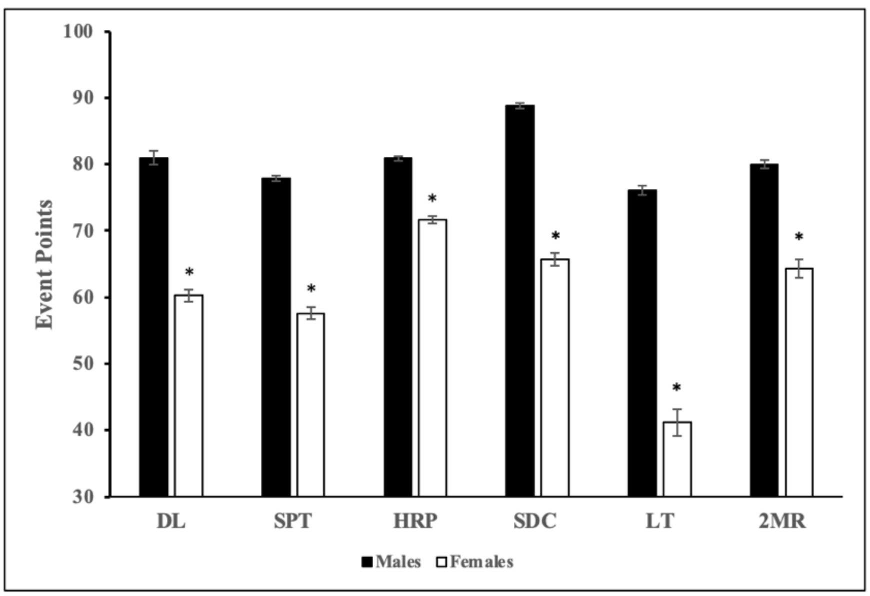

Figure 1: Differences between males and females for ACFT events scored on a 0-100 scale.

Data are presented as mean ± SE.
* = denotes significant difference ($p \leq 0.05$) from male cadets

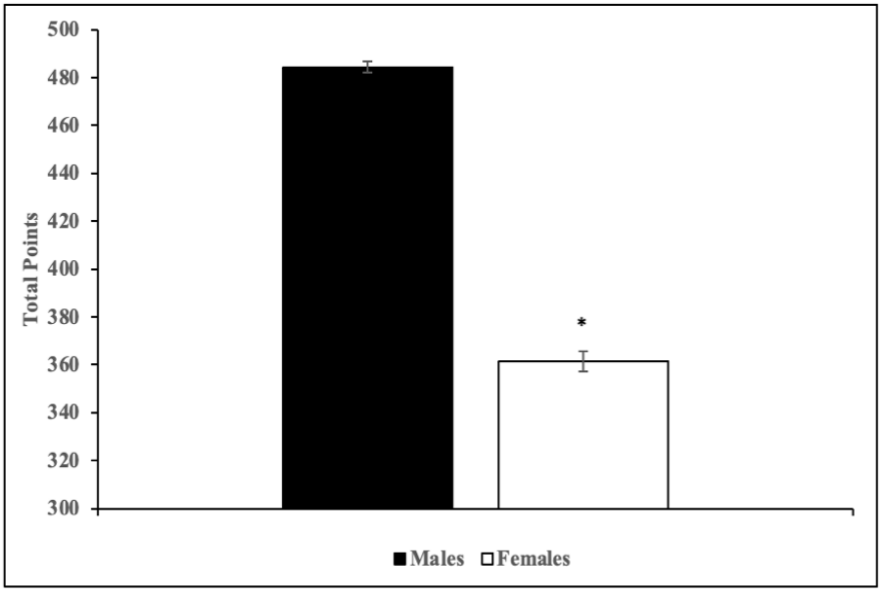

Figure 2: Differences between males and females for total ACFT score.

Data are presented as mean ± SE.
* = denotes significant difference ($p \leq 0.05$) from male cadets

Table 2: Differences between males and females for ACFT events scored on a 0-100 scale

	Males	Female	P-value	Effect size
DL Points	80.94 ± 28.38	60.26 ± 14.50	0.000*	0.92
SPT Points	77.89 ± 11.58	57.61 ± 15.01	0.000*	1.51
HRP Points	80.85 ± 10.35	71.64 ± 8.95	0.000*	0.95
SDC Points	88.78 ± 11.45	65.73 ± 14.83	0.000*	1.74
LT Points	76.12 ± 19.69	41.21 ± 32.27	0.000*	1.31
2MR Points	80.01 ± 17.67	64.33 ± 22.2	0.000*	0.78
Total Points	484.38 ± 64.14	361.46 ± 70.62	0.000*	1.82
Qual Level	2.11 ± 1.00	0.57 ± 0.62	0.000*	1.85

Data are presented as mean ± SE.
* = denotes significant difference ($p \leq 0.05$) from male cadets

Table 3: Differences between males and females for ACFT raw event performances

	Males	Female	P-value	Effect Size
DL Raw (lbs)	248.13 ± 59.28	155.02 ± 43.14	0.000*	1.79
DL Raw (%BW)	143.77 ± 32.66	114.64 ± 30.45	0.000*	0.92
SPT Raw (m)	9.21 ± 3.47	5.55 ± 1.28	0.000*	1.40
HRP Raw (reps)	40.41 ± 11.56	30.21 ± 11.00	0.000*	0.90
SDC Raw (s)	109.42 ± 18.12	147.99 ± 26.24	0.000*	1.71
LT Raw (reps)	9.48 ± 6.17	2.11 ± 3.15	0.000*	1.50
2MR Raw (s)	977.4 ± 143.47	1117.2 ± 148.76	0.000*	0.96

Data are presented as mean ± SE.
* = denotes significant difference ($p \leq 0.05$) from male cadets

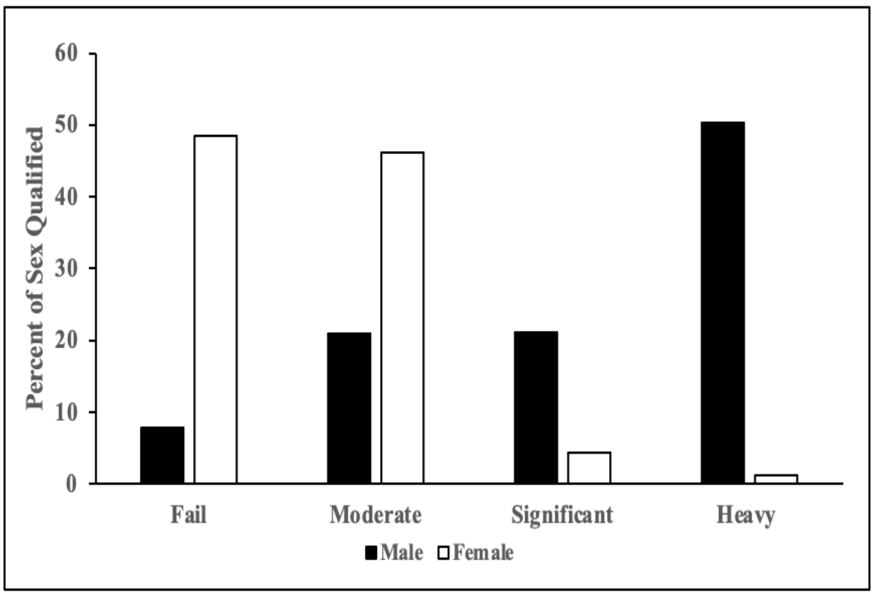

Figure 3: Percent of males and females qualified for each ACFT category

Table 4: Percent of failures per ACFT event and overall separated by sex

	MDL	SPT	HRP	SDC	LT	2MR	Overall
Male Failures	0.3%	0.7%	0.0%	0.1%	4.0%	4.2%	7.8%
Female Failures	5.5%	16.4%	0.8%	9.8%	39.8%	14.0%	48.4%

Table 5: Percent of cadets scoring in the "HEAVY" category per ACFT event and overall separated by sex

	MDL	SPT	HRP	SDC	LT	2MR	Overall
Male Heavy	81.8%	72.7%	86.1%	88.8%	77.1%	81.1%	50.0%
Female Heavy	13.3%	4.7%	57.8%	24.6%	20.7%	44.1%	1.2%

Military Science Class Comparisons

MS Class and Age, Height, Mass, BMI

MS-class designation was reported for 822 cadets with age and anthropometric averages presented in Table 6. A Kruskal-Wallis Test with individual Mann-Whitney U post-hoc analyses showed all MS-classes to be significantly different in age when compared using a manual Bonferroni calculated significance level of < 0.00833 (0.05 adjusted for number of comparisons). MS-classes showed no significant difference in height and only MS-2 vs. MS-4 cadets were significantly different in mass, with MS-4 cadets being heavier, on average (71.8kg ± 13.7kg v. 74.8kg ± 11.6kg, p=0.000). Likewise, MS-2 and MS-4 cadets presented significantly different body mass indexes (23.8 ± 3.5 v. 24.7 ± 2.8, p = 0.001) (Table 7-10).

Table 6: MS-class age and anthropometric variables

	MS-1	**MS-2**	**MS-3**	**MS-4**
N=	138	159	268	257
Age	18.9 ± 2.0	20.1 ± 2.2	22.2 ± 3.5	23.1 ± 3.3
Height (m)	1.7 ± 0.3	1.7 ± 0.2	1.7 ± 0.1	1.7 ± 0.1
Mass (kg)	73.4 ± 12.5	71.8 ± 13.7	74.0 ± 12.9	74.8 ± 11.6
BMI	24.1 ± 3.2	23.8 ± 3.5	24.5 ± 3.1	24.7 ± 2.8

Data are presented as mean ± SD.

Table 7: Age differences between MS-classes, observed effect sizes, and statistical power estimates

Age Comparison	**Mean ± SD**	**P-value**	**Effect Size**	**1 - β**
MS 1 v. MS2	18.9 ± 2.0 v. 20.1 ± 2.2	0.000*	0.57	0.99
MS 1 v. MS3	18.9 ± 2.0 v. 22.2 ± 3.5	0.000*	1.16	1.00
MS 1 v. MS4	18.9 ± 2.0 v. 23.1 ± 3.3	0.000*	1.54	1.00
MS 2 v. MS3	20.1 ± 2.2 v. 22.2 ± 3.5	0.000*	0.72	1.00
MS 2 v. MS4	20.1 ± 2.2 v. 23.1 ± 3.3	0.000*	1.07	1.00
MS 3 v. MS4	22.2 ± 3.5 v. 23.1 ± 3.3	0.000*	0.26	0.78

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 8: Height differences between MS-classes, observed effect sizes, and statistical power estimates

Height Comparison (m)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	1.7 ± 0.3 v. 1.7 ± 0.2	0.318	0.00	0.05
MS 1 v. MS3	1.7 ± 0.3 v. 1.7 ± 0.1	0.477	0.00	0.05
MS 1 v. MS4	1.7 ± 0.3 v. 1.7 ± 0.1	0.668	0.00	0.05
MS 2 v. MS3	1.7 ± 0.2 v. 1.7 ± 0.1	0.573	0.00	0.05
MS 2 v. MS4	1.7 ± 0.2 v. 1.7 ± 0.1	0.419	0.00	0.05
MS 3 v. MS4	1.7 ± 0.1 v. 1.7 ± 0.1	0.721	0.00	0.05

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 9: Mass differences between MS-classes, observed effect sizes, and statistical power estimates

Mass Comparison (kg)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	73.4 ± 12.5 v. 71.8 ± 13.7	0.120	0.12	0.12
MS 1 v. MS3	73.4 ± 12.5 v. 74.0 ± 12.9	0.727	0.05	0.06
MS 1 v. MS4	73.4 ± 12.5 v. 74.8 ± 11.6	0.138	0.12	0.13
MS 2 v. MS3	71.8 ± 13.7 v. 74.0 ± 12.9	0.051	0.17	0.25
MS 2 v. MS4	71.8 ± 13.7 v. 74.8 ± 11.6	0.001*	0.24	0.44
MS 3 v. MS4	74.0 ± 12.9 v. 74.8 ± 11.6	0.243	0.06	0.09

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 10: BMI differences between MS-classes, observed effect sides, and statistical power estimates

BMI Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	24.1 ± 3.2 v. 23.8 ± 3.5	0.285	0.09	0.09
MS 1 v. MS3	24.1 ± 3.2 v. 24.5 ± 3.1	0.278	0.13	0.15
MS 1 v. MS4	24.1 ± 3.2 v. 24.7 ± 2.8	0.030	0.20	0.29
MS 2 v. MS3	23.8 ± 3.5 v. 24.5 ± 3.1	0.013	0.21	0.38
MS 2 v. MS4	23.8 ± 3.5 v. 24.7 ± 2.8	0.001*	0.28	0.59
MS 3 v. MS4	24.5 ± 3.1 v. 24.7 ± 2.8	0.263	0.07	0.09

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

MS Class and Deadlift

Military science classes were not significantly different for the scaled score, raw score, or relative weight lifted on the three-repetition maximum deadlift (Table 11-13 and Figure 4a).

MS Class and Standing Power Throw

MS-1 and MS-2 cadets had significantly less standing power throw points and shorter raw throw distance than MS-3 and MS-4 cadets, with neither the underclassmen nor the upperclassmen being significantly different from one another (Table 14-15 and Figure 4b).

MS Class and Hand-Release Pushups

Military science classes were not significantly different for both the scaled score and raw score hand-release pushups (Table 16-17 and Figure 4c).

MS Class and Sprint-Drag-Carry

MS-1 cadets had significantly lower scores (78.1 ± 19.0 v. 83.9 ± 15.5, $p = 0.002$; 78.1 ± 19.0 v. 84.5 ± 14.6, $p < 0.001$) and slower times (127.1 ± 30.2 v. 117.6 ± 25.4, $p = 0.002$; 127.1 ± 30.2 v. 115.9 ± 22.8, $p < 0.001$) than the MS-3 and MS-4 cadets, respectively, on the sprint-drag-carry (Table 18-19 and Figure 4d).

MS Class and Leg Tuck

Military science classes were not significantly different for both the scaled score and raw score leg tuck event (Table 20-21 and Figure 4e).

MS Class and 2-Mile Run

MS-1 cadets had significantly lower scores (69.6 ± 26.4 v. 79.2 ± 13.2, $p = 0.004$; 69.6 ± 26.4 v. 79.1 ± 16.4, $p = 0.001$) and slower times (1059.8 ± 195.0 v. 995.0 ± 125.7, $p = 0.002$; 1059.8 ± 195.0 v. 991.7 ± 130.3, $p = 0.002$) than the MS-3 and MS-4 cadets, respectively, on the two-mile run (Table 22-23 and Figure 4f).

MS Class and Total ACFT Scores

MS-1 cadets also had significantly lower total scores on the ACFT than MS-3 and MS-4 cadets (426.9 ± 93.6 v. 463.2 ± 75.4, $p < 0.001$; 426.9 ± 93.6 v. 464.3 ± 70.1, $p < 0.001$, respectively) (Table 24 and Figure 5). MS-4 cadets had the highest total average but were not statistically different than MS-3 cadets.

MS Class and Qualification Category

MS-1 cadets had significantly lower average qualification categories, on a scale of 0-3, than MS-3 and MS-4 cadets (1.38 ± 1.17 v. 1.86 ± 1.11, $p < 0.001$; 1.38 ± 1.17 v. 1.90 ± 1.02, $p < 0.001$, respectively) (Table 25). The MS-3 class had the greatest percent of cadets, 41%, reach the heavy qualification category and the MS-1 class had the greatest percent of cadets, 28%, fail the test. The MS-4 class had only 9% failures, the least out of all classes (Table 26).

Table 11: MDL points differences between MS-classes

DL Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	71.2 ± 19.3 v. 74.6 ± 13.7	0.349	0.20	0.35
MS 1 v. MS3	71.2 ± 19.3 v. 77.0 ± 14.4	0.016	0.34	0.85
MS 1 v. MS4	71.2 ± 19.3 v. 75.8 ± 13.8	0.104	0.27	0.67
MS 2 v. MS3	74.6 ± 13.7 v. 77.0 ± 14.4	0.158	0.17	0.35
MS 2 v. MS4	74.6 ± 13.7 v. 75.8 ± 13.8	0.467	0.09	0.13
MS 3 v. MS4	77.0 ± 14.4 v. 75.8 ± 13.8	0.387	0.09	0.15

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 12: MDL raw differences between MS-classes

DL Raw Comparison (lbs)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	211.1 ± 73.4 v. 220.3 ± 63.6	0.370	0.13	0.19
MS 1 v. MS3	211.1 ± 73.4 v. 231.8 ± 68.4	0.018	0.29	0.73
MS 1 v. MS4	211.1 ± 73.4 v. 225.8 ± 66.5	0.111	0.21	0.45
MS 2 v. MS3	220.3 ± 63.6 v. 231.8 ± 68.4	0.148	0.17	0.36
MS 2 v. MS4	220.3 ± 63.6 v. 225.8 ± 66.5	0.470	0.08	0.12
MS 3 v. MS4	231.8 ± 68.4 v. 225.8 ± 66.5	0.384	0.09	0.16

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 13: Relative MDL differences between MS-classes

DL Raw Comparison (%BW)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	129.9 ± 36.5 v. 136.0 ± 35.4	0.345	0.17	0.41
MS 1 v. MS3	129.9 ± 36.5 v. 140.2 ± 33.9	0.051	0.29	0.86
MS 1 v. MS4	129.9 ± 36.5 v. 134.2 ± 32.8	0.518	0.12	0.31
MS 2 v. MS3	136.0 ± 35.4 v. 140.2 ± 33.9	0.347	0.12	0.32
MS 2 v. MS4	136.0 ± 35.4 v. 134.2 ± 32.8	0.654	0.05	0.13
MS 3 v. MS4	140.2 ± 33.9 v. 134.2 ± 32.8	0.111	0.18	0.64

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 14: SPT points differences between MS-classes

SPT Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	68.0 ± 16.5 v. 71.4 ± 13.8	0.171	0.22	0.42
MS 1 v. MS3	68.0 ± 16.5 v. 74.0 ± 16.0	0.000*	0.37	0.90
MS 1 v. MS4	68.0 ± 16.5 v. 75.2 ± 13.1	0.000*	0.48	0.99
MS 2 v. MS3	71.4 ± 13.8 v. 74.0 ± 16.0	0.007*	0.17	0.36
MS 2 v. MS4	71.4 ± 13.8 v. 75.2 ± 13.1	0.002*	0.28	0.74
MS 3 v. MS4	74.0 ± 16.0 v. 75.2 ± 13.1	0.699	0.08	0.14

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 15: SPT Raw differences between MS-classes

SPT Raw Comparison (m)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	7.5 ± 2.3 v. 7.9 ± 2.3	0.190	0.17	0.28
MS 1 v. MS3	7.5 ± 2.3 v. 8.5 ± 2.4	0.000*	0.42	0.96
MS 1 v. MS4	7.5 ± 2.3 v. 8.6 ± 2.3	0.000*	0.48	0.99
MS 2 v. MS3	7.9 ± 2.3 v. 8.5 ± 2.4	0.007*	0.26	0.68
MS 2 v. MS4	7.9 ± 2.3 v. 8.6 ± 2.3	0.002*	0.30	0.80
MS 3 v. MS4	8.5 ± 2.4 v. 8.6 ± 2.3	0.647	0.04	0.07

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 16: HRP points differences between MS-classes

HRP Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	76.8 ± 9.1 v. 78.6 ± 10.4	0.184	0.18	0.31
MS 1 v. MS3	76.8 ± 9.1 v. 79.3 ± 10.7	0.018	0.25	0.61
MS 1 v. MS4	76.8 ± 9.1 v. 78.2 ± 11.9	0.237	0.13	0.21
MS 2 v. MS3	78.6 ± 10.4 v. 79.3 ± 10.7	0.403	0.26	0.66
MS 2 v. MS4	78.6 ± 10.4 v. 78.2 ± 11.9	0.816	0.04	0.06
MS 3 v. MS4	79.3 ± 10.7 v. 78.2 ± 11.9	0.259	0.10	0.18

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 17: HRP Raw differences between MS-classes

HRP Raw Comparison (reps)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	36.0 ± 10.5 v. 37.8 ± 11.7	0.185	0.16	0.25
MS 1 v. MS3	36.0 ± 10.5 v. 39.0 ± 11.5	0.014	0.27	0.67
MS 1 v. MS4	36.0 ± 10.5 v. 37.2 ± 14.3	0.214	0.24	0.56
MS 2 v. MS3	37.8 ± 11.7 v. 39.0 ± 11.5	0.346	0.10	0.16
MS 2 v. MS4	37.8 ± 11.7 v. 37.2 ± 14.3	0.884	0.05	0.07
MS 3 v. MS4	39.0 ± 11.5 v. 37.2 ± 14.3	0.249	0.14	0.31

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 18: SDC points differences between MS-classes

SDC Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	78.1 ± 19.0 v. 82.3 ± 14.8	0.085	0.25	0.50
MS 1 v. MS3	78.1 ± 19.0 v. 83.9 ± 15.5	0.002*	0.33	0.85
MS 1 v. MS4	78.1 ± 19.0 v. 84.5 ± 14.6	0.000*	0.38	0.91
MS 2 v. MS3	82.3 ± 14.8 v. 83.9 ± 15.5	0.206	0.11	0.16
MS 2 v. MS4	82.3 ± 14.8 v. 84.5 ± 14.6	0.115	0.15	0.28
MS 3 v. MS4	83.9 ± 15.5 v. 84.5 ± 14.6	0.880	0.04	0.07

Data are presented as mean ± SD.
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 19: SDC Raw differences between MS-classes

SDC Raw Comparison (s)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	127.1 ± 30.2 v. 120.4 ± 25.0	0.078	0.24	0.49
MS 1 v. MS3	127.1 ± 30.2 v. 117.6 ± 25.4	0.002*	0.34	0.85
MS 1 v. MS4	127.1 ± 30.2 v. 115.9 ± 22.8	0.000*	0.42	0.96
MS 2 v. MS3	120.4 ± 25.0 v. 117.6 ± 25.4	0.184	0.11	0.18
MS 2 v. MS4	120.4 ± 25.0 v. 115.9 ± 22.8	0.103	0.19	0.41
MS 3 v. MS4	117.6 ± 25.4 v. 115.9 ± 22.8	0.877	0.07	0.12

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 20: LT points differences between MS-classes

LT Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	63.2 ± 30.6 v. 65.0 ± 28.1	0.714	0.06	0.08
MS 1 v. MS3	63.2 ± 30.6 v. 69.9 ± 26.9	0.021	0.23	0.54
MS 1 v. MS4	63.2 ± 30.6 v. 71.5 ± 23.0	0.025	0.31	0.77
MS 2 v. MS3	65.0 ± 28.1 v. 69.9 ± 26.9	0.029	0.18	0.38
MS 2 v. MS4	65.0 ± 28.1 v. 71.5 ± 23.0	0.047	0.25	0.64
MS 3 v. MS4	69.9 ± 26.9 v. 71.5 ± 23.0	0.805	0.06	0.10

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 21: LT Raw differences between MS-classes

LT Raw Comparison (reps)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	6.8 ± 6.3 v. 7.0 ± 5.9	0.605	0.03	0.06
MS 1 v. MS3	6.8 ± 6.3 v. 8.3 ± 6.5	0.025	0.23	0.55
MS 1 v. MS4	6.8 ± 6.3 v. 8.1 ± 6.3	0.025	0.21	0.44
MS 2 v. MS3	7.0 ± 5.9 v. 8.3 ± 6.5	0.052	0.21	0.49
MS 2 v. MS4	7.0 ± 5.9 v. 8.1 ± 6.3	0.070	0.18	0.38
MS 3 v. MS4	8.3 ± 6.5 v. 8.1 ± 6.3	0.869	0.03	0.06

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 22: 2MR points differences between MS-classes

2MR Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	69.6 ± 26.4 v. 75.2 ± 20.2	0.232	0.24	0.47
MS 1 v. MS3	69.6 ± 26.4 v. 79.2 ± 13.2	0.004*	0.46	0.98
MS 1 v. MS4	69.6 ± 26.4 v. 79.1 ± 16.4	0.001*	0.43	0.97
MS 2 v. MS3	75.2 ± 20.2 v. 79.2 ± 13.2	0.108	0.23	0.58
MS 2 v. MS4	75.2 ± 20.2 v. 79.1 ± 16.4	0.038	0.21	0.50
MS 3 v. MS4	79.2 ± 13.2 v. 79.1 ± 16.4	0.533	0.01	0.05

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 23: 2MR raw differences between MS-classes

2MR Raw Comparison (s)	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	1059.8 ± 195.0 v. 1014.5 ± 174.8	0.168	0.24	0.49
MS 1 v. MS3	1059.8 ± 195.0 v. 995.0 ± 125.7	0.002*	0.39	0.93
MS 1 v. MS4	1059.8 ± 195.0 v. 991.7 ± 130.3	0.002*	0.41	0.95
MS 2 v. MS3	1014.5 ± 174.8 v. 995.0 ± 125.7	0.112	0.13	0.22
MS 2 v. MS4	1014.5 ± 174.8 v. 991.7 ± 130.3	0.087	0.15	0.27
MS 3 v. MS4	995.0 ± 125.7 v. 991.7 ± 130.3	0.873	0.03	0.06

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 24: Total points differences between MS-classes

Total Points Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	426.9 ± 93.6 v. 451.4 ± 96.5	0.146	0.26	0.54
MS 1 v. MS3	426.9 ± 93.6 v. 463.2 ± 75.4	0.000*	0.43	0.97
MS 1 v. MS4	426.9 ± 93.6 v. 464.3 ± 70.1	0.000*	0.45	0.98
MS 2 v. MS3	451.4 ± 96.5 v. 463.2 ± 75.4	0.019	0.13	0.24
MS 2 v. MS4	451.4 ± 96.5 v. 464.3 ± 70.1	0.028	0.15	0.29
MS 3 v. MS4	463.2 ± 75.4 v. 464.3 ± 70.1	0.830	0.02	0.05

Data are presented as mean ± SD.
* = denotes significant difference (p ≤ 0.00833) between MS-classes

Table 25: Total Qualification differences between MS-classes

Total Qual Comparison	Mean ± SD	P-value	Effect Size	1 - β
MS 1 v. MS2	1.38 ± 1.17 v. 1.58 ± 1.16	0.139	0.17	0.42
MS 1 v. MS3	1.38 ± 1.17 v. 1.86 ± 1.11	0.000*	0.42	0.99
MS 1 v. MS4	1.38 ± 1.17 v. 1.90 ± 1.02	0.000*	0.47	1.00
MS 2 v. MS3	1.58 ± 1.16 v. 1.86 ± 1.11	0.017	0.25	0.78
MS 2 v. MS4	1.58 ± 1.16 v. 1.90 ± 1.02	0.006	0.29	0.88
MS 3 v. MS4	1.86 ± 1.11 v. 1.90 ± 1.02	0.776	0.04	0.11

Data are presented as mean ± SD
* = denotes significant difference ($p \leq 0.00833$) between MS-classes

Table 26: Percent of each MS-class that scored in each qualification category

	MS-1	MS-2	MS-3	MS-4
Fail	28.3%	22.6%	14.2%	9.3%
Moderate	32.6%	28.3%	27.2%	29.2%
Significant	11.6%	17.0%	17.2%	23.3%
Heavy	27.5%	32.1%	41.4%	38.1%

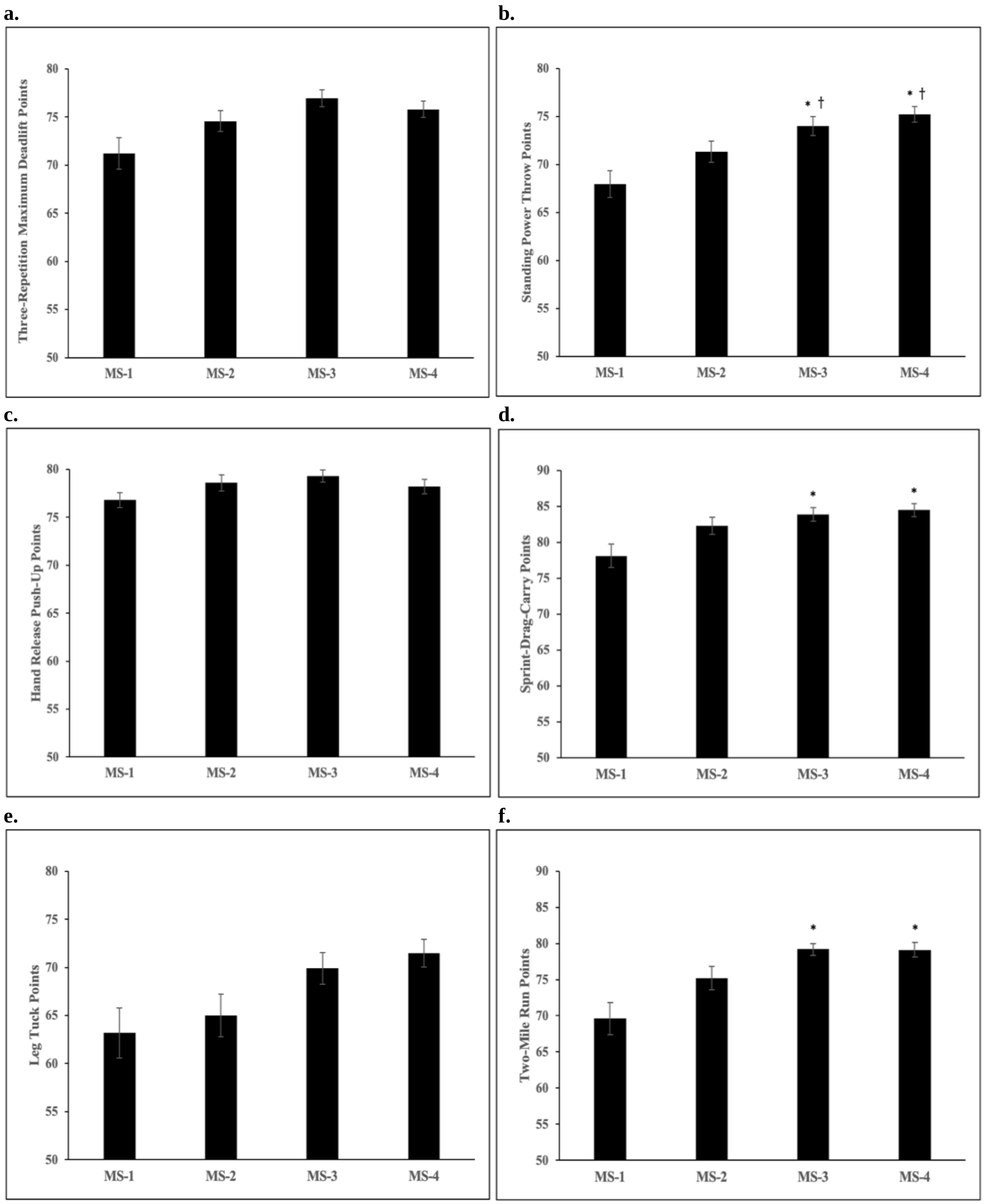

Figure 4: Differences between MS-classes for individual events points

Data are represented as mean ± SE**.**
* = denotes significant difference ($p \leq 0.00833$) compared to MS-1 class
† = denotes significant difference ($p \leq 0.00833$) compared to MS-2 class

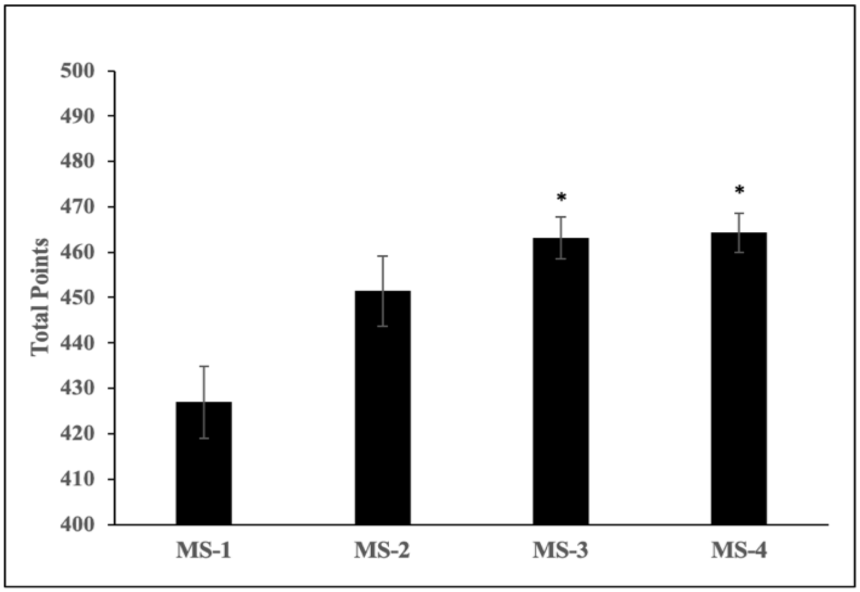

Figure 5: Differences between MS-classes for total ACFT scores

Data are represented as mean ± SE.
* = denotes significant difference ($p \leq 0.00833$) compared to MS-1 class

Anthropometric Correlations

Height, mass, and BMI were all significantly correlated ($p < 0.05$) with ACFT outcomes except for BMI vs. leg tuck raw scores, leg tuck points, and the two-mile run (Table 27-28). Height had moderate positive correlations ($R_s > 0.5$) with three-repetition maximum deadlift raw scores and standing power throw raw scores and a moderate negative correlation with sprint-drag-carry raw scores (R_s = 0.502, 0.618, and -0.593, respectively) (Table 27). Additionally, height had moderate positive correlations with standing power throw points, sprint-drag-carry points, and total ACFT scores (R_s = 0.604, 0.584, and 0.514, respectively) (Table 27). Mass had moderate positive correlations with three-repetition maximum deadlift raw scores and standing power throw raw scores and moderate negative correlations with sprint-drag-carry raw scores (R_s = 0.583, 0.643, and -0.512, respectively) (Table 28 and Figure 7a-c). Furthermore, mass had moderate positive correlations with three-repetition maximum deadlift points, standing power throw points, and sprint drag carry points (R_s = 0.580, 0.634, and 0.503, respectively) (Table 28 and Figure 7d-f). Total ACFT scores exhibited a moderate correlation with mass, $R_s = 0.444$ (Figure 6). When separated by sex, anthropometric measures showed statistically significant, but weak to moderate, correlations with ACFT outcomes (Table 29-32).

Table 27: Spearman Rho values for correlations between ACFT event points and anthropometric variables

	DL Pts	SPT Pts	HRP Pts	SDC Pts	LT Pts	2MR Pts	Total Pts	Qual
Height	0.499*	**0.604***	0.178*	**0.584***	0.257*	0.320*	**0.514***	0.470*
Mass	**0.580***	**0.634***	0.195*	**0.503***	0.172*	0.155*	0.444*	0.379*
BMI	0.343*	0.323*	0.107*	0.173*	-0.002	-0.065	0.148*	0.098*

* = denotes statistically significant ($p \leq 0.05$)
Bold = denotes moderate (> 0.5) correlations

Table 28: Spearman Rho values for correlations between ACFT event raw scores and anthropometric variables

	DL Raw	SPT Raw	HRP Raw	SDC Raw	LT Raw	2MR Raw
Height	**0.502***	**0.618***	0.166*	**-0.593***	0.245*	-0.325*
Mass	**0.583***	**0.643***	0.182*	**-0.512***	0.173*	-0.150*
BMI	0.343*	0.325*	0.099*	-0.176*	0.009	0.077*

* = denotes statistically significant ($p \leq 0.05$)
Bold = denotes moderate (> 0.5) correlations

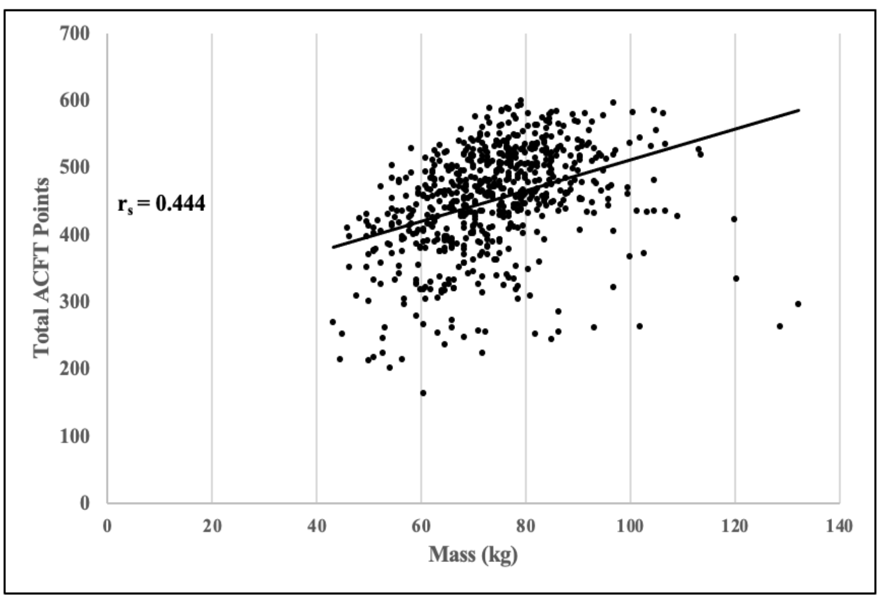

Figure 6: Total ACFT scores plotted versus mass, with trendline

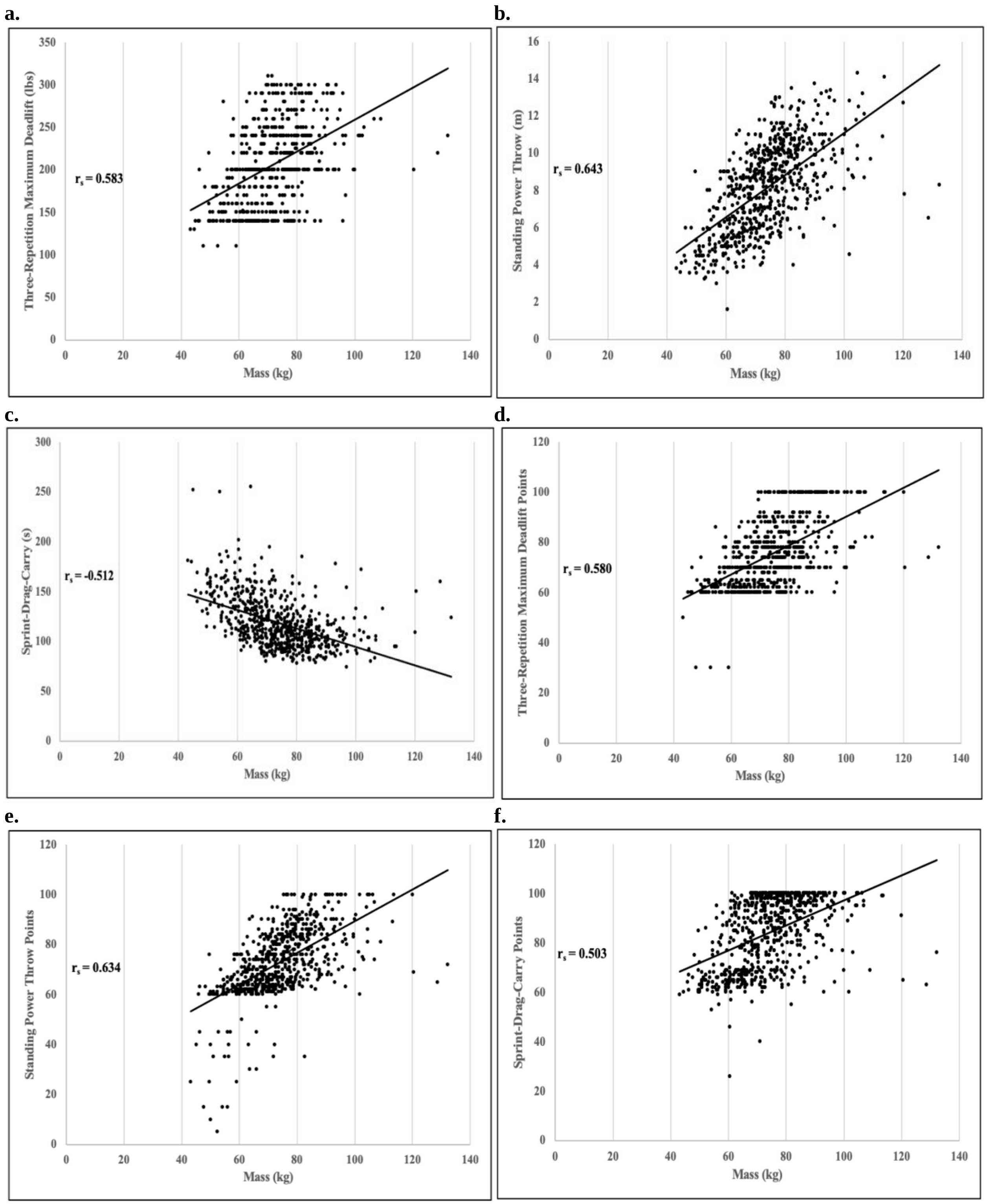

Figure 7a-f: ACFT event raw and scaled scores plotted versus mass (kg), with trendline

Table 29: Spearman's Rho values for correlations between males' ACFT event points and anthropometric variables

Males	DL Pts	SPT Pts	HRP Pts	SDC Pts	LT Pts	2MR Pts	Total Pts	Qual
Height	0.181*	0.351*	-0.065	0.315*	0.026	0.081	0.188*	0.169*
Mass	0.411*	0.488*	0.046	0.263*	-0.074	-0.062	0.206*	0.142*
BMI	0.354*	0.325*	0.080	0.087	-0.113*	-0.129*	0.098*	0.035

* = denotes statistically significant ($p \leq 0.05$)

Table 30: Spearman's Rho values for correlations between males' ACFT event raw scores and anthropometric variables

Males	DL Raw	SPT Raw	HRP Raw	SDC Raw	LT Raw	2MR Raw
Height	0.188*	0.356*	-0.066	-0.324*	0.009	-0.085
Mass	0.419*	0.489*	0.044	-0.271*	-0.076	0.072
BMI	0.357*	0.322*	0.079	-0.090*	-0.099*	0.150*

* = denotes statistically significant ($p \leq 0.05$)

Table 31*:* Spearman's Rho values for correlations between females' ACFT event points and anthropometric variables

Females	DL Pts	SPT Pts	HRP Pts	SDC Pts	LT Pts	2MR Pts	Total Pts	Qual
Height	0.238*	0.291*	-0.126	0.259*	-0.069	-0.056	0.098	0.069
Mass	0.282*	0.405*	-0.118	0.243*	-0.114	-0.283*	-0.028	-0.026
BMI	0.197*	0.275*	-0.028	0.118	-0.112	-0.277*	-0.097	-0.082

* = denotes statistically significant ($p \leq 0.05$)

Table 32: Spearman's Rho values for correlations between females' ACFT event raw scores and anthropometric variables

Females	DL Raw	SPT Raw	HRP Raw	SDC Raw	LT Raw	2MR Raw
Height	0.231*	0.341*	-0.151*	-0.286*	-0.073	0.041
Mass	0.273*	0.442*	-0.149*	-0.251*	-0.117	0.259*
BMI	0.192*	0.283*	-0.054	-0.116	-0.114	0.258*

* = denotes statistically significant ($p \leq 0.05$)

Pre and Post ACFT Comparisons

Pre- and post-semester test scores were obtained for 95 cadets from two universities, which included 74 males and 21 females. A Wilcoxon signed ranks test was used to determine if differences results from training across the Fall semester. Males significantly improved 3RM deadlift raw scores, SDC raw scores, 3RM deadlift points, SDC points, and 2MR points ($p < 0.001$) (Table 33 and Figure 8). Additionally, male cadets significantly improved their total ACFT score and their average qualification category ($p < 0.001$) (Figure 10). Female cadets significantly improved their SDC raw score, LT raw score, SPT points, SDC points, and LT points (Table 34 and Figure 9). Furthermore, female cadets significantly improved their total ACFT score and their average qualification category ($p < 0.001$ and $p = 0.046$, respectively) (Figure 10).

Table 33: Differences between males' pre- and post-semester ACFT scores

Males	Pre	Post	P-Value	Effect Size	1 - β
DL Raw (lbs)	236.96 ± 61.89	255.00 ± 62.95	0.000*	0.30	0.67
SPT Raw (m)	8.31 ± 2.04	8.45 ± 1.99	0.198	0.07	0.08
HRP Raw (reps)	38.51 ± 11.42	39.22 ± 10.59	0.478	0.06	0.08
SDC Raw (s)	124.17 ± 20.12	113.92 ± 15.7	0.000*	0.56	0.99
LT Raw (reps)	7.91 ± 6.49	8.47 ± 6.40	0.161	0.09	0.11
2MR Raw (s)	1021.28 ± 155.43	1002.77 ± 126.97	0.074	0.13	0.17
DL Points	77.93 ± 12.42	81.65 ± 12.91	0.000*	0.29	0.64
SPT Points	73.84 ± 10.67	74.91 ± 11.15	0.109	0.10	0.12
HRP Points	79.30 ± 10.09	79.72 ± 9.82	0.521	0.04	0.06
SDC Points	79.51 ± 12.54	85.35 ± 11.33	0.000*	0.49	0.97
LT Points	67.15 ± 28.93	70.89 ± 25.99	0.053	0.14	0.19
2MR Points	74.22 ± 22.03	77.66 ± 16.91	0.024*	0.17	0.27
Total Points	451.95 ± 72.39	470.18 ± 64.93	0.000*	0.26	0.55
Qualify Level	1.66 ± 1.09	2.04 ± 1.05	0.000*	0.35	0.80

Data are presented as mean ± SD
* = denotes statistical difference ($p \le 0.05$) between pre- and post- semester scores

Table 34: Differences between females' pre- and post-semester ACFT scores

Females	Pre	Post	P-value	Effect Size	1 - β
DL Raw (lbs)	141.90 ± 15.04	136.67 ± 47.26	0.130	0.13	0.08
SPT Raw (m)	4.55 ± 1.25	4.93 ± 0.83	0.092	0.34	0.28
HRP Raw (reps)	28.24 ± 7.95	31.05 ± 7.57	0.052	0.36	0.31
SDC Raw (s)	178.00 ± 50.78	147.29 ± 26.34	0.000*	0.70	0.80
LT Raw (reps)	0.81 ± 1.36	1.57 ± 2.13	0.032*	0.41	0.37
2MR Raw (s)	1118.57 ± 196.67	1098.86 ± 184.98	0.159	0.10	0.07
DL Points	59.14 ± 7.03	56.05 ± 18.75	0.130	0.19	0.12
SPT Points	39.43 ± 27.26	53.76 ± 14.79	0.034*	0.60	0.68
HRP Points	70.14 ± 5.66	72.09 ± 6.41	0.094	0.32	0.25
SDC Points	48.95 ± 28.34	67.24 ± 13.6	0.000*	0.75	0.85
LT Points	18.43 ± 29.88	36.62 ± 32.72	0.011*	0.58	0.64
2MR Points	60.38 ± 32.13	61.38 ± 32.31	0.318	0.03	0.05
Total Points	296.48 ± 96.83	347.14 ± 80.36	0.000*	0.56	0.62
Qualify Level	0.19 ± 0.40	0.38 ± 0.50	0.046*	0.42	0.39

Data are presented as mean ± SD
* = denotes statistical difference (p ≤ 0.05) between pre- and post- semester scores

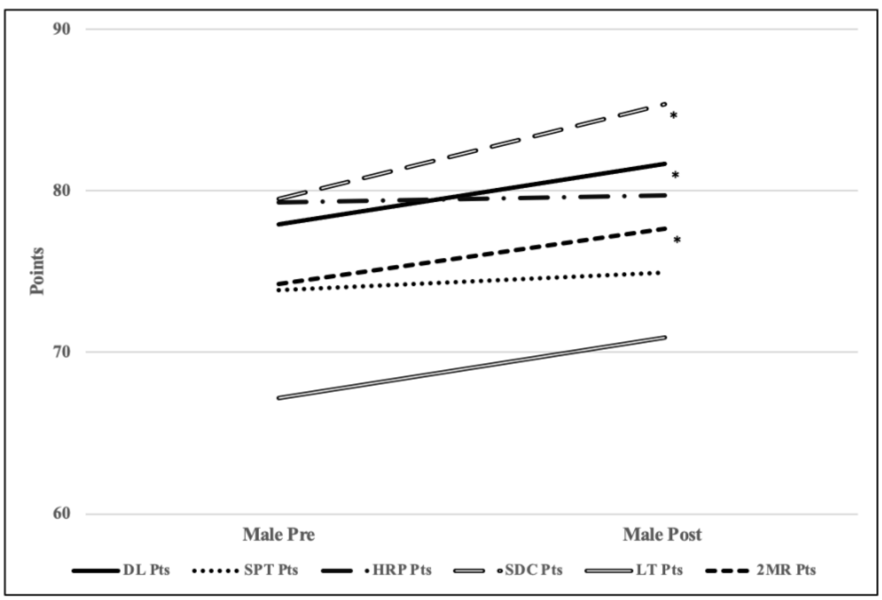

Figure 8: Differences in males' individual event scores from pre- to post-semester

* = denotes statistical difference (p ≤ 0.05) between pre- and post- semester scores

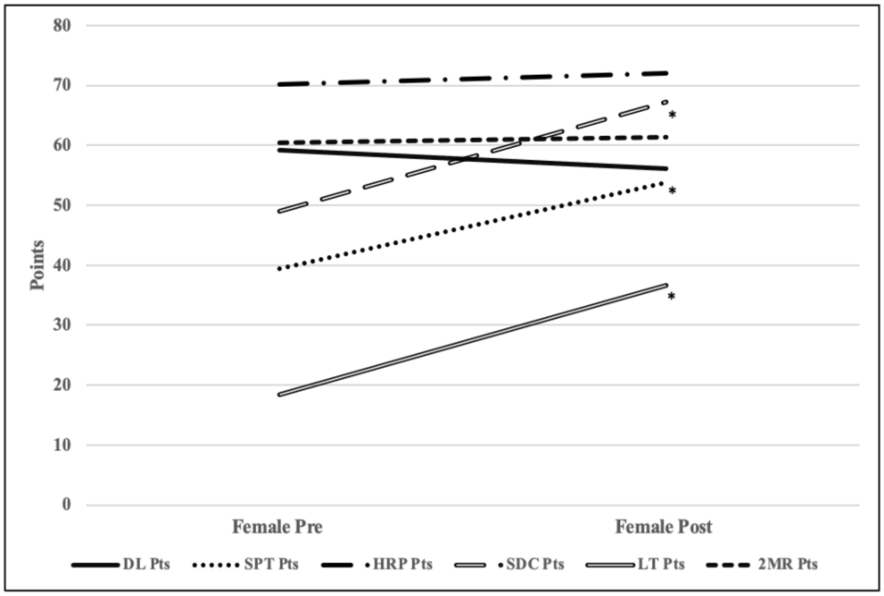

Figure 9: Differences in females' individual event scores from pre- to post- semester

*** =** denotes significant difference pre- to post-test ($p \leq 0.05$)

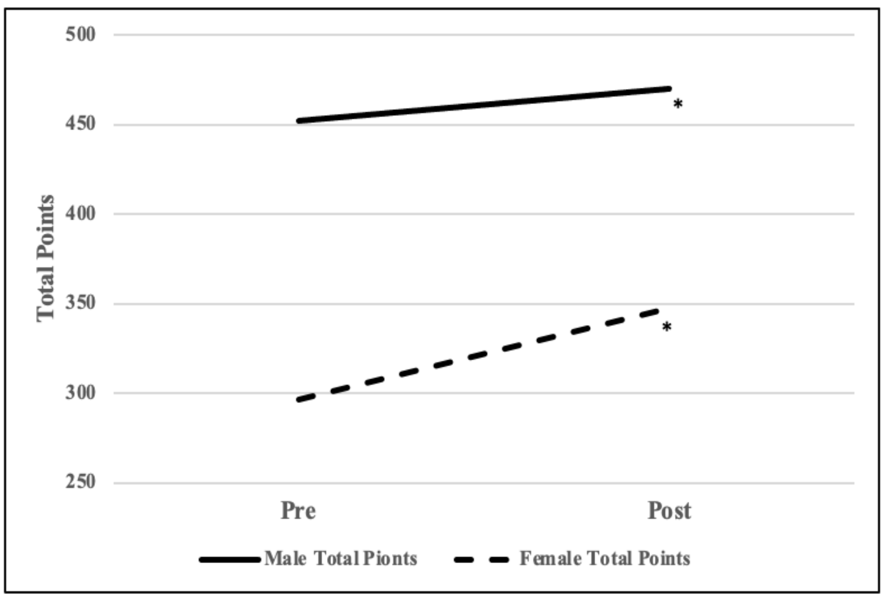

Figure 10: Differences between total ACFT scores from pre- to post- semester for each sex

* = denotes statistical difference between pre- and post- semester scores ($p \leq 0.05$)

Multinomial Logistic Regression

Cadets' military science class, height, and mass were reported for 424 males and 155 females. This group consisted of 194 MS-4, 188 MS-3, 112 MS-2, and 84 MS-1 cadets.

Three-Repetition Maximum Deadlift

Multinomial logistic stepwise regression was used to create a model that represented a significant improvement in fit relative to the null model for three-repetition maximum deadlift category qualified for (LR χ^2 (18) = 456.358, $p < 0.001$). Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 31.6% improvement in fit relative to the null model. The overall classification accuracy of the model was 80% with 0% correct failure, 65.8% correct moderate, 0% correct significant, and 94.5% correct heavy predictions.

Mass

Among all cadets, those with greater mass were less likely to "FAIL" (B = -0.309, Exp(B) = 0.734, $p < 0.001$), to score in the "MODERATE" category (B = -0.080, Exp(B) = 0.923, $p < 0.001$, or to score in "SIGNIFICANT" category (B = -0.065, Exp(B) = 0.937, p = 0.002) when compared to cadets with less mass (Table 35).

Sex

When considering MDL performance, in light of sex, females were significantly more likely to either score in the "MODERATE" (B = 2.776, Exp(B) = 16.059, $p < 0.001$) and "SIGNIFICANT" (B = 2.672, Exp(B) = 14.473, $p < 0.001$) categories compared to their male counterparts.

Table 35: Significant stepwise multinomial regression results for MDL

DL Level		B	SE	Sig.	Exp (B)
N = 579					
0	Intercept	14.724	4.533	0.001	
	Mass	-0.309	0.075	0.000	0.734
1	Intercept	4.037	1.031	0.000	
	Sex	2.776	0.297	0.000	16.059
	Mass	-0.080	0.014	0.000	0.923
2	Intercept	1.585	1.023	0.312	
	Sex	2.672	0.453	0.000	14.473
	Mass	-0.065	0.021	0.002	0.937
In reference to "3"					

Standing Power Throw

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for standing power throw category qualified for (LR $\chi^2(18) = 456.387$, $p < 0.001$). Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 35.7% improvement in fit relative to the null model. The overall classification accuracy of the model was 72.4% with 7.1% correct failure, 77.2% correct moderate, 0% correct significant, and 96.6% correct heavy predictions.

Sex

In SPT performance, females were far more likely to "FAIL" (B = 4.781, Exp(B) = 119.250, p < 0.001), score in the "MODERATE" (B = 4.067, Exp(B) = 58.364, p < 0.001) , or score in the "SIGNIFICANT" (B = 2.477, Exp(B) = 11.909, p < 0.001) category, compared to males (Table 36).

Height

Cadets with greater height are less likely to "FAIL" (B = 09.977, Exp(B) = 4.644 e-5, p = 0.015) or score in the "MODERATE" (B = -7.120, Exp(B) = 0.001, p = 0.006) category than shorter cadets.

Mass

Cadets with greater mass are less likely to “FAIL” (B = -0.176, Exp(B) = 0.838, $p < 0.001$), or score in the “MODERATE” (B = -0.102, Exp(B) = 0.903, p = 0.000) or “SIGNIFICANT” (B = -0.046, Exp(B) = 0.955, p = 0.003) category than cadets with less mass.

MS-Class

MS-4 cadets are less likely to “FAIL” (B = -3.291, Exp(B) = 0.037, $p < 0.001$) or score in the “MODERATE” (B = -1.945, Exp(B) = 0.143 p = 0.000) or “SIGNIFICANT” (B = -0.932, Exp(B) = 0.394, p = 0.015) category of the SPT. Similarly, MS-3 cadets are less likely to “FAIL” (B = -3.347, Exp(B) = 0.035, p = 0.012) or score in the “MODERATE” (B = -2.517, Exp(B) = 0.081, $p < 0.001$) or “SIGNIFICANT” (B = -1.445, Exp(B) = 0.236, p = 0.001) category. MS-2 cadets are less likely to “FAIL” the SPT than MS-1 cadets (B = -2.025, Exp(B) = 0.132, p = 0.012).

Table 36: Significant stepwise multinomial regression results for SPT

SPT Level		B	SE	Sig.	Exp (B)
N = 579					
0	Intercept	27.065	6.363		
	Sex	4.781	0.874	0.000	119.250
	Height	-9.977	4.087	0.015	4.644 e-5
	Mass	-0.176	0.037	0.000	0.838
	MS-4	-3.291	0.805	0.000	0.037
	MS-3	-3.347	0.759	0.000	0.035
	MS-2	-2.025	0.802	0.012	0.132
1	Intercept	19.219	4.270		
	Sex	4.067	0.511	0.000	58.364
	Height	-7.120	2.610	0.006	0.001
	Mass	-0.102	0.021	0.000	0.903
	MS-4	-1.945	0.473	0.000	0.143
	MS-3	-2.517	0.510	0.000	0.081
2	Intercept	7.776	3.449		
	Sex	2.477	0.498	0.000	11.909
	Mass	-0.046	0.015	0.003	0.955
	MS-4	-0.932	0.384	0.015	0.394
	MS-3	-1.445	0.418	0.001	0.236
In reference to “3”					

Hand Release Push-ups

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for hand release push-ups category qualified for (LR $\chi^2(12) = 94.170$, $p < 0.001$. Based on McFadden’s Pseudo R-Square, the full model containing our predictors resulted in a 12.3% improvement in fit relative to the null model. The overall classification accuracy of the model was 78.7% with 2% correct moderate, 4% correct significant, and 99.3% correct heavy predictions. Only one cadet failed the HRP. Thus, this data point and the “FAIL” category were removed from the model.

Sex

Female cadets were more likely to score in the "MODERATE" (B = 1.800, Exp(B) = 6.048, $p < 0.001$) or "SIGNIFICANT" (B = 1.799, Exp(B) = 6.042, $p < 0.001$) category for HRP than males (Table 37).

MS-Class

MS-4 cadets were less likely to score in the "SIGNIFICANT" (B = -0.867, Exp(B) = 0.420, p = 0.026) category of HRP. Similarly, MS-3 cadets were less likely to score in the "MODERATE" (B = -1.155, Exp (B) = 0.315, p = 0.040) or "SIGNIFICANT" (B = -1.335, Exp(B) = 0.263, p = 0.001) category.

Table 37: Significant stepwise multinomial regression results for HRP

HRP Level		B	SE	Sig.	Exp (B)
N = 578					
1	Intercept	-0.651	4.108	0.874	
	Sex	1.800	0.452	0.000	6.048
	MS-3	-1.155	0.562	0.040	0.315
2	Intercept	-7.069	3.576		
	Sex	1.799	0.375	0.000	6.042
	MS-4	-0.867	0.388	0.026	0.420
	MS-3	-1.335	0.400	0.001	0.263
In reference to "3"					

Sprint-Drag-Carry

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for sprint-drag-carry category qualified for (LR $\chi^2(15) = 349.045$, $p < 0.001$. Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 36.8% improvement in fit relative to the null model. The overall classification accuracy of the model was 79.1% with 0% correct failure, 41.8% correct moderate, 38.6% correct significant, and 94.6% correct heavy predictions.

Sex

Female cadets were more likely to score in the “MODERATE” ($B = 3.442$, $Exp(B) = 31.251$, $p < 0.001$) and “SIGNIFICANT” ($B = 2.751$, $Exp(B) = 15.651$, $p < 0.001$) category of the SDC than male cadets (Table 38).

Height

Cadets with greater height are less likely to “FAIL” ($B = -18.904$, $Exp(B) = 6.166$ e-9, $p = 0.000$) or score in the “MODERATE” ($B = -14.240$, $Exp(B) = 6.54$ e-7, $p < 0.001$) or “SIGNIFICANT” ($B = -6.992$, $Exp(B) = 0.001$, $p = 0.001$) categories of the SDC.

MS-Class

MS-4 cadets are less likely to “FAIL” ($B = -3.498$, $Exp(B) = 0.030$, $p = 0.001$) the SDC or score in the “MODERATE” ($B= -3.431$, $Exp(B) = 0.032$, $p < 0.001$) or “SIGNIFICANT” ($B = -1.832$, $Exp(B) = 0.160$, $p < 0.001$) categories. Likewise, MS-3 cadets are less likely to “FAIL” ($B = -3.441$, $Exp(B) = 0.032$, $p =0.001$) or score in the “MODERATE” ($B = -2.295$, $Exp(B) = 0.101$, $p < 0.001$) or “SIGNIFICANT” ($B = -1.999$, $Exp(B) = 0.135$, $p < 0.001$) categories of the SDC. MS-2 cadets are less likely to score in the “MODERATE” ($B = -1.534$, $Exp(B) = 0.216$, $p = 0.011$) or “SIGNIFICANT” ($B = -1.306$, $Exp(B) = 0.271$, $p = 0.005$) SDC categories.

Table 38: Significant stepwise multinomial regression results for SDC

SDC Level		B	SE	Sig.	Exp (B)
N = 579					
0	Intercept	12.104	7.496	0.106	
	Height	-18.904	4.638	0.000	6.166 e-9
	MS-4	-3.498	1.087	0.001	0.030
	MS-3	-3.441	1.077	0.001	0.032
1	Intercept	22.461	5.037	0.000	
	Sex	3.442	0.555	0.000	31.251
	Height	-14.240	2.898	0.000	6.54e-7
	MS-4	-3.431	0.669	0.000	0.032
	MS-3	-2.295	0.581	0.000	0.101
	MS-2	-1.534	0.606	0.011	0.216
2	Intercept	11.026	3.821	0.004	
	Sex	2.751	0.373	0.000	15.651
	Height	-6.992	2.156	0.001	0.001
	MS-4	-1.832	0.436	0.000	0.160
	MS-3	-1.999	0.450	0.000	0.135
	MS-2	-1.306	0.465	0.005	0.271
In reference to "3"					

Leg Tuck

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for leg tuck category qualified for (LR $\chi^2(18) = 241.900$, $p < 0.000$). Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 20.6% improvement in fit relative to the null model. The overall classification accuracy of the model was 69.6% with 22.6% correct failure, 48.3% correct moderate, 0% correct significant, and 90.5% correct heavy predictions.

Sex

Female cadets were more likely to "FAIL" (B = 4.471, Exp(B) = 87.465, $p < 0.001$) or score in the "MODERATE" (B = 3.248, Exp(B) = 25.737, $p < 0.001$) category of the LT (Table 39).

Height

Cadets with greater height were less likely to "FAIL" (B = -8.063, Exp(B) = 0.000, p = 0.006) the LT.

Mass

Cadets with greater mass were more likely to "FAIL" (B = 0.099, Exp(B) = 1.104, p < 0.001) or score in the "MODERATE" (B = 0.044, Exp(B) = 1.045, p = 0.001) category of the LT.

MS-Class

MS-4 cadets were less likely to "FAIL" (B = -2.147, Exp(B) = 0.117, p < 0.001) or score in the "MODERATE" (B = -0.897, Exp(B) = 0.408, p = 0.016) LT category. Likewise, MS-3 cadets were less likely to "FAIL" (B = -1.781, Exp(B) = 0.168, p = 0.001) or score in the "MODERATE" (B = -1.231, Exp(B) = 0.292, p = 0.001) LT category.

Table 39: Significant stepwise multinomial regression results for LT

LT Level		**B**	**SE**	**Sig.**	**Exp (B)**
N = 579					
0	Intercept	4.009	4.447	0.367	
	Sex	4.471	0.578	0.000	87.465
	Height	-8.063	2.927	0.006	0.000
	Mass	0.099	0.020	0.000	1.104
	MS-4	-2.147	0.580	0.000	0.117
	MS-3	-1.781	0.537	0.001	0.168
1	Intercept	-1.200	3.145	0.703	
	Sex	3.248	0.379	0.000	25.737
	Mass	0.044	0.014	0.001	1.045
	MS-4	-0.897	0.374	0.016	0.408
	MS-3	-1.231	0.385	0.001	0.292
2	Intercept	-1.138	4.356	0.794	
In reference to "3"					

Two-Mile Run

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for two-mile run category qualified for (LR $\chi^2(15) = 150.653$, $p < 0.001$). Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 15.0% improvement in fit relative to the null model. The overall classification accuracy of the model was 73.4% with 14.3% correct failure, 12.9% correct moderate, 0% correct significant, and 97.2% correct heavy predictions.

Sex

Female cadets were more likely to "FAIL" (B = 2.769, Exp(B) = 15.939, $p < 0.001$) and score in the "MODERATE" (B = 3.181, Exp(B) = 24.080, $p < 0.001$) and "SIGNIFICANT" (B = 1.566, Exp(B) = 4.788, $p < 0.001$) categories of the 2MR (Table 40).

Mass

Those with greater mass were more likely to "FAIL" (B = 0.044, Exp(B) = 1.045, p = 0.007) and score in the "MODERATE" (B = 0.056, Exp(B) = 1.058, $p < 0.001$) category of the 2MR.

MS-Class

MS-4 cadets were less likely to "FAIL" (B = -2.417, Exp(B) = 0.089, $p < 0.001$) or score in the "MODERATE" (B = -1.194, Exp(B) = 0.303, p = 0.012) category of the 2MR. Likewise, MS-3 cadets were less likely to "FAIL" (B = -2.682, Exp(B) = 0.068, $p < 0.001$) or score in the "MODERATE" (B = -1.159, Exp(B) = 0.314, p = 0.011) category of the 2MR. MS-2 cadets were less likely to "FAIL" (B = -0.997, Exp(B) = 0.369, p = 0.042) and more likely to score in the "SIGNIFICANT" (B = 1.579, Exp(B) = 4.849, p = 0.044) category of the 2MR.

Table 40: Significant stepwise multinomial regression results for 2MR

2MR Level		B	SE	Sig.	Exp (B)
N = 579					
0	Intercept	-5.208	1.370	0.000	
	Sex	2.769	0.490	0.000	15.939
	Mass	0.044	0.017	0.007	1.045
	MS-4	-2.417	0.564	0.000	0.089
	MS-3	-2.682	0.573	0.000	0.068
	MS-2	-0.997	0.490	0.042	0.369
1	Intercept	-6.569	1.158	0.000	
	Sex	3.181	0.403	0.000	24.080
	Mass	0.056	0.013	0.000	1.058
	MS-4	-1.194	0.473	0.012	0.303
	MS-3	-1.159	0.454	0.011	0.314
2	Intercept	-4.250	1.311	0.001	
	Sex	1.566	0.381	0.000	4.788
	MS-2	1.579	0.783	0.044	4.849
In reference to "3"					

Overall Qualification Category

Multinomial logistic stepwise regression was able to create a model that represented a significant improvement in fit relative to the null model for the overall ACFT qualified category for (LR $\chi^2(18) = 345.850$, $p < 0.001$). Based on McFadden's Pseudo R-Square, the full model containing our predictors resulted in a 22.5% improvement in fit relative to the null model. The overall classification accuracy of the model was 54.9% with 40.7% correct failure, 47.7% correct moderate, 0% correct significant, and 93.8% correct heavy predictions.

Sex

Female cadets were more likely to "FAIL" (B = 6.180, Exp(B) = 483.000, $p < 0.001$) and score in the "MODERATE" (B = 4.787, Exp(B) = 119.882, $p < 0.001$) and "SIGNIFICANT" (B = 2.719, Exp(B) = 15.169, $p = 0.012$) categories of the ACFT (Table 41).

Height

Cadets with greater height were less likely to “FAIL” (B = -10.885, Exp(B) = 1.873 e-5, $p < 0.001$) or score in the “MODERATE” (B = -4.653, Exp(B) = 0.010, p = 0.022) category.

Mass

Cadets with greater mass are more likely to “FAIL” (B = 0.041, Exp(B) = 1.042, $p < 0.001$) the ACFT than score in the “HEAVY” category.

MS-Class

MS-4 cadets were less likely to “FAIL” (B = -2.863, Exp(B) = 0.057, p = 0.000) or score in the “MODERATE” (B = -1.025, Exp(B) = 0.359, p = 0.005) category of the ACFT. Similarly, MS-3 cadets were less likely to “FAIL” (B = -2.874, Exp(B) = 0.056, $p < 0.001$) or score in the “MODERATE” (B = -1.659, Exp(B) = 0.190, $p < 0.001$) category.

Table 41: Significant stepwise multinomial regression results for overall qualification

Overall		**B**	**SE**	**Sig.**	**Exp (B)**
N = 579					
0	Intercept	15.509	4.284	0.000	
	Sex	6.180	1.082	0.000	483.000
	Height	-10.885	2.755	0.000	1.873e-5
	Mass	0.041	0.018	0.019	1.042
	MS-4	-2.863	0.521	0.000	0.057
	MS-3	-2.874	0.497	0.000	0.056
1	Intercept	9.443	3.243	0.004	
	Sex	4.787	1.039	0.000	119.882
	Height	-4.653	2.037	0.022	0.010
	MS-4	-1.025	0.365	0.005	0.359
	MS-3	-1.659	0.384	0.000	0.190
2	Intercept	4.846	3.247	0.136	
	Sex	2.719	1.086	0.012	15.169
In reference to “3”					

Discussion

The purpose of this study was to identify whether sex-based and experience-based differences exist in ACFT scores within a cohort that regularly conducts military-based physical training and if present, where those differences occurred. Additionally, the purpose of this study was to identify possible relationships between anthropometric variables and ACFT outcomes. Upon receiving pre- and post- semester ACFT scores for two universities, an additional purpose of this study was to analyze whether cadets' ACFT performances were affected by a semester of physical training. This investigation was a cross-sectional retrospective data analysis that examined the ACFT scores of cadets from Cadet Command's 5th BDE. The main findings were: 1) Males performed significantly better than females on the ACFT. 2) Upperclassmen cadets performed significantly better on the ACFT than underclassmen cadets. 3) Height, mass, and BMI were significantly, and in some cases, moderately correlated with ACFT performance. 4) Cadets performed significantly better in some, but not all, ACFT events after a semester of training. 5) Odds ratios determined sex to be the most likely factor in determining ACFT performance.

Age and Anthropometric Variables

Sex Differences

Males and females were statistically significantly different in age. However, the observed effect size was minimal at 0.14. In the context of this study, age was not used as a potential predictor of military experience because an MS-1 (freshman) cadet could be older than an MS-4 (senior) cadet. Therefore, age differences, though significant, are not of great importance in the current study.

Consistent with previous research, males and females had significantly different heights, with males being taller on average. The observed effect size of this difference, 1.93, was the largest amongst the anthropometric variables examined. Foundational research on the physiological differences between sexes identified height disparities stemming from pre-pubertal and pubertal growth characteristics (Marshall & Tanner, 1969, 1970). The greater height that males possess may provide an advantage in performance of physical fitness events. This concept will be examined at length later in the discussion.

The variables in this study were limited to easily observable anthropometrics. Thus, FM and FFM were not obtained from the cadets. Body mass and BMI are still currently used by the Army. Therefore, these variables provide the researchers with the same information that university ROTC programs and Army units commonly have access to. As differences in height would lead one to expect, males and females also had significantly different body masses with males being approximately 14kg heavier. While exact composition of the body mass cannot be determined, the overall greater average mass of males may be attributed to more muscle mass than the female cadets. Likewise, the lower body mass observed in the females may be attributed to, but cannot be exactly determined, less muscle mass than the male cadets.

Male cadets, on average, had a BMI that was 1.0 kg/m^2 greater than that of female cadets. This statistically significant difference resulted in a small effect size of 0.31. Regardless, this result is congruent with some literature and incongruent with other literature. Similar to the current study, male enlisted soldiers and officers had greater BMI than female enlisted soldiers and officers (Canino et al., 2018). However, in an analysis of 101st Airborne soldiers, females were found to have greater BMI than males (Allison et al., 2015). A recent investigation using a similar population as the current study, found no difference between the sexes for BMI in ROTC cadets (Draicchio et al., 2020). The current findings are important additions to the literature surrounding BMI differences and if they consistently exist between male and female cadets and soldiers.

MS-Class Differences

To our knowledge, there is no published research related to differences between military science classes based on anthropometric variables or fitness performance. Therefore, the results of this study are unique and may provide valuable insight into MS-class similarities and differences. As expected, MS-classes were all significantly different in average age with MS-1 cadets being the youngest, MS-4 being the oldest, and MS-2 and MS-3 falling in line between. As previously mentioned, differences in age do not necessarily correspond with differences in military experience. Therefore, no further analysis was conducted analyzing age and its effects.

MS-classes were not significantly different in height, despite upperclassmen cadets being significantly older than underclassmen cadets. This may be due to males and females reaching their peak height during or before their first year of ROTC participation. Body mass was only significantly different between MS-2 and MS-4 classes. The MS-4 cadets were, on average, 3kg heavier. MS-1 cadets, though not statistically different, had greater body mass than MS-2 cadets, who had the least body mass out of all classes. This decrease in body mass from MS-1 to MS-2 may be due to physical training that serves to decrease the body fat of freshmen. Consequently, this same physical training may also increase the muscle mass of cadets the longer they are in ROTC, as MS-3 cadets' mass was greater than MS-2, and MS-4 was greater than MS-3. As previously mentioned with sex differences, differences in body mass may be due to FM or FFM. Therefore, this study lacks the ability to determine if the statistically greater mass in the seniors is advantageous. Not surprisingly, BMI coincided with the body mass as MS-4 and MS-2 cadets were also the only class comparison that was significantly different.

Three-Repetition Maximum Deadlift

The three-repetition maximum deadlift (MDL) was designed by the Army to test lower body strength while also reflecting a movement similar to lifting and carrying a casualty on a litter and moving one's own equipment (Headquarters Center for Initial Military Training, 2019). Isometric peak deadlift force has been connected with successful performance on dummy casualty evacuations and lower body strength has been significantly correlated with occupational tasks such as box lifts and load carriage (Hydren et al., 2017; Poser et al., 2019; Szivak et al., 2014). Lower body strength, therefore, is essential to successful completion of job duties for military personnel.

Sex Differences

Congruent with previous literature, the males performed significantly better on the MDL than the females. When comparing raw weight lifted, males lifted on average almost 100 pounds more than females, 248.00lbs ± 59.28lbs vs. 155.02lbs ± 43.14lbs. This resulted in an observed effect size of 1.79, higher than any other individual event comparison in this study. These results are similar to an experiment analyzing cadets' performance on the OPAT that found males had significantly higher average one-repetition maximum deadlift raw values compared to females, 92.82kg ± 13.56kg vs. 76.31kg ± 16.84kg, with an observed effect size of 1.72 (Draicchio et al., 2020). These results are also comparable to that of a meta-analysis that also found a very large effect size for lower body strength differences in males and females at 1.62 (Courtright et al., 2013). When compared on the Army's 100-point scale, males still performed significantly better than females by 20 points, with a large observed effect size of 0.92. As this scaled score is what ultimately goes on record for the cadet/soldier, differences in these scaled scores are equally, if not more, impactful on job qualifications and promotion evaluations.

In this study, the average MDL raw and scaled values for females, 155.02lbs ± 43.14lbs and 60.26 ± 14.5, respectively, would only reach the lowest qualification category, "MODERATE", whereas males average MDL raw value qualifies them for "HEAVY". Only 13.2% of females were able to meet the minimum MDL requirement for "HEAVY" while, in contrast, 82.7% of males met the minimum. On the OPAT, 50% of females scored "HEAVY" on the deadlift, while 70% of males were able to reach the category (Draicchio et al., 2020). The difference is even greater, and more comparable to the current study, in basic trainees, where only 15% of females were able to reach "HEAVY" while 80% of males were (J. Pierce et al., 2018). Therefore, while they are on different scales, the ACFT requiring three-repetitions at 200lbs and the OPAT requiring one-repetition at 220lbs, males' and females' ability to reach the heaviest category in the deadlift is vastly different. Reaching this "HEAVY" category is a requirement for cadets and soldiers to serve in the more physically rigorous combat arms branches. Therefore, despite all MOS's being recently opened to female soldiers, a majority of females would not qualify for "HEAVY" jobs based on their deadlift performance alone.

Even when deadlift performance is expressed relative to the cadets' body mass, ([deadlift (kgs) / body mass (kgs)] * 100%), males were able to lift a significantly greater percent of their body mass than females were. The difference between sexes in deadlift performance relative to body mass is compatible with previous literature that concluded male and female soldiers are significantly different in knee flexion and extension strength when expressed relatively (Allison et al., 2015). Additionally, previous research on strength trained individuals has also observed significant differences in conventional deadlift one-repetition maximum when expressed relative to body mass, with males able to lift on average 230% and females 160% of their body mass (Jones et al., 2016). Even if soldiers' occupational workloads could be adjusted to be relative to their body mass, male cadets' ability to deadlift 143% and females' 115% of their body mass would most likely result in differing levels of fatigue. One's total body weight would likely be more fatiguing for females than it would be for males, even if the absolute loads were different, because it is a higher percent of a female's maximal ability.

Military Science Class Differences

Military science classes were not significantly different in raw, scaled, or relative MDL performance. The null finding between cadets in this event was unanticipated. Considering, the OPAT has been used by the Army since 2017, upperclassmen cadets have had more exposure to a deadlift type fitness test event and one would assume, upperclassmen would thus perform better. However, before the establishment of the ACFT, there may not have been formal training for the OPAT, and thus the deadlift, as it was for a one-time test and not a semi-annual fitness test for record. If military experience was the sole regulator of MDL performance, MS-4 cadets would have the greatest average performances. However, this was not the case. The MS-3 cadets had the best deadlift performances out of any of the classes. Though not statistically significant, this may be due to the fact that cadets' junior (MS-3) year fitness test scores are factored into their overall order of merit to receive their desired MOS and component. This may point to a motivational component as part of the explanation for this finding.

Anthropometric Correlations

The MDL raw and scaled scores were significantly correlated with all anthropometric variables. Moderate correlations were seen between MDL raw with height and mass as well as MDL points and mass. BMI was significantly correlated with deadlift outcomes but only displayed moderate correlations. A previous study using basic trainees observed a positive trend between BMI and deadlift outcomes on the OPAT (J. Pierce et al., 2018). Despite only being a moderate correlation, we observed this trend as well. With males having significantly greater BMI's than females it is no surprise that males performed better on the deadlift event. More notably, increased mass presented moderate correlations with deadlift performance. Body mass was significantly different between males and females and thus supports the finding of males performing better on the MDL.

Previous research has shown that when strength is corrected for differences in muscle mass, i.e. relative to FFM, sex differences disappear (Miller et al., 1993; Sharp, 1993). While FFM data was not available in this study, expressing MDL relative to body mass was possible. Even when expressing relative to body mass, there were significant differences between males and females for this test of lower body strength. Therefore, we can draw the conclusion that composition of the males' body mass was different than that of the females. Males, on average, have more FFM and less FM than females. Females, on average, have less FFM and more FM than males (Wells & Plowman, 1983). Increased muscle mass often equates to greater strength and thus, despite adjusting relative to total body mass, the composition of that body mass plays an integral role in deadlift performance. When Spearman's Rho (ρ) was calculated for the individual sexes, males exhibited larger ρ values for mass vs. DL raw and mass vs. DL points than females; $\rho = 0.419$ & 0.411 vs. $\rho = 0.273$ & 0.282. Though all statistically significant, the effects of mass on deadlift performance may be different for males and females. This could be due to females' increased mass being due to increased FM, rather than FFM, and males' increased mass being a result of greater FFM.

Pre- & Post-Semester Outcomes

Male cadets' raw and scaled MDL scores significantly changed from pre- to post-semester but only exhibited moderate effect sizes, 0.3 and 0.29, respectively. Female cadets did not significantly change over the course of a semester with post-semester averages declining from the pre-semester test. It is plausible that female cadets did not respond to the physical training regimen in the same manner male cadets did. The training that was conducted may not have been adequate to induce the physiological changes necessary, such as gaining FFM, for increasing deadlift performance. Previous research has shown males and females progress similarly in muscle size and strength over the course of a training protocol when both sexes trained at similar intensities (Abe et al., 2000; Staron et al., 1994). Therefore, it is possible that female cadets were not performing deadlift and/or lower body strength training at the same relative intensity as males and thus, did not elicit similar strength gains. ROTC programs are expected to adhere to published Army physical training protocol. However, physical training programming is planned and conducted by the cadets and is not conducted in a controlled, laboratory environment. Thus, divergences from Army suggested methods and proper strength and conditioning practices may occur.

Multinomial Logistic Regression Odds Ratios

Analysis of the multinomial logistic regression results shows that females had significantly greater odds of scoring in the "MODERATE" or "SIGNIFICANT" category than the "HEAVY" category. It can be concluded then that males are significantly more likely to score in the "HEAVY" category than females. This may be due to a variety of factors. However, mass was the only other significant variable in the regression equation. Those with greater mass are less likely to "FAIL" or score in any other category than the "HEAVY" category. The odds ratios for mass are pointedly less than that of sex (ex. 16.059 vs. 0.923) and therefore may not have as much of an effect on MDL outcome as sex does. Regardless, mass was the only significant predictor for the failure group. Thus, those with less mass are more likely to "FAIL" than those with greater mass. Fourteen females (about 5%) failed the MDL, making it the second least failed event by females behind the hand-release push-ups. Despite the relatively low number of failures, females' odds ratios point to a greater likelihood of not scoring in the "HEAVY" category, thus limiting the MOS they qualify for at the moment.

Standing Power Throw

The SPT was designed by the Army to test explosive power while also reflecting a movement similar to helping a soldier scale a wall, throwing a hand grenade, jumping across a ditch, and moving equipment over an obstacle (Headquarters Center for Initial Military Training, 2019). The backwards, overhead SPT has been evaluated as a valid and reliable test to assess explosive power (Stockbrugger & Haennel, 2001). The ACFT's SPT is used to assess full body explosive power while the OPAT's events of standing long jump and seated medicine ball power throw evaluate lower body power and upper body power separately. As the only test of explosive power in the ACFT, performance in the SPT may provide key insight into a soldier's coordination and ability to perform specific occupational tasks as outlined by the Army.

Sex Differences

Congruent with previous literature, the males performed significantly better on the SPT than the females. When comparing raw distance thrown, males threw, on average, almost 4m further than females (9.21m ± 3.47 vs. 5.55m ± 1.28, respectively). This resulted in a large observed effect size of 1.40. Our findings are similar to the results of Draicchio et al. (2020), who analyzed cadets' performance on the OPAT and found males performed significantly better on the explosive power tests of standing long jump and seated medicine ball power throw when compared to females, with observed effect sizes of 1.24 and 2.28, respectively (Draicchio et al., 2020). Similarly, significant differences in these OPAT events were found between male and female basic trainees, where males jumped 25% and threw 34% further than females (J. Pierce et al., 2018). These results are also comparable to that of a meta-analysis that calculated effect sizes for power differences in males and females to be 1.11 (Courtright et al., 2013). Even when compared on the Army's 0-100 scale, males still performed significantly better than females by 20 points, with a large observed effect size of 1.51. Since males and females often differ in their amounts of FFM and muscular strength, force production and explosive power output may also be different between the two sexes.

In this study, the SPT raw values for females (5.55m ± 1.28m) would, on average, only reach the lowest qualification category, "MODERATE", whereas males average SPT raw value qualifies them for "HEAVY". Female cadets' average SPT points (57.61 ± 15.01) would not reach the minimum qualification level requiring 60 points and would result in failure. The SPT was the second most failed event by female cadets, with 16.4% of females failing, as compared to < 1% of males. Therefore, performance in this event may be a limiting factor in a cadet passing the ACFT at the minimum level. On the OPAT, 83% and 56% of female ROTC cadets scored "HEAVY" on the standing long jump and seated power throw, respectively, while 99% and 97% of male cadets were able to reach this category (Draicchio et al., 2020). However, in the current study, only 4.7% of females scored "HEAVY" in the SPT compared to 73.4% of male cadets. The SPT had the least number of females and males reach the "HEAVY" standard out of all events. Therefore, the standards required for a cadet or soldier to reach the "HEAVY" category in the ACFT may be more difficult than in the OPAT for explosive power tests and may be a limiting factor in qualifying for the more physically demanding MOSs.

Military Science Class Differences

Upperclassmen cadets performed significantly better on the SPT than their underclassmen counter parts. MS-3 and MS-4 cadets threw the medicine ball, on average, 1m further than MS-1 cadets and just over 0.5m further than MS-2 cadets. Unlike the deadlift, where MS-3 cadets had the highest averages, the MS-4 cadets claimed the highest average raw and scaled SPT scores. Significant differences between the MS-classes may be due to exposure to explosive power training. Training for the MDL event requires weightlifting equipment that is expensive and cumbersome to transport and set up. Therefore, cadets may not have had as much experience in MDL training. In contrast, training explosive power can be done with body weight exercises and plyometrics. Because the Army has traditionally trained using calisthenics and other body weight training mechanisms, MS-4 cadets may have performed the best because they have been training explosive power the longest despite the ACFT being a new test. Additionally, the OPAT includes lower and upper body explosive power tests. Upperclassmen cadets with more experience training for and/or taking the OPAT will have more experience with explosive power type tests.

Anthropometric Correlations

SPT performance outcomes were significantly correlated with all anthropometric variables. The correlations between SPT and height and mass were stronger than any other correlations observed in this study. A taller cadet will have a greater release height of the medicine ball which may contribute to an increased throw distance per the tenants of projectile motion. Likewise, increased mass was correlated with increased SPT raw and scaled scores.

Because this increase in mass is correlated with better performance, we can postulate that the greater mass is likely due to more FFM capable of strength and power output. Therefore, it is no surprise that males, who on average are taller and heavier than females, perform significantly better on this event. When analyzing the sexes separately, SPT performance still had significant correlations with height and mass but were only moderate correlations compared to combined analysis. However, height and mass were not more strongly correlated with SPT for only the males, unlike the deadlift. Therefore, height and mass may have an equal effect on SPT performance regardless of sex.

A previous study has associated higher BMI with decreased performance, or a negative correlation, by basic trainees in the standing long jump of the OPAT (J. Pierce et al., 2018). In contrast, the current study observed significant, moderate positive correlations between BMI and SPT performance. This may be due to the standing long jump requiring soldiers to propel one's entire body mass forward while the SPT consists of propelling only a 10lb medicine ball. In this case, increased body mass, leading to increased BMI, would not have as much of a detriment on performance and may help increase momentum behind the medicine ball.

Similar to the results of this study, in basic trainees, increased BMI was positively correlated with seated medicine ball throw performance (J. Pierce et al., 2018). Higher BMI, therefore, may not have as much of a negative effect, and may serve as advantageous, in fitness test events utilizing a medicine ball for resistance instead of one's own body mass. Additionally, ROTC cadets' increased BMI may be due to increased FFM when compared to basic trainees as they have had months and/or years more physical training experience.

Pre- & Post-Semester Outcomes

While male cadets did not perform significantly better on the SPT at the end of the semester, female cadets significantly increased their SPT scaled score by almost 15 points. Increases in SPT performance over the course of a semester could be due to a couple of factors such as: increased muscle mass, increased BMI, and/or better throwing form leading to increased release height and optimal throw angle. The sub-population in which pre- and post-semester ACFT data was gathered did not report anthropometric variables. Therefore, a connection between body mass or BMI changes and SPT performance cannot be made. However, it is likely that over the course of a semester, female cadets' throwing form improved leading to further throw distances. SPT raw distance scores were not significantly different from pre- to post-semester for female cadets, despite scaled scores significantly changing. This leads to an analysis of the 0-100 scoring scale developed by the Army. The minimum throw distance for the "MODERATE" category is 4.5m for 60 points. A 0.4m increase in distance from here would only results in a one-point increase to 61 points. However, below the 60-point minimum threshold, an increase from 4.4m to 4.5m would result in a 5-point increase, from 55 to 60 points. With the SPT being the second most failed event by females and average SPT points pre-semester being 39.43 ± 27.26, even small, insignificant changes in raw distance thrown will result in large, significant changes in points when performances are below 60 points.

Multinomial Logistic Regression Odds Ratios

Female cadets were significantly more likely to "FAIL" or score in the "MODERATE" or "SIGNIFICANT" categories of the SPT. Thus, male cadets were more likely to score in the "HEAVY" category than females. With high odds ratios, such as Exp (B) = 119.25 and 58.364, sex may be the most important variable determining SPT performance. Height may play a key role in this disparity as males are generally taller than females and cadets with increased height are significantly less likely to “FAIL” or score in the “MODERATE” category. Likewise, body mass may play a key role as increased mass was associated with decreased likelihood of failing or scoring in the “MODERATE” and “SIGNIFICANT” categories.

Significant differences between MS-classes were calculated and thus, MS-class has shown to also be a significantly predictor for SPT performance. MS-4 and MS-3 cadets being less likely to score in any category besides “HEAVY” when compared to MS-1 cadets. MS-2 cadets were less likely to “FAIL” when compared to MS-1 cadets. Therefore, it can be concluded, that being an MS-1 cadet, along with being a female, shorter, and/or having smaller body mass, leads to increased odds of failing the SPT.

Hand-Release Push-Up

The HRP was designed by the Army to test upper body and core muscular endurance while also reflecting a movement similar to hand-to-hand combat, pushing up from a prone firing position, and pushing a stuck vehicle (Headquarters Center for Initial Military Training, 2019). Due to the hand-release and arms reaching out to a “T”, the HRP engages not only chest, shoulder and core muscles but also upper back. The Army’s previous fitness test, the APFT, included a timed two-minute push-up event. Despite the HRP inclusion of a hand release to a “T”, there is some carry over from the APFT to the ACFT.

Sex Differences

Males performed significantly better than females when comparing both HRP raw and scaled scores. On average, males performed 10 more repetitions and scored 9 more points. Large effect sizes were observed for the difference between males’ and females’ raw and scaled score, 0.90 and 0.95, respectively. Previous research analyzing sex differences in APFT performance also found there to be significant differences between males and females for the timed two-minute push-up event (Draicchio et al., 2020; Knapik et al., 2001; J. Pierce et al., 2018). The largest difference between male and female basic trainees' APFT performance was 52% difference in the push-up event before BCT and 33% after BCT. Differences in the 2MR and sit-ups were less (J. Pierce et al., 2018). In this current study, females completed approximately 25% less HRP repetitions than males.

However, out of the six ACFT events, males and females scored most similarly on the HRP; the gap between scaled scores was more than 10 points less than all other events. It is the only event in which the average point value for females exceeded the minimum for the "HEAVY" category (70 points). More females scored in the "HEAVY" category of the HRP than any other event (151 females, or 59.0%). Additionally, the HRP was the least failed event of the ACFT with only 2 female failures and 0 male failures. All of this may be due to a variety of factors. First, the HRP and the APFT's push-up event are similar so cadets have been training these muscle groups and events longer. Second, training the HRP requires no equipment and can be done anywhere and therefore can be included into every physical training session if necessary. Third, the HRP necessitates significant core strength because a requirement for successful repetition completion is a straight back, such as in a plank position. Males and females have been shown to not be significantly different in core strength (Courtright et al., 2013; Draicchio et al., 2020). Therefore, it would seem to not be a limiting factor in HRP performance; only upper body muscular endurance and strength would be.

Military Science Class Differences

MS-classes were not significantly different in HRP raw or scaled scores. Although reasonable to hypothesize upperclassmen cadets would perform better due to more experience/training in preparation for a push-up type event in the APFT (previously) or the ACFT (currently), this was not the case. While MS-3 and MS-4 cadets had greater points averages, the differences were not large enough to be statistically significant. Without additional data, the underlying reasons for this finding cannot be confidently postulated or explained.

Anthropometric Correlations

The HRP event was significantly correlated with all anthropometric variables but only presented low, weak correlations. Height, mass, and BMI did not present negative correlations with HRP performance. Therefore, in the scope of this study and a combined population, increased mass and BMI presented as neither a detriment nor an advantage for HRP performance. When separated by sex, the correlations became weaker for both males and females and insignificant for males. The correlations between HRP and height became negative for males and females once separated by sex but were only significant for females and HRP raw scores. This may be because increased height necessitates a greater demand from core strength to keep the body straight during the push-up. Interestingly, once separated by sex, correlations between mass and HRP raw performance also became negative for females. This correlation was significant. Therefore, increased mass in females was not necessarily advantageous to HRP performance. If the increased mass is due to increased FM, and not increased FFM, a negative relationship would be expected. However, without body composition data, this certainty of our current finding remains unclear.

Pre- & Post-Semester Outcomes

Neither males nor females significantly improved their HRP performance over the course of a semester of training. Because push-ups have been included in previous fitness tests and HRP is one of the most “passed” tests of the ACFT, it is plausible little attention was given to this event during the course of training. With both males and females failing other ACFT events at a higher rate, physical training may have been focused more on improving performance in those events rather than the HRP.

Multinomial Logistic Regression Odds Ratios

There were no significant predictors that put a cadet more or less likely to “FAIL” the HRP. This may be due to only 2 cadets out of the whole population failing this event. As for the other qualification categories, females were more likely to score in the “MODERATE” and “SIGNIFICANT” category than males. Subsequently, males were more likely to score in the “HEAVY’ category than females. More females scored in the ‘HEAVY’ category in the HRP than any other event and therefore, it may not be as much of a limiting factor for overall qualification category as the other ACFT events. Despite there being no significant difference between MS-classes for raw or scaled scores, MS-3 cadets were more likely to score in the “MODERATE” and “SIGNIFICANT” categories than MS-1 cadets. Similarly, MS-4 cadets were more likely to score in “SIGNIFICANT” than MS-1 cadets. Therefore, being an upperclassman does offer some statistical advantage in performance of the HRP.

Sprint-Drag-Carry

The SDC was designed by the Army to test anaerobic capacity as well as muscular endurance and strength as soldiers move through a variety of weighted and unweighted forward and lateral movements. The SDC was created to reflect movements similar to reacting to direct and indirect fire, extracting a casualty and moving them to safety, and other high intensity, relatively brief tasks (Headquarters Center for Initial Military Training, 2019). A Blue Ribbon Panel has rated moving with agility, dragging/carrying heavy loads, and moving quickly over short distances as four of the nine common military tasks (Nindl, Alvar, et al., 2015). The SDC attempts to test more of the common military tasks than any other ACFT event. Thus, high performance in the SDC may be integral for mission safety and success.

Sex Differences

Males completed the SDC significantly faster and scored significantly more points on the event than females. The observed effect size for SDC raw and points were very large at 1.71 and 1.74, respectively. Through the various shuttles, the SDC tests speed and agility during the sprints and laterals as well as lower and upper body strength during the sled drag and farmer's carry. The literature shows males and females to have significantly different lower and upper body strength as well as speed and agility (Courtright et al., 2013). In an event that combines so many realms of fitness, differences between the sexes were expected. No official Army fitness test in the 21st Century has incorporated events to assess anaerobic capacity. However, laboratory tests performed on soldiers, like the Wingate test, have revealed male basic trainees to have 28.6% higher anaerobic fitness than females (Yanovich et al., 2008). Likewise, active duty 101st Airborne soldiers showed significant differences in anaerobic power and capacity between the sexes, as tested by the Wingate test (Allison et al., 2015). While direct comparisons between the SDC and the Wingate test cannot be made at this time, if the SDC truly tests anaerobic capacity, the sex differences in performance align with the literature.

When assessing point averages, males scored the most points in the SDC compared to any other ACFT event. Their 88.78 ± 11.45 points performance was about 8 points higher than any other event. Only 1 male cadet failed this event. Additionally, more males scored in the "HEAVY" category in this event than any other event; 679, or 89.7%. Similarly, females scored the second highest on the SDC (HRP was the highest overall) with 25 failures (9.7%) and 63 (24.6%) females scoring in the "HEAVY" category. Given this information, the SDC may be the most well performed event in this cohort of cadets and may not be a limiting factor in achieving desired MOS qualification categories.

Military Science Class Differences

Upperclassmen cadets performed significantly better on the SDC than MS-1 cadets with MS-4 cadets having the highest scores and lowest times of all classes. Despite no recent Army fitness test involving an anaerobic, shuttle run type event, the cadets with more military experience performed better on this test. This may be a testament to the overall better fitness of upperclassmen cadets when compared to the freshmen. Even if MS-3 and MS-4 cadets had not been practicing and training for the SDC for 3-4 years, they had been partaking in other military training exercises that could have some carry over to this event. Moving a casualty to safety during a mission or field training exercise is similar to dragging the 90lb sled. Carrying two standard five-gallon jerry cans of water is similar to the two, 40lb kettlebell farmer's carry in this event. Therefore, without direct exposure to the SDC, upperclassmen cadets have still had more experience with similar activities which may have led to superior performance on this event. Also, one can assume that being in ROTC for a longer period of time may result in overall higher fitness levels than the freshman cadets who have been in the program for no more than three-months.

Anthropometric Correlations

The SDC raw time had significant, moderate negative correlations with height and mass. As height and mass increase, SDC raw time decreases, a desired outcome. Similarly, as height and mass increased, SDC points increased; a significant, moderate positive correlation. Male cadets were significantly taller than female cadets which may provide them an advantage in this event. As for body mass, visual analysis of the scatterplots reveals a ceiling effect. SDC performances tend to increase as mass increases up to around 120kg. Past the 120kg point, performances tend to decline. This may be because the additional mass above 120kg is due to FM and not FFM, thus, these cadets have more metabolically inactive tissue. It is possible too, that beyond a given quantity of FFM, there are diminishing returns for force production versus force required to maintain velocity of a heavier body. Nevertheless, additional body mass up to 120 kg may be beneficial if it is due to increased FFM.

BMI was significantly, but weakly correlated with SDC outcomes. Previous literature has concluded higher BMI to be correlated with reduced speed and agility, tested through a 400m sprint, 300m shuttle run, and Illinois agility test, in both males and females (J. R. Pierce et al., 2017). In contrast, this study found higher BMI to be significantly positively correlated with SDC points and significantly negatively correlated with SDC time. It is plausible that because the SDC incorporates weighted movements (sled drag and kettlebell farmer's carry) a higher BMI is more advantageous than if the test was unweighted. Higher BMI has been associated with improvements in muscular strength and power in both males and females (J. R. Pierce et al., 2017). Therefore, inclusion of shuttles requiring strength, in addition to speed, may had led to the significant correlations observed between higher BMI and better SDC performance.

When separated by sex, correlations between SDC outcomes and anthropometric outcomes decrease in strength. Correlations are still significant between SDC outcomes and mass and height for both sexes. BMI also remained statistically significant in its correlation with SDC raw scores for males. All other BMI correlations were not significant.

Pre- & Post-Semester Outcomes

Males and females both significantly improved their SDC performance by the end of the semester. Males decreased their time by approximately 10 seconds and females by 30 seconds.

For the male cadets in this sub-population, the SDC event was their highest scoring event during the pre-semester test. Despite that, it was one of the two events they significantly improved their raw score in. In contrast, the SDC was one of the female cadets' worst events, just behind the LT and the SPT and was one of the two events in which females improved their raw score. Because ROTC units, like squads and platoons, often perform physical training together, it is possible that increased time was spent training for the SDC by both sexes because females performed poorly at the beginning of the semester. This would lead males to continually increase their performance in the event as well. It is also possible that because no Army fitness test has included an event like the SDC, more time was given to it during training sessions to familiarize the cadets.

The pre-semester scores would result in females, on average, failing the event because the minimum score is 60 and females only achieved 48.95 ± 28.34 points. Females' post-semester scores, however, would qualify them for the "SIGNIFICANT" category in the SDC as there was an increase of almost 20 points. Males qualified for the "HEAVY' category in this event at the pre-semester and post-semester test. So, increases in performance for the males did not result in further qualification increase, only point increases.

Multinomial Logistic Regression Odds Ratios

Height was a significant predictor in the multinomial logistic regression equation analyzing all categories versus "HEAVY". Those who are shorter are more likely to "FAIL" the SDC event than those who are taller. Similarly, height was a significant factor in the "MODERATE" and "SIGNIFICANT" equations with those being shorter more likely to score in those categories rather than in the "HEAVY" category. Height had the strongest correlations with SDC outcomes when compared to mass and BMI. Overall, cadets with greater height are less likely to score in any category but "HEAVY" in the SDC because those who are taller often have longer legs and subsequently longer stride length. These may be equated with increased speed, an integral component of this shuttle style event.

Increased height is connected to sex with males being taller than females on average. The correlations between height and SDC outcomes decreased when separated by sex. So, when analyzed together (not separated by sex), increased height being associated with better performance may also be a function of sex and the physiological differences between sexes that could not be measured in this current study. Nevertheless, sex was an important predictor in these equations. Females were significantly more likely to score in both the “MODERATE” and “SIGNIFICANT” categories than in the “HEAVY” category. Accordingly, males were more likely to score in the “HEAVY” category than any other. These predictors are congruent with the number of males and females that failed and scored in the “HEAVY” category of this event.

Consistent with previous results, MS-class was a significant variable in these predictor equations. Upperclassmen were significantly less likely to “FAIL” or score in the “MODERATE” or “SIGNIFICANT” category than freshmen. MS-3 and MS-4 cadets performed significantly better on the SDC than the MS-1 cadets and thus, it is not unexpected for this measure of military experience to be significant. Despite not scoring significantly different in previous analyses, MS-2 cadets were less likely than MS-1 cadets to score in the ‘MODERATE” or “SIGNIFICANT” categories. Hence forth, simply not being an MS-1 enhances a cadet’s chance to perform better. In a practical sense, this may necessitate giving special attention to training up MS-1 cadets their first semester involved in the program.

Leg Tuck

The LT was designed by the Army to test grip, upper body, core, and hip flexor strength with hopes that these will translate to climbing tasks, surmounting obstacles, and low crawling

(Headquarters Center for Initial Military Training, 2019). This event requires not only strength but coordination as soldiers are expected to flex their arms and hips at the same time and in a controlled manner. The Army has identified strong, well-conditioned core muscles as helpful to decreasing injuries to the upper and lower back during load carriage (Headquarters Center for Initial Military Training, 2019). Performance in the LT event may be indicative of a soldiers' ability to move and hold their own body weight in its entirety, something that is not tested by any other event in the ACFT or in the APFT and OPAT.

Sex Differences

Overall, the LT proved to be the toughest event with both males and females having their lowest points scores out of all ACFT events. Nonetheless, males still performed significantly better than females. Males were able to perform enough leg tucks to qualify them for the heaviest category. In contrast, females' average points would not meet the minimum qualification standard. Interestingly, females' average LT repetitions, 2.11 ± 3.15, would meet the minimum qualification of just one repetition. The LT event, like the HRP events, tests core strength as well as upper body strength. However, the differences between males and females for the LT were larger than during the HRP event; average 35 points versus 9 points difference, respectively. This may be because the LT requires more pulling strength while the HRP requires pushing strength.

Previous research has identified differences in push-to-pull ratios among males and females. In recreationally active adults, males' push-to-pull ratio was 1.57:1 whereas females' was 2.72:1 (Negrete et al., 2013). This means that females pulling muscles are a lot weaker than their pushing muscles when compared to males. In other words, males can push 157% of what they can pull, and females can push 272% of what they can pull. Though the LT is not strictly a pulling event, larger push-to-pull ratios may be a possible cause why females performed so poorly on this event and why the gap between sexes is so large when compared to the HRP.

The Army has not included a pull-up type event in any 21st Century fitness tests. However, other tactical populations, such as law enforcement officers (LEO) and the United States Marine Corps, have included pull-up events in their fitness tests. Analysis of LEO trainees pull-up performance during their police academy concluded almost 90% of female recruits scored in the bottom three percentiles of pull-ups while only 23.5% of males scored in those percentile ranges (Lockie et al., 2020). Being in the bottom three percentiles corresponded with performing 5 or less repetitions. 80% of females performed 2 or less pull-up repetitions. While statistical significance of the differences between the sexes was not analyzed, when expressed as percentile ranks, females performed worse than males in that event. Similarly, female Marines performed, on average, only 25% of the pull-up repetitions that males performed on the USMC physical fitness test, 16 ± 4 v. 4 ± 2 (Jameson et al., 2015).

Given the large difference between sexes in performance on pulling type events, the high rate of failures among women was expected. There were 102 female cadets failed the LT in this current study as compared to 30 males, or 39.8 % v. 4.0%, respectively. This was the most failed event by females and the second most failed event by males. When the LT event was eliminated from the test, only 75 females failed the ACFT as opposed to 124 females when the event was included (29.29% vs. 48.4% of the female cadet population, respectively). Removal of this event resulted in 41 more females qualifying for the "MODERATE" category and 8 more females qualifying for the "SIGNIFICANT" category. Surprisingly, removal of this event did not bump any females up into the "HEAVY" category. Therefore, while LT may be negatively impacting females' ability to pass the minimum standards, it does not hinder qualifying for the heaviest category. In fact, 54 females, or 21.1% of females, qualified for the “HEAVY” category of the LT, more than both the SPT and the MDL. In contrast, eliminating the LT did decrease male failures by 2.6% and increase those qualifying for the heaviest category by 5%. Removal of the event, however, did not have as large of an effect on males as it did females.

Military Science Class Differences

MS-4 cadets had the highest average LT points with each subsequent MS-class having less. Despite that, MS-classes were not significantly different in LT performance. This may be due to the relative new nature of the LT event. As previously stated, no Army fitness test during the 21^{st} Century has included a hanging pull-up type event. Therefore, upperclassmen and underclassmen have had similar exposure to the LT as a fitness test event. The Army has recommended doing “Climbing Drills”, such as alternating grip pull-ups, leg tucks, and straight arm pulls, as part of regular physical training in both 2012 and 2020 versions of FM 7-22 (Headquarters Department of the Army, 2020). Despite recommendations, implementation of climbing drills may not have been common practice within ROTC programs as they did not directly relate to the APFT and thus, upperclassmen cadets do not perform significantly better on the LT.

Anthropometric Correlations

LT performance was significantly, but not moderately, correlated to height and mass. Increased height provides no clear advantage towards LT performance. Therefore, height’s correlation with LT performance may be a product of males being significantly taller than females and the other physiological differences that give males’ an advantage rather than simply just increased height. Increased mass may provide an advantage in LT performance if the mass is primarily FFM and not FM, or dead weight. A previous study has found total body mass and lean body mass (FFM) to be significantly correlated with one repetition maximum (1RM) pull-up weight in males and females (Johnson et al., 2009). FFM could not be assessed in this current study. Like height, mass's correlation with LT may again be a product of males weighing significantly more than females and therefore, the other physiological differences that give males' an advantage rather than simply just increased mass. This is confirmed by insignificant correlations between LT outcomes and height and mass when separated by sex. In other words, height and mass were not significantly correlated with LT outcomes when comparing males to males and females to females.

Interestingly, when analyzing the population as a whole, BMI did not present significant correlations with LT outcomes. When separated by sex, BMI became significantly negatively correlated with males' LT raw and scaled scores. That is, as BMI increased in male cadets, LT performance decreased. Females' BMI correlations with LT also became negative but were insignificant. It's plausible that the lower variability in scores and lower magnitude of repetitions performed by females may be reducing the chance of a significant finding in this regard. Negative correlations between BMI and pull-up performance have been seen previously. In male Army cadets, BMI was negatively correlated with pull-ups; -0.22 (p=0.002) (Nikolaidis, 2014). While the correlations were not as strong in this study, BMI was negatively correlated with male cadets' LT raw and scaled performance in the ACFT, -0.09 and -0.113, respectively.

Pre- & Post-Semester Outcomes

Over the course of a semester, female cadets significantly improved their LT performance while male cadets did not. Females' average leg tuck repetitions performed was less than one during the pre-semester test and was greater than one at the end of the semester. This resulted in the average points scored by females doubling; 18.43 ± 29.88 to 36.62 ± 32.72

points. The Army's 0-100 scaling system does not include any point values less than 60 points for the leg tuck as 1 repetition is worth 60 points and 0 repetitions is worth 0 points. Therefore, while the average female repetitions increased to above the minimum for "MODERATE" qualification, the average points included females that still could not complete one repetition and therefore, decreased the average points to well below the "MODERATE" standard. Moderate effect sizes were observed for females' improvement in the LT with 0.70 and 0.58 observed for LT raw and LT scaled scores, respectively.

Male cadets not significantly improving in the LT over the course of the semester is unexpected. LT was the lowest point scoring event for males during the pre-semester test. It remained the lowest point scoring event during the post-semester test. This is in contrast to earlier conclusions that tied male increases in SDC performance to more unit focus on training that event because females performed poorly. Poor female LT performance during the pre-test did not seem to effect males in the same way. Males increased their average repetitions by approximately 0.5 repetitions and females by approximately 0.75 repetitions. Despite a similar increase, 0.5 repetitions represent a smaller percent of total repetitions for males than it does for females. Thus, they were not significantly different post-test.

Multinomial Logistic Regression Odds Ratios

Females were at far greater odds to "FAIL" and/or score in the "MODERATE" category, rather than the "HEAVY" category, in comparison to their male counterparts. Congruent with previous analysis, almost 40% of females failed the LT and only 4% of males did. In contrast, 21% of females and 78% of males scored in the "HEAVY" category. Therefore, while males are at far greater odds to score in the "HEAVY" category, some female cadets are beating the odds. Height proved to be a significant predictor as cadets with greater height were less likely to

“FAIL” the LT event. However, as discussed previously, this may be a result of the other physiological differences that accompany males’ significantly greater height than females. Thus, height is only a predictor when analyzing the “FAIL” category and not any other. In contrast to previous correlations, cadets with greater mass were more likely to “FAIL” or score in the “MODERATE” category in this event than in the “HEAVY” category. Correlations resulted in mass being positively correlated with LT performance. However, stepwise multinomial logistic regression resulted in mass being a negative predictor.

MS-3 and MS-4 cadets were less likely to “FAIL” or score in the “MODERATE” category than MS-1 cadets. While no significant differences were observed between classes, the upperclassmen cadets did have more average leg tuck repetitions and points than the MS-1 cadets did. Therefore, being an upperclassman is generally advantageous in this event. Surprisingly, no variables were significant predictors for being more or less likely to score in the “SIGNIFICANT” category. This may be due to small range in which cadets would qualify for the “SIGNFICANT” category. Many cadets struggled to get one repetition, so there are numerous failures (repetitions = 0) and “MODERATE” (repetitions = 1-2) qualifications. Once a cadet is able to do multiple repetitions, the minimum for “HEAVY” qualification is well within reach; 5 to 20 repetitions. To qualify for the “SIGNIFICANT” category, a cadet would have to perform 3 or 4 repetitions. The range for this qualification category is small. Thus, only about 6 percent of the entire population in this study scored in the “SIGNIFICANT” category in this event.

Two-Mile Run

The 2MR was designed by the Army to test aerobic endurance, a trait deemed necessary for dismounted movement, ruck marching, and infiltration (Headquarters Center for Initial

Military Training, 2019). The 2MR is the only event on the ACFT that was also a part of the APFT. The standards for passing the 2MR on the ACFT are lower than on the APFT with 18:00 minutes being the threshold for “HEAVY” qualification whereas 18:00 would have resulted in males younger than 36 years old failing this event in the APFT.

Sex Differences

Congruent with previous literature, males performed significantly better on the 2MR event. Males ran the 2MR over 120s faster and, as a result, scored 15 more points than females, on average. In another recent investigation, male Army ROTC cadets taking the APFT, also ran the 2MR over 120s, or two-minutes, faster than females did, 903s ± 126s vs. 1041s ± 151s (Draicchio et al., 2020). The observed effect size between sexes in that APFT 2MR was 1.02. Similarly, in the current study, the observed effect size was also large at 0.96. Significant differences have also been seen in basic trainees after BCT completion. Despite months of physical training, an 18% difference in 2MR performance persisted between the sexes (J. Pierce et al., 2018). Similarly, when comparing raw seconds, female cadets ran the 2MR approximately 15% slower than male cadets did on the ACFT.

A study conducted at the U.S. Army Research Institute of Environmental Medicine (USARIEM) concluded in 1987 that a maximal effort 2MR was significantly and strongly correlated to VO_{2max} (Mello et al., 1987). While this current study does not attempt to correlate the 2MR with other fitness predictors or occupational performance, highlighting the relationship between the 2MR event and VO_{2max} is important for analyzing why there are differences in performance between the sexes. On average, females demonstrate a VO_{2max} from 15-30% lower than males (Drinkwater, 1973; Yanovich et al., 2008). When examining 101st Airborne soldiers, males displayed significantly greater relative VO_{2max} on a treadmill graded exercise test when compared to female soldiers (Allison et al., 2015). Differences in absolute and relative VO_{2max} have been attributed to physiological differences between the sexes like heart size and hemoglobin content. While muscle mitochondrial content, capillary density, and blood volume (positive adaptations for increased VO_{2max}) can be increased due to physical training, smaller heart size, an inherent characteristic of females, cannot be changed. Thus, differences in 2MR performance will persist between sexes in a trained population.

Despite the 2MR being an event on the APFT, it was the most failed event by male cadets taking the ACFT. The 2MR was the third most failed event for females. In ROTC Cadets taking the APFT, the 2MR was the most failed event by both male and female cadets, with 24% and 35% failing, respectively (Draicchio et al., 2020). In the current study, approximately 4% of males and 14% of females failed the 2MR. Due to the five other events that are performed before the 2MR (the final event), the standards for this event are relaxed in comparison to the APFT. Nonetheless, there remains an approximate 10% difference in failure rates between sexes. Because of the sheer number of physical tasks a cadet must perform before the 2MR, cadets may have performed worse than if they ran the 2MR by itself. So, despite this event being a part of the old fitness test, cadets may struggle to perform their best due to fatigue. Thus failure rates are higher than one would expect for the more relaxed standards. Nevertheless, 45% of females were able to run fast enough to score in the “HEAVY” category, second greatest percent behind the HRP. The majority (82%) of males were able to score in the “HEAVY” category. Thus, despite accumulated fatigue, the 2MR standards are set at a point in which both sexes are still able to run fast enough to score in the “HEAVY” category.

Military Science Class Differences

MS-3 and MS-4 cadets ran the 2MR significantly faster and scored significantly more points than MS-1 cadets. The most likely cause is better overall cardiorespiratory, or aerobic endurance, fitness because the upperclassmen cadets have been performing military physical training for 2-3 more years than the MS-1 cadets. The APFT and the OPAT both contain events that test aerobic endurance; the 2MR and the interval aerobic run. Upperclassmen cadets, though with similar ACFT experience, would have taken the APFT and the OPAT considerably more times than the MS-1 cadets. Therefore, if sheer exposure to aerobic endurance tests determined 2MR performance outcomes on the ACFT and not VO_{2max}, MS-3 and MS-4 cadets would be at an advantage. MS-2 cadets, though not significantly different, ran the 2MR faster and scored more points in the event than MS-1 cadets. MS-4 cadets had the fastest 2MR times and MS-1 cadets the slowest. There is an observed trend that the longer a cadet has been in ROTC, the faster they will run the 2MR.

Anthropometric Correlations

The 2MR raw scores had significant, weak to moderate correlations with mass, height and BMI while the 2MR scaled score had significant, weak to moderate correlations with just mass and height. Of all three anthropometric variables, height had the strongest correlations with 2MR performance. Again, this may be a function of males being significantly taller than females, along with, the aforementioned physiological differences between the two sexes rather than height alone accounting for the variability in 2MR run performance. When separated by sex, height’s correlation with 2MR performance becomes insignificant and extremely weak.

Mass presented with significant yet weak, negative correlations with 2MR raw performance when analyzing both sexes together. Increased mass, therefore, is correlated with faster 2MR times. As with height, this may be a function of other sex differences and thus must be analyzed when the sexes are separated. The correlation between mass and 2MR performance for males became insignificant once the sexes were separated. In contrast, the correlations between mass and 2MR performance for females became significant. Females' mass was moderately correlated with 2MR raw scores; the greater mass a female has, the slower her 2MR time is. Likewise, females' mass was negative and moderately correlated with 2MR points. Greater mass for females was correlated with less 2MR points. It is important to note how direction of correlations with mass changes when the analysis was separated by sex. When analyzed together, increased mass seems advantageous. However, for females in this study, it turns out to be harmful to performance. It can be postulated that the reason increased mass is not harmful to males is because their mass is due to increased FFM, metabolically active tissue. In contrast, females increased mass may be due to FM as it has been shown that females', on average, have significantly more FM than males (Knapik et al., 2001). Therefore, interpretation of correlations that are not separated by sex is cautioned.

BMI presented a significant positive correlation with 2MR raw scores for the combined population. At 0.077, the correlation is very weak but can be interpreted as higher BMI is correlated with longer 2MR duration (or slower 2MR average run speed). Higher BMI was correlated to decreased performance by basic trainees in the OPAT interval aerobic run, a test of aerobic endurance (J. Pierce et al., 2018). Conversely, in Army ROTC, cadets' 2MR performance was not significantly correlated with BMI (Draicchio et al., 2020). When analyzed by sex, BMI's possible effect on 2MR performance becomes clearer. BMI became significantly and positively correlated to 2MR raw performance for both males and females when separated by sex. Additionally, BMI became significantly and negatively correlated with 2MR scaled score for both sexes. Thus, greater BMI is correlated with decreased 2MR performance. Researchers have previously concluded that BMI greater than 25 kg/m^2 has a negative influence on cardiovascular performance. Additionally, high BMI was correlated with increased 2MR time in female basic trainees (Kazman et al., 2015; J. Pierce et al., 2018). While correlations between the 2MR and height and mass decreased in strength when separated by sex, the correlations with BMI increased in strength. Thus, BMI may be a better predictor for 2MR performance than height and mass alone.

Pre- & Post-Semester Outcomes

Only males significantly altered their 2MR performance during the semester. The male cadets' post-semester 2MR scaled score was significantly better than their pre-semester scaled score with an average 3.5 point increase. The observed effect size, 0.17, is minimal and has weak practical relevance. Consequently, no changes in 2MR performance during the semester made any noteworthy impact on the end of the semester ACFT. Because the 2MR was an event on the APFT, more attention may have been given to other ACFT events that cadets are less familiar and/or performed worse in. The 2MR is the only event on the ACFT that tests aerobic endurance. All five other events are designed to test muscular endurance, muscular strength, power, and/or anaerobic capacity. It is then plausible that gains in performance on a muscular strength-based event, such as the MDL, may lead to performance gains in another lower body strength and power-centric event, like the SPT or SDC. Performance increases in the 2MR may only be significantly impacted by training for the 2MR. If aerobic endurance training is decreased to focus on the other less familiar events, performance gains may be attenuated. Because the 2MR was the most failed event by male cadets, a decrease in aerobic training to focus on other ACFT events is not recommended.

Multinomial Logistic Regression Odds Ratios

Female cadets were more likely to “FAIL” or score in the “MODERATE” and “SIGNIFICANT” categories than males. Subsequently, males had greater odds of scoring in the “HEAVY” category. This is congruent with the sex differences that this and other studies have observed when it comes to the 2MR (Draicchio et al., 2020; J. Pierce et al., 2018). Mass presented as a significant predictor in this equation. Greater mass was associated with greater odds of failure and scoring in the “MODERATE” group rather than the “HEAVY” group. While analysis of mass alone was cautioned earlier in the discussion, analysis of mass in regard to a specific sex may be useful. Females overall have greater odds of failing. Based on this regression equation, females who have greater mass are at even greater odds of failing or only scoring “MODERATE”. This is congruent with the significant positive correlations observed between female cadets’ body mass and 2MR raw scores.

MS-class was a significant predictor in this equation. MS-1 cadets were more likely to “FAIL” or score in the “MODERATE” category than MS-3 and MS-4 cadets. This is consistent with the significant differences observed between upperclassmen and MS-1 cadets in 2MR performance. MS-2 cadets were less likely to “FAIL” and more likely to score in the “SIGNIFICANT” category than MS-1 cadets. While MS-1 and MS-2 cadets were not significantly different when 2MR performances were compared, MS-2 cadets had better performances on average. Thus, their average overall 2MR qualification level would be higher.

Total ACFT Score

Total ACFT scores are a combination of the scaled 0-100 scores from each of the six events. The total score out of 600 does not qualify a soldier for a specific MOS. Rather, minimum qualification categories, which will be discussed in the next section, must be met in every event for a soldier to achieve a particular qualification standard. Nonetheless, total ACFT points may be important for comparing soldiers vying for the same job (such as ROTC cadets applying for a specific MOS) or for soldiers' promotions.

Sex Differences

Male and females had significantly different total ACFT scores, with males scoring over 100 points more than females did on average. Comparing the ACFT with the APFT is difficult because the APFT used a sex- and age-based scoring scale. A perfect score for males on the APFT consisted of more push-ups, sit-ups, and a faster 2MR than a perfect score for females on the APFT. However, it can be concluded that because females performed significantly worse on the push-up and 2MR events compared to males, if they were graded on the same scoring scale, females would have a lower overall score than males (Draicchio et al., 2020; J. Pierce et al., 2018). The OPAT did not include a 0-100 scoring scale and was based on minimum standards for qualification categories. Therefore, the current ACFT total score results cannot easily be compared to the OPAT either.

The observed effect size for the differences between male and female cadets' total ACFT scores was large at 1.82. This is the second largest effect size observed for sex differences in this study, only behind total qualification level differences. The large effect size observed for total ACFT scores is congruent with the literature. A meta-analysis investigating sex-based musculoskeletal and cardiovascular differences in physically demanding occupations (military, LEO, firefighters, etc). found observed a large, 2.13 effect size for differences in the ability to produce muscle tension (Courtright et al., 2013). The study defined muscle tension as exerting force against objects in pushing, pulling, lifting, etcetera. Five out of the six ACFT events incorporate use of “muscle tension” as defined by these authors, with the 2MR being the only outlier. Thus, overall differences in total ACFT and a large effect size were anticipated.

Such large differences between the sexes for average total ACFT, if unchanged or un-scaled, may put males at an advantage for obtaining their desired MOS or getting promoted. This could become problematic if fitness scores are the only component separating a male and female soldier. Thus, care should be taken in interpreting total ACFT score as it relates to a soldier’s ability to perform a job and/or their overall state of fitness.

Military Science Class Differences

MS-3 and MS-4 cadets had significantly higher total ACFT scores than MS-1 cadets. Moderate effect sizes were observed for both; 0.43 and 0.45, respectively. Due to significantly more time partaking in military physical training, differences between the MS-classes for total ACFT score were expected. Upperclassmen cadets have completed 2-3 more years of physical training than the freshman. So, regardless of the ACFT being a new test, the MS-3 and MS-4 cadets are generally more fit. Cadets do not vie for their potential MOS until their MS-3 year. Therefore, performing significantly worse as a freshman would not necessarily adversely impact their career in the Army at that time. Total ACFT scores increased for each MS-class with the MS-4 cadets having the greatest average total score. Thus, one would expect MS-1 cadets, despite their performance as freshman, to score higher in subsequent years.

Anthropometric Correlations

Total ACFT points were significantly and positively correlated with height, mass, and BMI. Height was moderately correlated, $r_s = 0.514$, with total ACFT points. This significant correlation persists for only males once analysis was separated by sex. When comparing males to other males, however, being taller may be advantageous. Unfortunately, that is an inherent trait that cannot be changed through physical training. Additional height was not advantageous for females in regard to higher total ACFT scores. Therefore, the moderate correlation between height and total ACFT scores when not separated by sex may be a function of males being significantly taller and the inherent physiological advantages, such as higher VO_{2max} and greater average FFM, when compared to females.

Mass was significantly and moderately correlated with total ACFT scores as seen in the Figure 6 scatterplot. Mass seems to have a ceiling effect when it comes to performance advantages. Through visual observation, one can note that cadets at and above 120kg scored similarly on the total ACFT to cadets with less mass. These heavier cadets, therefore, are contrary to the notion that heavier is advantageous in ACFT performance. Regardless, increased mass has proved to be advantageous on various individual ACFT events and thus may play a role in total ACFT scores. When separated by sex, mass remained significantly and positively correlated with total ACFT scores in male cadets only. Though insignificant, the correlation between mass and total ACFT score for females became negative. That is, greater mass was associated with poorer performance on the ACFT. With all this taken into account, it is important to note that increased mass is only advantageous and weakly correlated with better performance for male cadets. A measure of body composition seems warranted for better understanding of the patterns observed.

Similar to height and weight, BMI was significantly and positively correlated to total ACFT scores but became insignificant for females once separated by sex. Increased BMI remained advantageous for males. So, all anthropometric measures became insignificantly correlated with total ACFT scores in female cadets once analysis was separated by sex. This leads to the conclusion that it is not height, mass, or BMI alone that determines ACFT performance in female cadets. Other variables that were not measured in this study, such as FFM, number of training days per week, or type of physical training, set female cadets apart from one another.

Previous research has shown BMI to be negatively correlated with APFT scores while other studies have shown BMI to have not been significantly correlated with APFT scores or with warrior task and battle drill obstacle courses (Anderson et al., 2014; Kazman et al., 2015; J. R. Pierce et al., 2017; Steed et al., 2016). This study found BMI to be positively correlated with total ACFT scores when analyzing both sexes together and with total ACFT scores when males are analyzed alone. Thus, this current study refutes an overall negative effect of higher BMI on ACFT performance and concludes that for male cadets, it may have a positive effect.

Pre- & Post-Semester Outcomes

Both male and female cadets significantly improved their overall ACFT score by the end of the semester. A moderate effect size was observed for the change in female cadets' performance while a smaller effect size was observed in male cadets. This may be because female cadets increased their average point total by approximately 50 points while males improved by just less than 20 points. Regardless, a semester of physical training was adequate for Army ROTC cadets to significantly improve their total ACFT score. For males, total score improvements were likely the result of the significant improvements they made in MDL points, SDC points, and 2MR points. Whereas for females, total score improvements were likely the result of the significant improvements made in SPT points, SDC points, and LT points. In support of our findings, it has been shown that over the course of BCT, basic trainees significantly improve their APFT scores (J. Pierce et al., 2018). It is therefore not surprising that, ROTC cadets in the current investigation improved their total ACFT scores over a time frame longer than that of BCT.

ACFT Qualification Category

The Army designed the ACFT similar to the OPAT in that soldiers and cadets must meet minimum requirements in each event in order to qualify for a specific MOS. There are three different qualification categories with "MODERATE" being the lowest, "SIGNIFICANT" being the middle, and "HEAVY" being the highest qualification category possible. In both the OPAT and the ACFT 2.0, a soldier or cadet must meet the minimum requirement for a category in all of the events. The lowest scoring event then determines what qualification category a soldier achieves. Despite all MOS being open to females, the ACFT requires females to reach the same minimum standard as males must for all jobs. Preparing female soldiers and cadets to reach the heaviest standard is of utmost importance if the Army wishes to keep the ACFT and all jobs open to females.

Sex Differences

When calculated as an average, males and females qualified for a significantly different category level. The observed effect size between qualification categories was the largest out of all sex differences, 1.85, as males' average qualification category was 2.11 ± 1.00 while females' average qualification category was 0.57 ± 0.62. In other words, males' average qualification category was "SIGNIFICANT" while females' was "FAIL". When calculated as a percentage of the total population, 48.4% of females failed the ACFT compared to only 7.8% amongst males. In contrast, 1.1% of females qualified for the heaviest category while 50% of all males did. A stark difference in the number of males and females qualifying for the heaviest category is observed. Figure 3 illustrates the percentages of each sex that qualified for each category. 94% of females scored in the “MODERATE” or “FAIL” category while males were more distributed across the four categories.

The ACFT standards to reach “HEAVY” appear to be more difficult than the previous test, the OPAT, that utilized a category-based scoring scale. In Army ROTC cadets, 29% of females were able to qualify for the heaviest OPAT category overall while only 18% failed (Draicchio et al., 2020). This is in stark contrast to the respective 1.1% and 48.4% observed in the current study. Based on the OPAT, a female cadet may qualify to serve in an MOS that requires “HEAVY” standards but may not qualify for that same MOS when tested on the ACFT. Only 3 female cadets, from a pool of 20 universities, qualified for the “HEAVY” standard on the ACFT. As discussed previously, the LT event resulted in the greatest number of ACFT failures. However, the SPT was the event in which the lowest number of females reached the “HEAVY” standard, and the MDL was the event in which the second lowest number of females reached the “HEAVY” standard. Therefore, the lower body strength and explosive power of female cadets may be lower than anticipated and additional consideration and resources may be necessary for preparing females to meet the “HEAVY” standards in SPT and MDL.

Military Science Class Differences

When averaged, upperclassmen cadets had a significantly higher average qualification category, on a scale of 0-3, than MS-1 cadets. Additionally, MS-4 cadets had a significantly higher qualification level than MS-2 cadets. As increased military experience has played a role in better performance on many of the individual ACFT events as well as total ACFT score, it is expected that upperclassmen cadet would qualify, on average, for heavier categories. MS-3 cadets had the greatest percent of their total population qualify for the “HEAVY” category. MS-4 cadets had the smallest percent of their total population “FAIL” the ACFT. Therefore, while not only does military experience seem to play a role in qualifying for heavier categories but also acts as a protective mechanism against failing the test. Previous studies have not assessed how ROTC cadets of different MS-classes differ on performance of the OPAT. Thus, these results are unique in nature.

Anthropometric Correlations

The average level qualified for, from 0-3, was significant and positively correlated with height, mass and BMI. Greater height and mass presented moderate correlations while BMI was weak. When separating the analysis by sex, height and mass remained significantly correlated with total qualification level for males but BMI became insignificant. For female cadets, all anthropometric variables became insignificant in their correlation. Thus, height and mass may provide some advantage for male cadets when comparing them to one another. The insignificant results seen between female anthropometrics and qualification category are similar to those seen in the total ACFT. As previously concluded, there may be variables that were not measured by this study that increase performance when analyzing the test as a whole. Height, mass, and BMI alone provide no clear advantage for females in overall performance.

Pre- & Post-Semester Outcomes

Male and female cadets both significantly improved their average qualification category over the course of the semester. Observed effect sizes were moderate in nature. At the pre-test, males' average qualification level was "MODERATE" but increased to" SIGNIFICANT" after the post-test. While females' average qualification level still remained below 1 ("MODERATE") at the end of the semester, it increased in value. Thus, some female cadets were able to improve from "FAIL" to "MODERATE". A semester of physical training was an adequate enough stimulus for some female cadets to progress from failure to passing the ACFT at its lowest qualification standard. Unfortunately, over half of female cadets in this sub-population stilled failed at least one event during the post-semester test. Future studies should assess how longer periods of training, such as an entire school year, affect average qualification scores.

Multinomial Logistic Regression Odds Ratios

Female cadets are far more likely to “FAIL” the ACFT than score in the “HEAVY” category with an odds ratio of 483:1. This is the largest odds ratio observed in any of the multiple logistic regression analyses conducted in this study. Similarly, female cadets are far more likely to score in the “MODERATE” and “SIGNIFICANT” categories then they are to score in the “HEAVY” category. These results are congruent with the absolute amount and percentage of males and females that qualified for each category. Sex was the only significant predictor for the “SIGNIFICANT” category. Thus height, mass, and MS-class did not play a significant role in differentiating cadets between “SIGNIFICANT” and “HEAVY” categories. It is possible that other factors not measured in this study, such as FFM and training regimen, have a greater effect on this differentiation.

Those with greater height were less likely to “FAIL” or score in the “MODERATE” categories. As previously mentioned, use of height as a predictor is cautioned as it relates to other inherent differences between the sexes. Nevertheless, height may play a role in performance for both sexes on events such as the MDL, the SPT, and the SDC and therefore, have an effect on overall qualification category. Mass was only a significant predictor for the failure group. Those with greater mass were more likely to score in the “0” or “FAIL” group than in the “HEAVY” group compared to those with less mass. This is unexpected. Mass was significantly and positively correlated with overall qualification level for the population as a whole and for male cadets. However, the correlation became negative but insignificant for females when separated by sex. It is plausible that because so many female cadets failed the ACFT, this insignificant but negative correlation became a significant predictor in the regression equation. Those who ended up in the “FAIL” category may have had greater body mass than their higher scoring counterparts.

Being an upperclassman was advantageous and decreased the likelihood of failing or scoring in the “MODERATE” category. MS-3 and MS-4 cadets had significantly better average qualification category than MS-1 cadets. In contrast, MS-Class did not seem to play a role in distinguishing between those in the “SIGNIFICANT” versus “HEAVY” categories. Thus, military experience, as measured by MS-class, plays a significant role as a predictor in qualification category for odds of failing or scoring “MODERATE” but not in differentiating between the two heaviest categories.

Implications and Practical Applications

The ACFT was originally designed to improve the evaluation of fitness components relevant to the successful completion of occupational duties and Warrior Tasks and Battle Drills. Therefore, performance on the ACFT was expected to be indicative of a soldier’s or cadet’s ability to serve in a specific MOS and/or in certain operational environments. Because no studies have been published to date analyzing the correlations between ACFT performance and occupational tasks, this study assumes that the two are related. The ACFT was designed as a sex and age neutral test because a soldier, regardless of sex or age, needs to be able to meet minimum requirements to ensure safety and mission success. Sex-neutral fitness tests are becoming more common practice within the last decade as other militaries, law enforcement agencies, and first responders have realized the necessity of all personnel meeting a certain standard. This notion has not been argued. However, the events in which personnel are being tested is up for debate.

Researchers with the Canadian Forces Morale and Welfare Services and researchers at USARIEM analyzed common occupational physical fitness events in search of sex bias (Reilly et al., 2019). Some common events that are included in field and laboratory fitness tests, such as a 2-mile run, 300m sprint, bench press, leg press, push-ups, pull-ups, counter movement jump (CMJ), etc., were analyzed to see how much overlap there is between male and female performances. The authors caution the use of physical fitness test events that have poor overlap between the sexes because inherent physiological sex differences may result in a female being excluded from a job or MOS that they may have qualified for if a different event was chosen to test that realm of fitness. Examples of common tests that can be performed outside of a laboratory and have the most overlap between the sexes are the 2MR, one-minute of sit-ups, a flexed arm hang, and one-minute of push-ups. Examples of common tests that have poor overlap but are often used are 300m sprint, two-minutes of push-ups, pull-ups, and CMJ. Despite that, the ACFT incorporates the SDC (a 250m shuttle sprint even), a two-minute HRP event, the LT (a pull-up type event), and a standing power throw (similar to CMJ). The authors have presented other, arguably better, options for testing all realms of fitness. However, if the ACFT is going to stay how it is, conclusions need to be made based on how females are performing on this test and not how they would perform if the test was different. Nonetheless, based on calculated overlap between the sexes on events similar to the ACFT, one would expect females to perform significantly worse than males.

Female cadets did perform worse than male cadets on all individual events as well as total ACFT score and qualification category. Anecdotal evidence has shown that the leg tuck event is the hardest for female cadets. The results of this study concur with this observation. The leg tuck was by far the most failed event by females. However, exclusion of the leg tuck from calculation towards overall qualification category did not help any females reach the “HEAVY” standard. Reaching the “HEAVY” standard is the goal for females wishing to serve in the most physically demanding MOS, such as the combat arms branches that were recently opened to them. The SPT and MDL seem to present the greatest obstacle to females scoring in the heaviest category. Only 12 females threw the medicine ball far enough to qualify for “HEAVY” in the SPT. Similarly, only 34 females were able to lift enough weight to qualify for “HEAVY” in the MDL. Therefore, while the LT seems to be an impediment for females passing the minimum qualification standards, the standards for the lower body strength and explosive power events were barriers to females qualifying in the heaviest category. The standards for these events in relation to females’ abilities should be revisited and analyzed by future studies to determine the best course of action for protecting mission safety and success while also providing females with a chance at qualifying for a heavy MOS.

With less than 10% of male cadets failing the ACFT and 50% scoring in the “HEAVY” category, adverse impacts of this test on male cadets’ prospective jobs and promotions are of a lesser consequence compared to female cadets. From the results of this study, there are a few practical recommendations that can be made to help decrease the gap between the sexes and increase females’ overall performance on an un-altered ACFT. Based on the results of the subpopulation that reported pre-semester and post-semester ACFT values, it can be concluded that typical military physical training is adequate to elicit performance gains in total ACFT scores and individual qualification. Therefore, as long as cadets keep training and incorporating progressive overload, they should enhance their performance. This is also obvious in that upperclassmen cadets performed better, on average, than underclassmen. Thus, more physical training and experience seems to lead to better overall fitness and performance.

Female cadets, however, may need additional or altered physical training based on the findings presented in this study. Only 13% of females could score “HEAVY” in MDL and female cadets did not significantly improve their MDL over the course of the semester. It is possible that current physical training practices do not provide a stimulus large enough for females to sufficiently increase their performance to meet the specified MDL standards. Thus, future training should include emphasis on the MDL and improving lower body strength through progressive overload and periodized programming. Specifically, an intensity-based resistance training program should be utilized to ensure that female cadets are training with heavy enough loads to stimulate strength improvements. The use of periodized resistance training programs with female soldiers has been recommended numerous times in the literature (Nindl, Alvar, et al., 2015; Nindl et al., 2016; Nindl, Jones, et al., 2015; Sauers & Scofield, 2014). A study conducted by Kraemer et al. found that a 24 week periodized training program utilizing both total body strength/power (3-8) and strength/hypertrophy (8-12) repetition ranges were able to elicit significant performance gains in one-repetition maximum squat and bench press, squat jump, bench press throw, squat endurance, one repetition maximum box lift, and APFT scores in females (Kraemer et al., 2001). This periodized program was 24-weeks long with two 12-week meso-cycles and three four-week micro-cycles within each meso-cycle. Each micro-cycle increased in load and intensity based on percentages of 1RM. This periodized, intensity-based program was successful at eliciting performance gains in strength, power, muscular endurance, and occupational tests in these female subjects and may be beneficial for female ROTC cadets.

Likewise, the SPT proved to be one of female cadets' worst events and a barrier to scoring in the "HEAVY" category. Not only should emphasis be placed on increasing explosive power but also on coaching proper technique. The SPT requires cadets to throw the medicine ball in such a way that release height and angle will play a significant role in how far the ball goes. Thus, more explosive power training, proper instruction, and practice in throwing form could help female cadets perform better in this event.

Lastly, the leg tuck presented an obvious issue for female cadets. While female cadets did significantly improve their LT over the course of a semester, some were still unable to perform a single leg tuck during the post-semester test. As identified before, this is likely a result of poor upper body pulling strength. The Army has published climbing drills that may be helpful in training these pulling muscles. However, for female cadets and soldiers who struggle to pull their own body mass up during the LT, climbing drills may not be an often-selected training tool by the cadet or the best tool for eliciting positive adaptations. Use of a resistance training facility and weightlifting machines that allow females to train at percentages lower than their body mass may be helpful in gaining the strength necessary to be able to complete the climbing drills and the LT event. In a recent unpublished study from our lab, it was found that as little as four high intensity resistance training sessions was able to elicit strength gains above the established baseline 10RM for female cadets (Campbell et al., 2019). Thus, even a short-term resistance training protocol may be helpful for improving ROTC cadets' strength to better prepare them for the ACFT.

When separated by sex, body mass was significantly and positively correlated with performance on the MDL, SPT, and SDC for males and females. This may be the result of these events requiring lower body strength, power, and anaerobic capacity, all of which are positively affected by increased muscle mass. While this study did not obtain FFM data for cadets, it can be postulated that the reason for these correlations was that cadets with more FFM performed better. Therefore, a primary aim of all cadets, especially females, should be to increase FFM. Cadets may have to perform resistance training on their own. ROTC physical training routines only occur three times per week and are often conducted outside with minimal equipment due to large units. Analysis of performance in female Marines showed that females that are taller, heavier, and with greater estimated FFM generally outperformed their smaller counterparts on the Marine physical fitness test and combat fitness test (Kelly & Jameson, 2016). Thus, a supplementary resistance training protocol should be implemented for cadets who wish to increase strength and FFM and thereby increase their performance on the ACFT.

Increased mass was negatively correlated with 2MR performance for females. This may be the result of increased mass being due to FM or a result of improper programming of physical training. FM represents an increase “mechanical load” which increases the energy demand on the body from a work:energy perspective during an event such as the 2MR. Thus, the increased mass due to FM is disadvantageous as it will lead to fatigue earlier in the event, leading to a reduced run speed and slower time. Another possibility is that the increased mass in female cadets is FFM but so much time has been spent on strength training that aerobic endurance has been compromised to some degree. Without FFM values, the reasons behind this negative correlation can only be speculated. Regardless, programming physical training to focus on weak events but still include all realms of fitness is of utmost importance. Likewise, increasing FFM in all cadets, especially females, will help boost performance in the events that cadets have struggled in the most.

If the ACFT remains the fitness test of record for the Army and no adjustments are made to the events or scoring system, female cadets and soldiers must adapt and train to improve performance and decrease the gap between the sexes. One of the best ways to do this is increase females' FFM and program periodized training schedules that focus on weak areas without neglecting other ACFT events. There are inherent traits of females that cannot be changed through any amount of training such as: height, heart size, hormones, etc. However, body composition, force output, and a variety of cardiovascular variables can be altered through physical training. The ACFT is a vastly different test from the APFT and will require a period of familiarization across the force. Future studies should analyze the effects of various training protocols on ACFT performance to conclude the best way to prepare soldiers for this test and their occupational duties. As another study has also concluded, because some females did reach the "SIGNIFICANT" and "HEAVY" standard, the barriers to qualifying for more physically demanding occupations are not inherent but may be due to lack of training specificity (Jameson et al., 2015). Ultimately, preparing female soldiers to qualify for the MOS they wish to serve in, now that all are open to females, is most likely a matter of training and experience.

Limitations

One of the primary limitations of this study was that MS-class and anthropometric variables were not reported for all cadets. Sex and ACFT outcomes were reported for 1013 cadets while MS-class was reported for only 822 cadets and anthropometrics for 655 cadets. Additionally, the data used for the multiple logistic regression equations required sex, MS-class, height, and mass data be available for each cadet. Therefore, this reduced the total sample to only 580 cadets for use in the final equations. It is plausible that results may have been different if all cadets were capable of being included in the analyses. All variables were requested from each

ROTC program during the initial data collection phase. However, some data was unavailable or not recorded at the time the ACFT was taken. Therefore, the study proceeded with the data that was collected and shared. The results of this study can only be generalized to an Army ROTC population as no active duty, Reserve, National Guard, or basic trainee soldiers were included.

Another limitation of this study is that curvilinear model for body mass correlations was not included. Additionally, only body mass could be collected by each ROTC program. Devices to measure body composition, like DEXA scans, BIA, BIS, etc., are not readily available to ROTC programs. Other methods, like circumference measurements or skin calipers, take time and expertise that programs cannot accommodate for all cadets. While having body fat percent and FFM values would have helped bring more clarity to the conclusions drawn in this study, most ROTC programs and military units do not and/or cannot collect this information for their fitness tests. Most units do collect height and weight to calculate BMI. Thus, this current study analyzed the data that is readily available to ROTC programs and can then help make practical recommendations that are easily measurable.

Lastly, a limitation of this study was that the ROTC programs and the researchers could not control/account for physical activity that cadets conducted outside of their respective military physical training. Some cadets may do no physical training outside of the ROTC-mandated training, while others train regularly on their own. Because of this, classifying individuals solely as male/female or part of a specific MS-class may leave out important confounders to our findings. Physical training outside of required military physical training may play a role in performance on the ACFT and could have explained why females perform differently despite height, mass, and BMI not being correlated to total ACFT score. Because of the retrospective nature of this study and the large sample size, gathering this information was not possible at this time. Additionally, due to the fatigue from having to complete six events consecutively, each cadets' score may not be indicative of their absolute best performance if each event was performed alone. However, all cadets executed the test in the same manner, as outlined in the methods section, and would have been subject to similar fatigue.

Chapter V - Conclusion

The purpose of this study was multifold; to determine if and where sex-based and experience-based differences in ACFT performance exist, to analyze how anthropometric variables are related to ACFT performance, to determine if a semester of physical training is adequate to increase ACFT scores, and to present predictor variables for regression analyses of ACFT qualification categories. The five main alternate hypotheses presented were:

1. Males and female ROTC cadets will have significantly different scores on the ACFT.

Based on the results of the study, we reject the null hypothesis and accept the alternate hypothesis. Males and females scored significantly different on all ACFT events, total ACFT score, and average ACFT qualification category.

2. ROTC cadets in different military science classes will have significantly different ACFT scores.

We reject the null hypothesis and accept the alternate hypothesis. MS-3 and MS-4 cadets significantly outperformed MS-1 cadets on most ACFT events, total ACFT score, and average ACFT qualification category.

3. There will be significant correlations between BMI and ACFT outcomes.

We reject the null hypothesis and accept the alternate hypothesis. BMI was significantly correlated with several ACFT outcomes when analyzing the population as a whole and when analyzed by sex. Body mass and height also presented significant correlations with outcomes and should also be considered.

4. Male and female cadets will significantly improve their ACFT scores over the course of a semester.

We reject the null hypothesis and accept the alternate hypothesis. Male and female cadets significantly improved their total ACFT scores over the course of a semester as well as several individual event scores.

5. Sex, MS-class, and anthropometrics will play a significant role in cadets' overall qualification category.

We reject the null hypothesis and accept the alternate hypothesis. Sex, MS-class, and anthropometric variables of height and mass presented as significant predictors in some or all multinomial logistic regression equations for qualification category in individual events and in the overall ACFT.

Significance and Application

To our knowledge, this is the first study to date analyzing differences between sexes and military science class on ROTC cadets' performance of the ACFT. This study provides unique information regarding how commonly measured anthropometric variables may be related to ACFT performance of male and female ROTC cadets. Additionally, analysis of a sub-population and their pre-semester ACFT scores versus post-semester ACFT scores allows us to make impactful conclusions about the adequacy of training protocols to elicit performance gains. With all the given information, we were able to develop multiple logistic regression equations for each ACFT event and total ACFT qualification that allowed us to analyze the effects of sex, MS-class, and anthropometrics individually and together on outcomes. With this information, we can make practical recommendations that may help decrease the gap in performance between sexes. Cadets improved their scores over the course of a semester but differences between sexes persisted. Additional intensity-based resistance training protocols should be employed by ROTC programs to increase FFM and strength in female cadets and thus, increase likelihood of qualifying for heavier categories. Qualifying for and maintaining fitness to remain in a more physically demanding job is of great importance for employing females across the force in all military occupational specialties.

Future Studies

Due to the relatively new nature of the ACFT, future studies should be conducted analyzing a multitude of questions surrounding the validity and reliability of the test, application to occupational duties, and performance differences between sexes.

The body composition information available for this current study was limited. Thus, future studies should collect FFM and FM data to possibly draw more accurate and nuanced conclusions related to sex differences and performance outcomes. The results of this current study can only be generalized to an Army ROTC cadet population. Therefore, future studies should include other military populations (Active Duty, Reserve, National Guard, basic trainee, etc.) as they may have different ACFT performance outcomes than ROTC cadets. This will provide an opportunity to see if sex, military experience, and body composition/anthropometrics play a different role compared to ROTC cadets. Additionally, due to the retrospective nature of this study, outside factors, such as time spent in physical training and extracurricular activity, nutritional intake, and sleep quality could not be controlled for. Future studies should attempt to control/account for these extraneous factors and test the effectiveness of various training protocols on ACFT performance. Lastly, this study only assessed the effectiveness of one semester of physical training on ACFT performance outcomes and utilized a cross-sectional cohort for military experience. Future studies should analyze how multiple semesters and/or years of physical training impact cadets' anthropometrics, body composition, and ACFT outcomes in order to better map the progress made over the course of a cadets' tenure in an ROTC program and beyond.

References

Abe, T., DeHoyos, D. V., Pollock, M. L., & Garzarella, L. (2000). Time Course for Strength and Muscle Thickness Changes Following Upper and Lower Body Resistance Training in Men and Women. *European Journal of Applied Physiology*, *81*(3), 174–180. https://doi.org/10.1007/s004210050027

Allison, K. F., Keenan, K. A., Sell, T. C., Abt, J. P., Nagai, T., Deluzio, J., McGrail, M., & Lephart, S. M. (2015). Musculoskeletal, Biomechanical, and Physiological Gender Differences in the US Military. *U.S. Army Medical Department Journal*, 22–32.

Anderson, M. K., Grier, T., Canham-Chervak, M., Bushman, T. T., & Jones, B. H. (2014). Risk Factors Associated with Higher Body Fat in US Army Female Soldiers. *U.S. Army Medical Department Journal*, 75–82.

Anderson, M. K., Grier, T., Canham-Chervak, M., Bushman, T. T., Nindl, B. C., & Jones, B. H. (2017). Effect of Mandatory Unit and Individual Physical Training on Fitness in Military Men and Women. *American Journal of Health Promotion*, *31*(5), 378–387. https://doi.org/10.1177/0890117116666977

Batchelor, J. (2008). *The Applicability of the Army Physical Fitness Test in the Contemporary Operating Environment* [U.S. Army Command and General Staff College]. https://apps.dtic.mil/dtic/tr/fulltext/u2/a483001.pdf

Bedard, A., Bigelman, K., & Nowels, R. (2019). The Army Combat Fitness Test—Redefining Physical Training for the United States Army. *NSCA TSAC Report*, *52*, 4.

Cain, P. A., Ahl, R., Hedstrom, E., Ugander, M., Allansdotter-Johnsson, A., Friberg, P., & Arheden, H. (2009). Age and gender specific normal values of left ventricular mass, volume and function for gradient echo magnetic resonance imaging: A cross sectional study. *BMC Medical Imaging*, *9*, 2. https://doi.org/10.1186/1471-2342-9-2

Campbell, J., Larson, R., Black, C., & McGuire, K. (2019). *Performance Recovery in Army Reserve Officer Training Corps Cadets Following a Bout of High-Intensity Exercise*. Medicine and Science in Sport and Exercise.

Canino, M. C., Cohen, B. S., Redmond, J. E., Sharp, M. A., Zambraski, E. J., & Foulis, S. A. (2018). The Relationship Between Soldier Performance on the Two-Mile Run and the 20-m Shuttle Run Test. *Military Medicine*, *183*(5–6), e182–e187. https://doi.org/10.1093/milmed/usx081

Charkoudian, N., & Joyner, M. J. (2004). Physiologic Considerations for Exercise Performance in Women. *Clinics in Chest Medicine*, *25*(2), 247–255. https://doi.org/10.1016/j.ccm.2004.01.001

Courtright, S. H., McCormick, B. W., Postlethwaite, B. E., & Reeves, C. J. (2013). A Meta-Analysis of Sex Differences in Physical Ability: Revised Estimates and Strategies for Reducing Differences in Selection Contexts. *Journal of Applied Psychology*, *98*(4), 623–641.

Draicchio, C., Martin, J. R., Fyock-Martin, M. B., & Merrigan, J. J. (2020). Retrospective Cohort Analysis of the Army Physical Fitness Test and the Occupational Physical Assessment Test in Reserve Officer Training Corps Cadets: A Brief Report. *Military Medicine*, *185*(7–8), e937–e943. https://doi.org/10.1093/milmed/usz489

Drinkwater, B. L. (1973). Physiological Responses of Women to Exercise. *Exercise and Sport Sciences Reviews*, *1*, 125–153.

Epstein, Y., Yanovich, R., Moran, D. S., & Heled, Y. (2013). Physiological Employment Standards IV: Integration of Women in Combat Units Physiological and Medical considerations. *European Journal Of Applied Physiology*, *113*(11), 2673–2690. https://doi.org/10.1007/s00421-012-2558-7

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fleg, J. L., O'Connor, F., Gerstenblith, G., Becker, L. C., Clulow, J., Schulman, S. P., & Lakatta, E. G. (1995). Impact of age on the cardiovascular response to dynamic upright exercise in healthy men and women. *Journal of Applied Physiology (Bethesda, Md.: 1985)*, *78*(3), 890–900. https://doi.org/10.1152/jappl.1995.78.3.890

Grier, T., Canham-Chervak, M., Anderson, M. K., Bushman, T. T., & Jones, B. H. (2015). The Effects of Cross-Training on Fitness and Injury in Women. *U.S. Army Medical Department Journal*, 33–41.

Grier, T., Canham-Chervak, M., Sharp, M., & Jones, B. H. (2015). Does Body Mass Index Misclassify Physically Active Young Men. *Preventative Medicine Reports*, *2*, 483–487. https://doi.org/10.1016/j.pmedr.2015.06.003

Headquarters Center for Initial Military Training. (2019). *Army Combat Fitness Test Initial Operation Capability*. https://www.army.mil/e2/downloads/rv7/acft/acft_ioc.pdf

Headquarters Department of the Army. (2019). *AR 600-9;The Army Body Composition Program*.

Headquarters Department of the Army. (2020). *FM 7-22: Holistic Health and Fitness*. https://armypubs.army.mil/ProductMaps/PubForm/Details.aspx?PUB_ID=1020968

Hicks, A. L., Kent-Braun, J., & Ditor, D. S. (2001). Sex Differences in Human Skeletal Muscle Fatigue. *Exercise and Sport Sciences Reviews*, *29*(3), 109–112. https://doi.org/10.1097/00003677-200107000-00004

Hydren, J. R., Borges, A. S., & Sharp, M. A. (2017). Systematic Review and Meta-Analysis of Predictors of Military Task Performance: Maximal Lift Capacity. *The Journal of Strength & Conditioning Research*, *31*(4), 1142–1164. https://doi.org/10.1519/JSC.0000000000001790

Jameson, J., Pappa, L., McGuire, B., & Kelly, K. R. (2015). Performance Differences Between Male and Female Marines on Standardized Physical Fitness Tests and Combat Proxy Tasks: Identifying the Gap. *U.S. Army Medical Department Journal*, 12–21.

Jeon, Y., Choi, J., Kim, H. J., Lee, H., Lim, J.-Y., Choi, S.-J., Jeon, Y., Choi, J., Kim, H. J., Lee, H., Lim, J.-Y., & Choi, S.-J. (2019). Sex- and Fiber-Type-Related Contractile Properties in Human Single Muscle Fiber. *Journal of Exercise Rehabilitation*, *15*(4), 537–545. https://doi.org/10.12965/jer.1938336.168

Johnson, D., Lynch, J., Nash, K., Cygan, J., & Mayhew, J. L. (2009). Relationship of Lat-Pull Repetitions and Pull-Ups to Maximal Lat-Pull and Pull-Up Strength in Men and Women. *The Journal of Strength & Conditioning Research*, *23*(3), 1022–1028. https://doi.org/10.1519/JSC.0b013e3181a2d7f5

Jones, M. T., Jagim, A. R., Haff, G. G., Carr, P. J., Martin, J., & Oliver, J. M. (2016). Greater Strength Drives Difference in Power between Sexes in the Conventional Deadlift Exercise. *Sports*, *4*(3), 43. https://doi.org/10.3390/sports4030043

Kazman, J. B., de la Motte, S., Bramhall, E. M. S., Purvis, D. L., & Deuster, P. A. (2015). Physical Fitness and Injury Reporting Among Active Duty and National Guard/Reserve

Women: Associations with Risk and Lifestyle Factors. *U.S. Army Medical Department Journal*, 49–57.

Kelly, K. R., & Jameson, J. T. (2016). Preparing for Combat Readiness for the Fight: Physical Performance Profile of Female US Marines. *The Journal of Strength & Conditioning Research*, *30*(3), 595–604. https://doi.org/10.1519/JSC.0000000000001269

Knapik, J. J., & East, W. B. (2014). History of United States Army Physical Fitness and Physical Readiness Training. *U.S. Army Medical Department Journal*, 5–19.

Knapik, J. J., Rieger, W., Palkoska, F., Van Camp, S., & Darakjy, S. (2009). United States Army Physical Readiness Training: Rationale and Evaluation of the Physical Training Doctrine. *The Journal Of Strength & Conditioning Research*, *23*(4), 1353–1362. https://doi.org/10.1519/JSC.0b013e318194df72

Knapik, J. J., Sharp, M. A., Canham-Chervak, M., Hauret, K., Patton, J. F., & Jones, B. H. (2001). Risk Factors for Training-Related Injuries Among Men and Women in Basic Combat Training. *Medicine & Science in Sports & Exercise*, *33*(6), 946–954.

Kraemer, W. j., Mazzetti, S. a., Nindl, B. c., Gotshalk, L. a., Volek, J. s., Bush, J. a., Marx, J. o., Dohi, K., Gomez, A. l., Miles, M., Fleck, S. j., Newton, R. u., & Hakkinen, K. (2001). Effect of Resistance Training on Women's Strength/Power and Occupational Performances. *Medicine & Science in Sports & Exercise*, *33*(6), 1011–1025.

Lockie, R. G., Dawes, J. J., Orr, R. M., & Dulla, J. M. (2020). Recruit Fitness Standards From a Large Law Enforcement Agency: Between-Class Comparisons, Percentile Rankings, and Implications for Physical Training. *The Journal of Strength & Conditioning Research*, *34*(4), 934–941. https://doi.org/10.1519/JSC.0000000000003534

Lowman, C. (2010). *Does Current Army Physical Fitness Training Doctrine Adequately Prepare Soldiers for War?* U.S. Army Command and General Staff College.

Malina, R. M., & Bouchard, C. (1991). *Growth, Maturation, and Physical Activity* (pp. xiii, 501). Human Kinetics Academic.

Marshall, W. A., & Tanner, J. M. (1969). Variations in Pattern of Pubertal Changes in Girls. *Archives of Disease in Childhood*, *44*(235), 291–303.

Marshall, W. A., & Tanner, J. M. (1970). Variations in the Pattern of Pubertal Changes in Boys. *Archives of Disease in Childhood*, *45*(239), 13–23.

Maud, P. J., & Shultz, B. B. (1986). Gender Comparisons in Anaerobic Power and Anaerobic Capacity Tests. *British Journal of Sports Medicine*, *20*(2), 51–54.

Maughan, R. J., Harmon, M., Leiper, J. B., Sale, D., & Delman, A. (1986). Endurance Capacity of Untrained Males and Females in Isometric and Dynamic Muscular Contractions. *European Journal of Applied Physiology and Occupational Physiology*, *55*(4), 395–400. https://doi.org/10.1007/BF00422739

Mello, R. P., Murphy, M. M., & Vogel, J. A. (1987). *Relationship between a Two Mile Run and Maximal Oxygen Uptake:* Defense Technical Information Center. https://doi.org/10.21236/ADA179343

Miller, A. E., MacDougall, J. D., Tarnopolsky, M. A., & Sale, D. G. (1993). Gender Differences in Strength and Muscle Fiber Characteristics. *European Journal of Applied Physiology and Occupational Physiology*, *66*(3), 254–262. https://doi.org/10.1007/BF00235103

Morgan, B., Woodruff, S. J., & Tiidus, P. M. (2003). Aerobic Energy Expenditure During Recreational Weight Training in Females and Males. *Journal of Sports Science & Medicine*, *2*(3), 117–122.

Murphy, M. M., Patton, J. F., & Frederick, F. A. (1986). Comparative Anaerobic Power of Men and Women. *Aviation, Space, and Environmental Medicine*, *57*(7), 636–641.

Negrete, R. J., Hanney, W. J., Pabian, P., & Kolber, M. J. (2013). Upper Body Push to Pull Strength Ratio in Recreationally Active Adults. *International Journal of Sports Physical Therapy*, *8*(2), 138–144.

Nikolaidis, P. (2014). Relationship of Body Mass Index with 1600 m Running, 50 m Swimming and Pull-ups Performance in Army Cadets. *SaudiJournal of Sports Medicine*.

Nindl, B. C., Alvar, B. A., R Dudley, J., Favre, M. W., Martin, G. J., Sharp, M. A., Warr, B. J., Stephenson, M. D., & Kraemer, W. J. (2015). Executive Summary From the National Strength and Conditioning Association's Second Blue Ribbon Panel on Military Physical Readiness: Military Physical Performance Testing. *The Journal of Strength & Conditioning Research*, *29 Suppl 11*, S216-220. https://doi.org/10.1519/JSC.0000000000001037

Nindl, B. C., Jones, B. H., Van Arsdale, S. J., Kellly, K., & Kraemer, W. J. (2015). Physical Training Strategies for Military Women's Performance Optimization in Combat-Centric Occupations. *The Journal Of Strength & Conditioning Research*, *29 Suppl 11*, S101–S106. https://doi.org/10.1519/JSC.0000000000001089

Nindl, B. C., Jones, B. H., Van Arsdale, S. J., Kelly, K., & Kraemer, W. J. (2016). Operational Physical Performance and Fitness in Military Women: Physiological, Musculoskeletal Injury, and Optimized Physical Training Considerations for Successfully Integrating Women Into Combat-Centric Military Occupations. *Military Medicine*, *181*(1 Suppl), 50–62. https://doi.org/10.7205/MILMED-D-15-00382

Pierce, J., Hauret, K., Alemany, J., Grier, T., Sharp, M., Redmond, J., Foulis, S., Cohen, B., Canino, M., & Jones, B. (2018). *Physical Performance on the Occupational Physical Assessment Test (OPAT), Army Physical Fitness Test (APFT), and Relationship to Body Mass Index During Initial Entry Training—OPAT Phase 1 | Request PDF* (Technical Report S.0047229-18a). U.S. Army Public Health Center.

Pierce, J. R., DeGroot, D. W., Grier, T. L., Hauret, K. G., Nindl, B. C., East, W. B., McGurk, M. S., & Jones, B. H. (2017). Body Mass Index Predicts Selected Physical Fitness Attributes but is not Associated with Performance on Military Relevant Tasks in U.S. Army Soldiers. *Journal of Science and Medicine in Sport*, *20 Suppl 4*, S79–S84. https://doi.org/10.1016/j.jsams.2017.08.021

Porto, L. G. G., Nogueira, R. M., Nogueira, E. C., Molina, G. E., Farioli, A., Junqueira, L. F., & Kales, S. N. (2016). Agreement Between BMI and Body Fat Obesity Definitions in a Physically Active Population. *Archives of Endocrinology and Metabolism*, *60*(6), 515–525. https://doi.org/10.1590/2359-3997000000220

Poser, W. M., Trautman, K. A., Dicks, N. D., Christensen, B. K., Lyman, K. J., & Hackney, K. J. (2019). Simulated Casualty Evacuation Performance Is Augmented by Deadlift Peak Force. *Military Medicine*, *184*(9–10), e406–e411. https://doi.org/10.1093/milmed/usz050

Reilly, T. J., Sharp, M. A., Cao, M., & Canino, M. C. (2019). A Database of Predictor Test Sex Bias for Development of Military Physical Employment Standards. *Work (Reading, Mass.)*, *63*(4), 591–601. https://doi.org/10.3233/WOR-192956

Rogol, A. D., Roemmich, J. N., & Clark, P. A. (2002). Growth at Puberty. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *31*(6 Suppl), 192–200. https://doi.org/10.1016/s1054-139x(02)00485-8

Saad, M. F., Damani, S., Gingerich, R. L., Riad-Gabriel, M. G., Khan, A., Boyadjian, R., Jinagouda, S. D., el-Tawil, K., Rude, R. K., & Kamdar, V. (1997). Sexual Dimorphism in Plasma Leptin Concentration. *The Journal of Clinical Endocrinology and Metabolism*, *82*(2), 579–584. https://doi.org/10.1210/jcem.82.2.3739

Sauers, S. E., & Scofield, D. E. (2014). Strength and Conditioning Strategies for Females in the Military. *Strength & Conditioning Journal*, *36*(3), 1–7.

Sharp, M. A. (1993). *Physical Fitness, Physical Training and Occupational Performance of Men and Women in the U.S. Army: A Review of Literature:* Defense Technical Information Center. https://doi.org/10.21236/ADA266297

Staron, R. S., Karapondo, D. L., Kraemer, W. J., Fry, A. C., Gordon, S. E., Falkel, J. E., Hagerman, F. C., & Hikida, R. S. (1994). Skeletal Muscle Adaptations During Early Phase of Heavy-Resistance Training in Men and Women. *Journal of Applied Physiology*, *76*(3), 1247–1255. https://doi.org/10.1152/jappl.1994.76.3.1247

Steed, C. L., Krull, B. R., Morgan, A. L., Tucker, R. M., & Ludy, M.-J. (2016). Relationship Between Body Fat and Physical Fitness in Army ROTC Cadets. *Military Medicine*, *181*(9), 1007–1012. https://doi.org/10.7205/MILMED-D-15-00425

Stockbrugger, B. A., & Haennel, R. G. (2001). Validity and Reliability of a Medicine Ball Explosive Power Test. *The Journal Of Strength & Conditioning Research*, *15*(4), 431–438.

St-Onge, M.-P. (2010). Are Normal-Weight Americans Over-Fat? *Obesity (Silver Spring, Md.)*, *18*(11). https://doi.org/10.1038/oby.2010.103

Szivak, T. K., Kraemer, W. J., Nindl, B. C., Gotshalk, L. A., Volek, J. S., Gomez, A. L., Dunn-Lewis, C., Looney, D. P., Comstock, B. A., Hooper, D. R., Flanagan, S. D., & Maresh, C.

M. (2014). Relationships of Physical Performance Tests to Military-Relevant Tasks in Women. *U.S. Army Medical Department Journal*, 20–26.

Tanner, J. M. (1965). The Relationship of Puberty to Other Maturity Indicators and Body Composition in Man. *Symp. Soc. Stud. Hum. Bio.*, *6*(211).

Wells, C. L. (1991). *Women, Sport and Performance: A Physiological Perspective* (2nd ed.). Human Kinetics. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1479015/

Wells, C. L., & Plowman, S. A. (1983). Sexual Differences in Athletic Performance: Biological or Behavioral? *The Physician and Sportsmedicine*, *11*(8), 52–63. https://doi.org/10.1080/00913847.1983.11708602

Wong, S. L., DePaoli, A. M., Lee, J. H., & Mantzoros, C. S. (2004). Leptin Hormonal Kinetics in the Fed State: Effects of Adiposity, Age, and Gender on Endogenous Leptin Production and Clearance Rates. *The Journal of Clinical Endocrinology and Metabolism*, *89*(6), 2672–2677. https://doi.org/10.1210/jc.2003-031931

Yanovich, R., Evans, R., Israeli, E., Constantini, N., Sharvit, N., Merkel, D., Epstein, Y., & Moran, D. S. (2008). Differences in Physical Fitness of Male and Female Recruits in Gender-Integrated Army Basic Training. *Medicine & Science in Sports & Exercise*, *40*(11), S654. https://doi.org/10.1249/MSS.0b013e3181893f30

Appendix A – ACFT 2.0 Scoring Scale

Appendix 3 to Annex A, HQDA EXORD 219-18

Army ACFT FY20 Standards (As of 1 Oct 19)

Points	MDL	SPT	HRP	SDC	LTK	2MR	
100	340	12.5	60	1:33	20	13:30	
99		12.4	59	1:36		13:39	
98		12.2	58	1:39	19	13:48	
97	330	12.1	57	1:41		13:57	
96		11.9	56	1:43	18	14:06	
95		11.8	55	1:45		14:15	
94	320	11.6	54	1:46	17	14:24	
93		11.5	53	1:47		14:33	
92	310	11.3	52	1:48	16	14:42	
91		11.2	51	1:49		14:51	
90	300	11.0	50	1:50	15	15:00	
89		10.9	49	1:51		15:09	
88	290	10.7	48	1:52	14	15:18	
87		10.6	47	1:53		15:27	
86	280	10.4	46	1:54	13	15:36	
85		10.3	45	1:55		15:45	
84	270	10.1	44	1:56	12	15:54	
83		10.0	43	1:57		16:03	
82	260	9.8	42	1:58	11	16:12	
81		9.7	41	1:59		16:21	
80	250	9.5	40	2:00	10	16:30	
79		9.4	39	2:01		16:39	
78	240	9.2	38	2:02	9	16:48	
77		9.1	37	2:03		16:57	
76	230	8.9	36	2:04	8	17:06	
75		8.8	35	2:05		17:15	
74	220	8.6	34	2:06	7	17:24	
73		8.5	33	2:07		17:33	
72	210	8.3	32	2:08	6	17:42	
71		8.2	31	2:09		17:51	
70	200	8.0	30	2:10	5	18:00	HVY
69		7.8	28	2:14		18:12	
68	190	7.5	26	2:18	4	18:24	
67		7.1	24	2:22		18:36	
66		6.8	22	2:26		18:48	
65	180	6.5	20	2:30	3	19:00	SIG
64	170	6.2	18	2:35		19:24	
63	160	5.8	16	2:40		19:48	
62	150	5.4	14	2:45	2	20:12	
61		4.9	12	2:50		20:36	
60	140	4.5	10	3:00	1	21:00	MOD
59				3:01		21:01	
58				3:02		21:03	
57				3:03		21:05	
56				3:04		21:07	
55		4.4	9	3:05		21:09	
54				3:06		21:10	
53				3:07		21:12	
52				3:08		21:14	
51				3:09		21:16	
50	130	4.3	8	3:10		21:18	