GLASS, STEPHEN MICHAEL, Ph.D. Strength, Balance, and Range of Motion as Modifiers of the Effect of External Load-Bearing on Functional Movement Behaviors in the Military Recruitment Population. (2015)
Directed by Dr. Scott E. Ross. 171 pp.

Clinical movement screens have become increasingly popular in strength and conditioning programs designed for the tactical athlete. Whereas conventionally administered movement screens are largely not sensitive to behaviors which impact soldier-relevant physical performance, recent evidence suggests that modified screens which incorporate external load-bearing strengthen the relationship between movement behaviors and performance outcomes. It remains unclear, however, which mechanisms may account for this improvement in association. Physical performance is considered a multidimensional construct influenced by several independent factors. Among the factors which influence military physical performance, movement screens may require high levels of strength, balance, and range of motion. This project used penalized interaction models to determine the role of strength, balance, and range of motion in modifying the effects of external load bearing on movement quality and movement. Additional confirmatory analyses examined differences in the abilities of FMS item scores to predict physical performance outcomes when those scores were obtained during control vs. external load-bearing conditions. Results suggest that the effect of load on movement complexity is modified by strength, balance, and range of motion whereas the effect on clinically rated movement quality is modified by only balance and range of motion. While the direction of the observed effects did not always coincide with our hypotheses, the present findings mirror those of previous research with respect to differential validity of weighted vs. control FMS item scores in predicting criterion performance measures.

STRENGTH, BALANCE, AND RANGE OF MOTION AS MODIFIERS OF THE EFFECT OF EXTERNAL LOAD-BEARING ON FUNCTIONAL MOVEMENT BEHAVIORS IN THE MILITARY RECRUITMENT POPULATION

by

Stephen Michael Glass

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2015

Approved by

____
Committee Chair

APPROVAL PAGE

This dissertation written by Stephen Michael Glass has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________
Scott E. Ross

Committee Members ______________________
Randy J. Schmitz

____
Christopher K. Rhea

____
Xiaoli Gao

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Predicting physical performance has historically been an elusive research goal. One practice which may hold promise in this area is clinical assessment of movement quality, which may be defined as how closely an individual approximates normative patterns of movement. Movement quality screens have recently gained traction in performance and rehabilitation circles as a cost-effective, convenient method of identifying biomechanical abnormalities or limitations. Such screens have been adopted in a number of military settings for purposes of determining injury risk and guiding training interventions, in addition to predicting physical performance.[1–5] There is currently limited evidence to suggest an association between clinically rated movement quality and physical performance outcomes. Further, while research has shown that movement quality is responsive to intervention programs, the performance implications of movement-based training are unclear. As part of the ongoing Human Performance Optimization (HPO) initiative,[6] the military is currently seeking ways to promote soldier performance capabilities despite recent cuts in defense spending. Field-expedient screens may help address HPO objectives; however, the lack of evidence and validation to support their use is a major logistical roadblock in this area.[7] In order to maximize the benefit derived from these tools, it is essential that we understand the relationship between clinically rated movement quality and physical performance outcomes. Additionally, we must establish an appropriate framework for interpreting these tests.

An important consideration in this regard is the theory which underlies the practice of movement screening. Many theories of motor performance attempt to identify an objectively optimal technique by which movement can be judged. Widely used movement screens are based on a similar approach, evaluating what their designers consider to be primitive and fundamental movement patterns.[8–10] These patterns are theorized to be generic, ubiquitous behaviors which are encoded in the central nervous system as motor programs and are considered to form the basis for optimal function in more complex neuromotor behaviors.[11–14] This position carries strong implications for evaluating and training movement quality which may lead to inappropriate practices. As an example, several intervention studies have been published in recent years which used movement quality scores as their primary outcome.[15] If optimal movement behaviors differ between individuals, applying movement quality indices in this way could be problematic.

An alternative theory which is becoming increasingly influential in the modeling of human physical performance suggests that behaviors arise dynamically from interacting constraints on the movement system.[16,17] From this point of view, the search for an optimal technique is misguided as any movement behavior can be considered an optimal solution given the comprehensive set of constraints limiting performance of the system at a given time. Rather than attempting to define optimal performance technique, the more relevant discussion may therefore revolve around the degree to which the movement system is constrained. This could enable us to develop more refined screening practices as well as training interventions designed to address the factors limiting performance.

Evidence of constraints impinging on dynamic human systems can be observed in the variability patterns of biological signals. Constraints on human movement behaviors exist on many scales and evolve continuously over time,[16] effectively imposing limits on the range of movement patterns that can be applied in solution of a given motor task. By analyzing the dynamic structure of variability in movement systems, we can infer the relative degree to which self-organization of behaviors is limited. Analysis of movement variability has been applied extensively in the assessment of aging and pathology.[18,19] Constraints on dynamic variability in movement behaviors are believed to increase the risk of injury, both acute and overuse,[20] and have recently received focus as a potential target for performance enhancement interventions.[21]

To the extent that these constraints are similarly visible through cost effective and field-expedient tests, such approaches can be quite useful. It is crucial, however, that the purpose of a clinical screen in this context be distinguished. Their utility is not derived from indicating proximity to a primal movement norm, but instead the ability to highlight characteristics of an individual that may be relevant to an outcome of interest. Of the traits assessed by the movement quality scales,[8,22,23] strength, balance, and range of motion also impact physical performance. Performance limitations related to deficits in these areas could therefore be identified using lightweight, field-expedient metrics. Despite this apparent overlap, early work on the topic of movement quality has largely been unable to identify a consistent association with performance outcomes.[24–27]

We conducted a preliminary small-n pilot study to investigate the relationship between the Functional Movement Screen™ (FMS), a widely used clinical movement quality assessment, and laboratory measures of postural control.[28] An unexpected finding was that subjects exhibiting different degrees of static and dynamic postural control, and different athletic profiles (mesomorph with BMI = 23, endomorph with BMI = 32), were able to achieve similarly high scores on the clinical assessment. We conjectured that the screen may suffer from a ceiling effect which limits its ability to discriminate performance levels as composite scores reached the 21-point upper limit. We subsequently completed a group study designed to impose a greater challenge on the movement system with the expectation that doing so would eliminate ceiling effects and increase the screen's sensitivity to behavioral differences which impact physical performance. Because our goal was to improve performance prediction in the military recruitment population, we administered the screen while subjects wore a standardized soldier load (18.1 kg) and used the results to predict performance in a battery of military-relevant physical tests. Our results indicated that FMS scores taken from the loaded condition were better predictors of physical performance outcomes than were scores from a control condition.[29] It is unclear, however, what mechanisms may have accounted for this improvement in performance prediction.

It stands to reason that the factors identified above—strength, balance, and range of motion—played a role in mediating the increase in the predictive validity of FMS scores observed in our latter study. Both strength and balance have been shown to correlate with performance outcomes comparable to those used in our investigation.[30,31] Similarly, athletes with range of motion restrictions may be forced to employ coordination strategies which are biased in favor of relatively inefficient joint motions. Because ROM, balance, and force output are adversely impacted by load carriage,[32–36] the experimental condition used in our preliminary study may have highlighted performance-relevant deficits in these areas. However, since these covariates were not observed concurrently, this cannot be stated conclusively. Including sufficient control measures, such as peak power, force plate measures of postural control, and clinical range of motion tests, may enable us to determine more precisely the extent to which deficits in each of these areas act to constrain movement quality and performance on military physical tests.

Simply identifying these constraints does not, however, provide conclusive support for the notion that movement quality should not be pursued in reference to an objective norm. It is important, therefore, that these inquiries be complemented with evidence of the role of movement variability in tasks commonly used to assess movement quality. Together, this information will establish a mechanistic understanding of the relationship between clinical ratings of movement behavior and physical performance while appropriately directing attention toward the underlying constraints in the context of intervention. Therefore, the purposes of this dissertation project were: 1) Determine the extent to which balance, strength, and range of motion protect against decreases in quality of FMS movement behaviors associated with standardized external loading. 2) Determine the extent to which balance, strength, and range of motion protect against decreased complexity of movement associated with standardized external loading in dynamic postural tasks. 3) Determine the extent to which soldier-relevant physical performance outcomes are predicted by FMS item scores obtained during a standardized external loading condition in comparison with item scores from a conventionally administered FMS.

1.1 Statement of Problem

Compared to conventional FMS testing, FMS testing with a standardized external load is associated with improved prediction of soldier-relevant physical performance outcomes; however, the mechanisms which account for this improvement in prediction are unclear.

1.2 Objectives and Hypotheses

(1) Determine the extent to which balance, strength, and range of motion protect against decreases in quality of FMS movement behaviors associated with standardized external loading. Hypotheses

- Decreases in FMS item scores related to external load-bearing will be more pronounced in individuals with high center of pressure resultant velocity.
- Decreases in FMS item scores related to external load-bearing will be more pronounced in individuals with low countermovement jump peak power and YMCA bench press tests repetitions.
- Decreases in FMS item scores related to external load-bearing will be more pronounced in individuals with low range of motion in the sit-and-reach, weight-bearing lunge, and Apley scratch tests.

(2) Determine the extent to which balance, strength, and range of motion protect against decreased complexity of movement associated with standardized external loading in dynamic postural tasks. Hypotheses

- During cyclic performance of the Deep Squat, Hurdle Step, and Inline Lunge, decreases in sample entropy related to external load-bearing will be greater in individuals with high resultant center of pressure velocity.
- During cyclic performance of the Deep Squat, Hurdle Step, and Inline Lunge, decreases in sample entropy related to external load-bearing will be greater in individuals with low countermovement jump peak power and YMCA bench press test .
- During cyclic performance of the Deep Squat, Hurdle Step, and Inline Lunge, decreases in sample entropy related to external load-bearing will be greater in individuals with lower range of motion in the sit-and-reach, weight-bearing lunge, and Apley scratch tests.

(3) Determine the extent to which soldier-relevant physical performance outcomes are predicted by FMS item scores obtained during a standardized external loading condition in comparison with item scores from a conventionally administered FMS. Hypotheses

- FMS item scores from the standardized external load condition will be associated with shorter 27.4 meter sprint times.
- FMS item scores from the standardized external load condition will be associated with shorter 400 meter run times.
- FMS item scores from the standardized external load condition will be associated with shorter completion times in the Mobility for Battle Assessment.

- FMS item scores from the standardized external load condition will be associated with shorter completion times in the Partner Rescue Drag task.
- FMS item scores from the standardized external load condition will be associated with shorter agility T-test times.
- With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter 27.4 meter sprint times.
- With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter 400 meter run times.
- With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter completion times in the Mobility for Battle Assessment.
- With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter completion times in the Partner Rescue Drag task.
- With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter agility T-test times

1.3 Limitations and Assumptions

- The findings of this project are limited to the population under investigation—healthy, college-age males and females.

- The equipment and procedures used for data collection are assumed to be sufficient to provide an accurate representation of movement variables. These include the AMTI Accusway Force Plate (AMTI Inc., Watertown, MA), Bertec Force Plate model 4060-NC (Bertec Corporation, Columbus, OH), and Brower Timing Gates (Brower Timing Systems, Draper, UT).
- Repeated FMS testing is assumed to be associated with minimal learning/fatigue effects and any related bias can be distributed across conditions through the use of counterbalanced trials.
- The effects of fatigue are further assumed to be negligible during the dynamic postural control tasks regardless of load condition.
- The performance outcomes assessed are adapted from previous scientific work and may be limited in their relation to military performance outside the laboratory.
- FMS scoring will be recorded in real-time by an experienced investigator with established reliability.

1.4 Delimitations

- Participation will be limited to healthy 18-34 year old men and women free from injury for at least 6 months prior to data collection.
- The observed relationships between FMS movement tasks and performance outcomes will be specific to the tests administered in the study.
- Timed tests of physical performance will be administered without external load.

- In order to standardize the tasks, optional modifications of FMS testing procedures will not be permitted.

1.5 Operational Definitions

- Quiet Single Leg Stance: A postural control task in which the subject is free from perturbing stimulus and is instructed to remain as motionless as possible on one foot. This task will require subjects to be barefoot with eyes closed and hands placed on hips.

- Dynamic Postural Control: Postural control incorporating movement demands. In order to distinguish between this concept and nonlinear system dynamics, the terms “dynamic,” “dynamics,” and “dynamical” will always refer to the latter unless included in the phrase “dynamic postural control.”

- Penalized Regression: Regression techniques making use of data-driven penalty parameters applied to coefficients at both the between-group and within-group levels. Penalized regression is particularly suited for model selection and prediction in high dimensional applications. Specialized penalty methods are used to accommodate predictor variables with ordinal level data.

- Young, Healthy Adults: 18-30 year old males and females with no recent history of injury (at least 6 months) who are medically fit to participate in vigorous physical activity.

1.6 Variables

1.6.1 Objective 1 Independent Variables

- **Categorical/Binary**
  - Condition: Wearing a weight vest or not wearing a weight vest (control).
- **Continuous**
  - Resultant Center of Pressure Velocity: During a single leg standing task with eyes closed, this quantity is the resultant (x, y) velocity of the point location of the ground reaction force vectors in a 2-dimensional Cartesian coordinate system
  - Countermovement Jump Peak Power: The maximum vertical power produced during the concentric phase of a countermovement jump task.[37]
  - Predicted 1-Repetition Maximum Bench Press: The predicted maximal load that can be lifted for one successful bench press repetition based on the modified YMCA bench press test prediction equations in Kim.[38]
  - Ankle dorsiflexion ROM: Maximal dorsiflexion range of motion measured as the distance from the great toe to the wall in the Weight-Bearing Lunge Test.[39]
  - Flexion ROM: Hip and trunk flexion range of motion measured as the forward reach distance achieved during the sit-and-reach Test.[40]
  - Shoulder ROM: The shortest distance measured between closed fists behind the back as in a modified version of the Apley Scratch Test.[41]

1.6.2 Objective 1 Dependent Variables

- **Ordinal**
  - FMS Item Scores:[8,9] For each item below, scores are assigned according to the following criteria: 0) Subject experienced pain during any portion of the movement; 1) Subject is unable to complete the movement; 2) Subject is able to complete the task with errors noted; 3) Subject is able to complete the task without error.
    * Deep Squat (DS): First FMS test item. Subjects begin with feet shoulder width apart and arms holding a dowel pressed overhead. When cued, subjects squat as deeply as possible while attempting to keep the spine straight and then return to the starting position.
    * Hurdle Step (HS): Second FMS test item. Subjects rest the dowel across their shoulders behind the head. When cued, they raise one leg over an obstacle placed at the height of the tibial tuberosity, touch their heel on the opposite side of the obstacle, and return to the starting position.
    * Inline Lunge (ILL): Third FMS test item. Subjects hold the dowel in place vertically behind their backs and stand with feet inline on a 2x6" board with a distance equal to the height of the tibial tuberosity separating the toe of the back foot from the heel of the front foot. When cued, subjects drop down into a lunge position and lightly touch the back knee to the board and then return to the starting position.

* Shoulder Mobility (SM): Fourth FMS test item. Subjects are instructed to make fists with their thumbs on the inside. When cued, they attempt to touch their fists together behind their backs by reaching overhead/down the back on one side and up the back on the other side.
* Active Straight Leg Raise (ASLR): Fifth FMS test item. Subjects lie on their backs with hips and knees fully extended. Maintaining a straight knee in both legs, they flex one hip as much as possible, hold briefly, and return to the starting position.
* Trunk Stability Push Up (TSPU): Sixth FMS test item. This test requires subjects to perform a push up while maintaining a rigid torso and legs (plank) with hands placed in a position slightly more superior than that of a conventional push up. For males, hands are placed on the ground such that a line connecting the thumbs would cross the middle of the forehead when in the face-down position. For females, hands are placed such that a line connecting the thumbs would cross the chin.
* Quadruped Rotary Stability (RS): Seventh FMS test item. Subjects are position "on all fours," i.e. with hands and knees on the ground. A 2x6" board is placed length-wise between the hands, knees, and feet, all of which must be in contact with the board on both sides. When cued, subjects attempt to reach forward with one hand while reaching backward with the ipsilateral foot, then touch the knee to the elbow, then reach out a second time, and finally return to the

starting position all while avoiding any contact with the floor on the working side.

1.6.3 Objective 2 Independent Variables

- **Categorical/Binary**
  - Condition: Wearing a weight vest or not wearing a weight vest (control).
- **Continuous**
  - Resultant Center of Pressure Velocity: During a single leg standing task with eyes closed, this quantity is the resultant (x, y) velocity of the point location of the ground reaction force vectors in a 2-dimensional Cartesian coordinate system
  - Countermovement Jump Peak Power: The maximum vertical power produced during the concentric phase of a countermovement jump task.[37]
  - Predicted 1-Repetition Maximum Bench Press: The predicted maximal load that can be lifted for one successful bench press repetition based on the modified YMCA bench press test prediction equations in Kim.[38]
  - Ankle dorsiflexion ROM: Maximal dorsiflexion range of motion measured as the distance from the great toe to the wall in the Weight-Bearing Lunge Test.[39]
  - Flexion ROM: Hip and trunk flexion range of motion measured as the forward reach distance achieved during the sit-and-reach Test.[40]
  - Shoulder ROM: The shortest distance measured between closed fists behind the back as in a modified version of the Apley Scratch Test.[41]

1.6.4 Objective 2 Dependent Variables

- **Continuous**
  - Multivariate Multiscale Sample Entropy: Sample entropy calculated on the resultant center of pressure displacement time series during the following tasks
  - Cyclic Deep Squat (DS): First FMS test item. Subjects begin with feet shoulder width apart and arms holding a dowel pressed overhead. When cued, subjects squat as deeply as possible while attempting to keep the spine straight and then return to the starting position. This version will be performed cyclically for 5 repetitions at a comfortable, self-selected pace.
  - Cyclic Hurdle Step (HS): Second FMS test item. Subjects rest the dowel across their shoulders behind the head. When cued, they raise one leg over an obstacle placed at the height of the tibial tuberosity, touch their heel on the opposite side of the obstacle, and return to the starting position. This version will be performed cyclically for 5 repetitions at a comfortable, self-selected pace.
  - Cyclic Inline Lunge (ILL): Third FMS test item. Subjects hold the dowel in place vertically behind their backs and stand with feet inline with a distance equal to the height of the tibial tuberosity separating the toe of the back foot from the heel of the front foot. When cued, subjects drop down into a lunge position and lightly touch the back knee to the board

and then return to the starting position. This version will be performed cyclically for 5 repetitions at a comfortable, self-selected pace.

1.6.5 Objective 3 Independent Variables

- **Ordinal**
  - FMS Item Scores:[8,9] For each item below, scores are assigned according to the following criteria: 0) Subject experienced pain during any portion of the movement; 1) Subject is unable to complete the movement; 2) Subject is able to complete the task with errors noted; 3) Subject is able to complete the task without error.
    * Deep Squat (DS): First FMS test item. Subjects begin with feet shoulder width apart and arms holding a dowel pressed overhead. When cued, subjects squat as deeply as possible while attempting to keep the spine straight and then return to the starting position.
    * Hurdle Step (HS): Second FMS test item. Subjects rest the dowel across their shoulders behind the head. When cued, they raise one leg over an obstacle placed at the height of the tibial tuberosity, touch their heel on the opposite side of the obstacle, and return to the starting position.
    * Inline Lunge (ILL): Third FMS test item. Subjects hold the dowel in place vertically behind their backs and stand with feet inline on a 2x6” board with a distance equal to the height of the tibial tuberosity separating the toe of the back foot from the heel of the front foot. When cued, subjects drop down into a lunge position and

lightly touch the back knee to the board and then return to the starting position.

∗ Shoulder Mobility (SM): Fourth FMS test item. Subjects are instructed to make fists with their thumbs on the inside. When cued, they attempt to touch their fists together behind their backs by reaching overhead/down the back on one side and up the back on the other side.

∗ Active Straight Leg Raise (ASLR): Fifth FMS test item. Subjects lie on their backs with hips and knees fully extended. Maintaining a straight knee in both legs, they flex one hip as much as possible, hold briefly, and return to the starting position.

∗ Trunk Stability Push Up (TSPU): Sixth FMS test item. This test requires subjects to perform a push up while maintaining a rigid torso and legs (plank) with hands placed in a position slightly more superior than that of a conventional push up. For males, hands are placed on the ground such that a line connecting the thumbs would cross the middle of the forehead when in the face-down position. For females, hands are placed such that a line connecting the thumbs would cross the chin.

∗ Quadruped Rotary Stability (RS): Seventh FMS test item. Subjects are position "on all fours," i.e. with hands and knees on the ground. A 2x6" board is placed length-wise between the hands, knees, and feet, all of which must be in contact with the board on both sides. When cued, subjects attempt to reach forward with one hand while reaching backward with the ipsilateral foot, then touch the knee to the elbow, then reach out a second time, and finally return to the starting position all while avoiding any contact with the floor on the working side.

1.6.6 Objective 3 Dependent Variables

- **Continuous**
  - 27.43 meter (30 yard) sprint: This task will be included as a test of speed.[42] Average speed will be recorded over five trials. Subjects begin with their foot depressing a start-on-release trigger mechanism and sprint 30 yards as quickly as possible. Final time is recorded by an infrared timing gate.
  - 400 meter run: This task measures short-duration aerobic/anaerobic endurance.[42] Subjects are given one trial. Beginning on the start-on-release trigger, subjects run 4.5 laps around the Coleman Gym as quickly as possible. Completion time is recorded by infrared timing gate.
  - Mobility for Battle Assessment: This task was developed to assess agility and mobility required for combat.[43] It incorporates shuttle runs, bear crawls, broad jumps, pushups, ammunition carries, and core strength work into a single, timed trial. Subjects begin on the start-on-release trigger and proceed through ordered stations which are marked by cones. Final time is recorded by timing gate.

- Partner Rescue Drag: This task is designed to simulate rescuing an injured partner. A load of 150 lbs. is dragged a distance of 100 ft. across the Coleman Gym.[42]
- T-test Agility Time: Beginning behind a line, subjects sprint forward 10 yards, shuffle left 5 yards, shuffle right 10 yards, shuffle left 5 yards, and backpedal 10 yards. Finish time is recorded when the subject crosses the start/finish line.[42]

CHAPTER II

REVIEW OF THE LITERATURE

This review will briefly discuss the priorities of the Human Performance Optimization (HPO) initiative along with the role of movement quality screening in supporting those efforts. The application of the Functional Movement Screen™ in military and non-military populations will then be addressed. Successes and limitations of the FMS as a predictor of physical performance, and an instrument for program design, will be reviewed. Finally, I will discuss the concepts of nonlinear dynamics and constrained optimization in the context of human physical performance and provide rationale for the dependent measures proposed for this investigation.

2.1 Total Force Fitness & Human Performance Optimization

The U.S. military has invested considerable time and resources into improving the performance of its most valuable asset, the warfighter. Following a 1998 Government Accountability Office[44] report encouraging the Department of Defense to study the factors underlying the military's high attrition rates, human performance scientists began to investigate the problem from several angles. According to this report, performance-related failure consistently accounted for a large proportion, if not a majority, of basic training attrition. The association between substandard physical fitness and attrition over varying time scales sparked interest in screening programs which could be implemented prior to enlistment.[45] A pilot program—the Assessment of Recruit Motivation and Strength—was implemented in this vein at Military Entrance Processing Stations (MEPS) around the country and was shown prospectively to reduce attrition rates at very low cost.[46–48] Despite its advantages, DOD terminated the program in 2009 as the weakening economy had resulted in a stronger recruiting environment.[49]

Researchers have also focused on methods to prevent attrition and improve training for those already enlisted. The Human Performance Optimization (HPO) initiative was formalized in the 2007[6] and specifically seeks solutions to maximize physical performance among soldiers. These efforts have led to a number of projects designed to address the performance requirements of the modern soldier with an applied focus. Examples programs include Ranger Athlete Warrior, the Tactical Athlete Program, the Eagle Tactical Athlete Program, the Mountain Athlete Program, Military Performance Power & Prevention, NSCA's Tactical Strength and Conditioning course, and the Army's Tactical Human Optimization, Rapid Rehabilitation and Reconditioning program.[50–53] Pilot programs which have been established to promote HPO objectives incorporate functional movement evaluation tools which can be used to rate physical ability, classify injury risk, track training progress, and assist with return to duty decisions. The Functional Movement Screen™ is arguably the most popular and its adoption in all branches was officially recommended in a 2011 Directive of the Joint Chiefs of Staff outlining the "Total Force Fitness Framework."[54]

The most recent assessments of the economic and defense climates suggest that the more favorable recruiting environment of the last several years has seen its end. The drop in defense budgets and concomitant increase in alternative job opportunities among the nation's recruit population are predicted to make recruitment goals difficult to achieve.[55] It is critical therefore that the defense department avoid unnecessary costs associated with performance or injury-related attrition, which can be considerable. For example, the cost associated with a single attrition from army basic combat training in 2009 was estimated at $57,500.[56] Performance related failures may account for as much as 29% of discharges from basic training across the service branches.[44] Low-fit soldiers are also more likely to attrite at 180 days' of service, a time when most are receiving advanced training for their occupational specialties.

Given the cost and frequency of performance related attrition, a logical solution might be to allocate more time to physical training and conditioning. However, whereas physical training is required for enhancing human performance, it also increases exposure. With excessive volume or intensity, the benefits of physical training reach a plateau while risk of injury continues to increase.[57] Injury rates among service members are already unacceptably high. Of the 600,000 soldiers who report musculoskeletal injuries on an annual basis, the majority result from overtraining or overuse.[7] Injuries not sustained during battle have accounted for the greatest proportion of U.S. soldier medical evacuations from Iraq and Afghanistan, the majority of which result from physical training or recreation.[58] Effectively promoting human performance while preventing musculoskeletal injuries therefore requires a measured approach.[7,59] The lack of evidence-based metrics which can be used to benchmark health and fitness among military personnel is arguably the greatest obstacle to achieving the HPO vision.[6,7]

The popularity of the functional training paradigm within the defense community is evident in the concepts underlying the aforementioned intervention programs. Despite a growing body of research evaluating this approach, validated methods are lacking and current practices in many cases are based on commercial claims.[6,7,60] As will be discussed further in the following sections, the FMS established a fair degree of predictive validity with respect to prospective injury risk in soldiers. Indeed, some of the most compelling data supporting the use of FMS testing comes from research in military cohorts. Even so, published data regarding its capacity to predict performance outcomes remains equivocal.

2.2 Fitness and Performance

Performance is considered a multidimensional construct consisting of relatively independent components. Much of the early work on performance (or, in some cases, "fitness") described several broad dimensions, each of which could be divided into smaller subdomains. Different authors offered varying accounts of what these dimensions were and how they were related to one another. Fleishman used factor analysis in military populations to arrive at a physical performance model featuring strength, flexibility, speed, balance, coordination, and endurance.[61] Hogan identified a three factor model which featured muscular strength, cardiovascular endurance, and movement quality.[62] Hogan's strength and movement quality constructs each included three subcomponents. Strength was composed of the ability to generate muscular tension, the ability to generate muscular tension quickly (i.e. muscular power), and the ability to generate sustained muscular tension (i.e. muscular endurance). Similarly, movement quality was further divided into flexibility, balance, and coordination. Shortly after Hogan's work was published, Myers et al. proposed a six factor solution more closely aligned with the earlier model of Fleishman.[63] The Myers et al. solution featured static strength, dynamic strength (which today might be referred to as local muscular endurance), explosive strength, trunk strength, stamina, and extent flexibility. They emphasized that Hogan's three factor solution was artificially simplistic in its grouping of constructs, potentially owing to a lack of statistical power and failure to include a sufficient array of tests. Further, they found no support for a movement quality construct as in Hogan's model.[63]

These models form the basis for all military guidance concerning physical training, the majority of which focuses on cardiovascular endurance, muscular strength and endurance, and flexibility/mobility. For example, the Army's perspective on the structure of physical performance includes factors identified as strength, endurance, and mobility, along with subdomains analogous to those outlined by Hogan.[64,65] Similar factors feature prominently in both the Navy Physical Readiness Training Program Instruction[66] and Air Force Fitness Program Instruction.[67] Each of the performance constructs emphasized in military physical training guidance documents is both modifiable[52] and testable through a variety of approaches that translate to occupationally relevant tasks.[63] The focus of training in recent years has shifted toward increased specificity to the mission,[60] which may confer benefits in the way of promoting fitness without increasing risk of injury.[7] However, whereas theories of fitness and performance are becoming increasingly influential in physical training, the evaluation of fitness in the military remains relatively antiquated.

2.2.1 Performance Screening and Testing in the Military

Currently, the Marine Corps is the only military branch that evaluates a candidate's fitness prior to enlistment.[56] Other basic training academies rely primarily on written tests, medical history, and limited background investigation to identify candidates with an elevated risk of substandard performance or behavioral problems which may prevent them from meeting their service commitments. All candidates must be cleared for enlistment (or commission) through a Military Entrance Processing Station (MEPS). Where pre-accession fitness testing has been implemented, it is conducted either at MEPS (as in the case of the Army's Assessment of Recruitment Motivation and Strength) or at recruiters' stations (as in the case of the Marine Corps Initial Strength Test). The Assessment of Recruit Motivation and Strength consists of a maximum repetition push-up test lasting one minute and modified Harvard step test. The step test portion requires prospective recruits to step up and down from a box (12" high for women, 16" high for men) for 5 minutes at a pace of 30 steps per minute. By comparison, the Marine Corps Initial Strength Test consists of a set of pull ups (males) or flexed-arm hang (females), a 2-minute effort of abdominal crunches for maximum repetitions, and a timed 1.5 mile run.

Fitness standards during the basic training academies are similar in structure. The Marine Corps fitness test is the same as its Initial Strength Test with the exception that the run is extended to 3 miles. The Army, Air Force, and Navy fitness tests each include push-ups (Air Force: 1minute, Army: 2 minutes, Navy: 2 minutes), sit ups (Air Force: 1minute, Army: 2 minutes, Navy: 2 minutes), and running (Air Force: 1.5 miles, Army: 2 miles, Navy: 1.5 miles). Two additional measures, waist circumference and sit-and-reach, are unique to the Air Force and Navy, respectively. Even when considered together, these tests do not paint comprehensive, multidimensional picture of fitness corresponding to the performance constructs identified in the previous section. Thus, it could be argued that there is a disconnect between military fitness doctrine and the testing procedures in place across the branches.

While this limitation with respect to fitness testing procedures has been noted, it is important to consider the restrictions associated with conducting fitness assessments on such a large scale.[68] Military installations are required to administer thousands of fitness examinations every year. Doing this in a cost-effective manner while avoiding tests which introduce undue bias against certain body types is logistically very difficult. The benefits of any modifications to existing test procedures must outweigh any costs associated with time, personnel training, and equipment. As a lightweight and field-expedient tool, movement screening has the potential to address the limitations of the current fitness standards while adding minimal overhead to the process.

2.2.2 Why Movement Screening?

Outcome-based performance measures seek to quantify the construct underlying a given fitness test. Thus, a comprehensive fitness assessment might include a one-repetition maximum weight lift test to evaluate strength, a sprint test to evaluate speed, and so forth. Scores on these measures provide information concerning an individual's performance capability, but may overlook valuable information regarding the strategy used to achieve the outcome. The creators of the FMS argue that two individuals who use different movement strategies to achieve a similar score on a performance test should not necessarily be considered equal.[8] Specific movement behaviors may be associated with greater risk of injury. For example, dynamic knee valgus during landing or cutting maneuvers is thought to increase the risk of ACL rupture.[69] The same may be true of movement behaviors with respect to performance.[70] Thus, identifying deficits relating to the efficiency or stability of one's movement behaviors could complement other sources of information used to establish baseline performance data, predict injury, and guide training objectives.

The notion of the movement strategy being an independent constituent of performance outcomes has gained considerable traction in recent years.[11,70–73] The number of strategies one can use to approach a given movement task is potentially limitless.[74] It is certainly conceivable that some of these strategies will make relatively more efficient use of biomechanical degrees of freedom . Proponents of movement screening argue that there are "optimal" strategies which could serve as a benchmark for assessment and program design,[71,73,75,76] a line of reasoning which predates the recent rise in the popularity of tools like the FMS. Indeed, there is a large body of research which attempts to model movement behaviors in relation to some type of cost function which determines the solution to a motor task. Thus, computer models describe optimal movement behaviors based on minimizing energy expenditure or mechanical strain.

That there exists a set of objectively optimal movement behaviors which might be appropriate for different tasks, let alone different individuals, represents a key distinction between this framework and theories of movement behavior grounded in dynamical systems theory. The traditional medical model rates health or performance in the context of population norms.[77] However, the variability of an outcome may in some cases be more telling than an average.[78] Dynamical systems theorists note that many human movement behaviors—even those of elite athletes—are characterized by nonrandom variability. Further, individuals attempting to adopt an exemplary movement pattern may perform worse than when using more naturally occurring behaviors.[79] Thus, objectively ideal or optimal movement strategies may not exist within or between individuals. Even if an optimal strategy exists for a particular task, it is not currently possible to model the complex array of constraints which define that strategy at any given point in time.[16]

The more relevant discussion for evaluating movement behaviors may therefore revolve around the degree to which a movement system is constrained. This theoretical distinction need not impact the practical administration of movement screens like the FMS. As will be discussed further in the sections on complexity and the movement system, the important takeaway is that the purpose of the clinical examination is to screen for constraints—like strength or range of motion restrictions—which could prevent a candidate from meeting the performance demands of military training rather than grading an individual's movement behavior in relation to a template of perfect function. In the context of training interventions, the implications of the distinction between perspectives are more far reaching. Rather than training individuals to behave according to a template of optimal movement as in FMS-based programs,[80] the appropriate focus of training from the constraint-based perspective is modifying the constraints themselves. The FMS, then, is a window into the constraints impinging on the system. As opposed to outcome-based performance measures, the approach of movement screening gets us closer to understanding what the most important limiting factors are in any particular case.

2.2.3 Summary

The emphasis on Human Performance Optimization reflects the military's increasing focus on evidence-based identification, measurement, and training of factors which promote resilience on the battlefield. Perennially high injury rates combined with recent shifts in the recruiting forecast suggest a requirement for refined methods of predicting and improving physical performance. Physical fitness is a multidimensional construct which is insufficiently measured in basic training fitness standards and largely unmeasured prior to enlistment. Movement screening may be a feasible adjunct to current fitness standards which can be used to assess constraints impacting an individual's physical performance potential.

2.3 The Functional Movement Screen™: Design & Administration

The Functional Movement Screen consists of 7 movement tasks administered in a standardized order. Each of the tests is assigned a score from 0-3 based on the following criteria: 0) Pain experienced during the movement task, 1) Inability to complete the movement task, 2) The examinee is able to complete the task with movement compensation or is able to complete an accepted modified version of the task, 3) The examinee is able to complete the task as prescribed without movement compensation.

In addition to the 7 scored tests, examinees are required to complete 3 "clearing tests" during the screening process. These clearing tests were included to prevent the administration of certain test items which may be contraindicated and/or to prevent potentially "false" high scores. The 7 scored tests and 3 clearing tests appear in the following order:

(1) Deep Squat
(2) Hurdle Step
(3) Inline Lunge
(4) Shoulder Mobility
(5) Impingement Clearing Test
(6) Active Straight-Leg Raise
(7) Trunk Stability Pushup
(8) Press-Up Clearing Test
(9) Rotary Stability
(10) Posterior Rocking Clearing Test

Of those tests that are administered bilaterally, the lower of the two scores is counted toward the total. In most conventional applications the total final score is recorded and represents a cumulative assessment of the examinee's movement quality.

2.3.1 FMS & Injury Risk

The Functional Movement Screen (FMS)[8,9] has been used as a screen for risk of injury,[23] a test by which to plan and evaluate intervention programs,[13,80,81] and as a predictor of physical performance.[82] Its application spans a variety of populations, including youth,[83] high level athletes,[23] military and public safety,[4,24] and middle aged adults.[84] It was intended to address a gap in research and clinical practice in the area of corrective exercise. The goal was to provide a theoretical framework and test for normal biomechanical function which could serve as 1) as complement to the pre-participation examination, and 2) an indicator of how to proceed with remediation when movement deficits are identified. Whereas other tests may provide a quantification of the outcome, such as time to completion of a task or distance covered, the FMS attempts to assess the strategy that led to the outcome.

The screen is based on the underlying theory that the elemental unit of human biomechanics is the movement pattern.[11,85] According to the theory, proper biomechanical function is contingent upon a subset of movement patterns which are encoded as programs in the central nervous system and constitute the foundation for movement behaviors. The patterns are analogous to the 7 FMS test items which are intended to assess whether or not they are intact. Dysfunctions noted in the screen are understood to reflect deficiencies in motor programming which cannot be further reduced (e.g. to the level of a muscle or joint). It is suggested that sport or exercise participation should be considered only after screening of fundamental movement patterns has been conducted and any deficiencies addressed.[11] The projects detailed in this dissertation were based on dynamical systems theory, which is largely at odds with the perspective of motor programming. Notwithstanding any theoretical discrepancies, the FMS offers familiar and easily accessible tests which can still provide useful information. However, as will be discussed, the results may be interpreted and applied differently under the dynamics framework.

In research settings, the screen has primarily been applied as a predictor of injury.[23,86] Inability to complete a given test as prescribed is assumed to reflect functional deficits which increase injury risk. The proposed mechanism by which this occurs involves the secondary movement strategies that arise to compensate for the deficits identified by the screen. These compensatory movement strategies redistribute tissue strain and expend musculoskeletal resources which might otherwise be used to prevent a potentially injurious situation. Injury prediction models commonly use the composite score to discriminate between individuals at high or low risk of injury. A composite score of 14 or less was retrospectively associated with increased risk of injury over the course of a season in American football players.[23] Another study in student-athletes of varying sports also found low composite scores to be retrospectively predictive of increased injury risk, this time at a cut score of 17.[87] Prospective investigations have shown that a score of less than or equal to 14 was associated with increased risk of injury in female college athletes[86] and male Marine officer candidates.[3] In the latter study, a cut score of 18 or greater was unexpectedly also associated with increased risk of injury. Other investigations have been unable to identify prospective associations between FMS composite scores and injury in recreational runners,[85] NCAA basketball players,[25] or high school basketball players.[88]

Results of the FMS are also used in designing and evaluating interventions. Several studies have been conducted in the area of corrective or rehabilitative exercise using FMS scores as an outcome measure. One investigation observed increases in FMS composite scores in a sample of special operators undergoing a comprehensive functional exercise training intervention.[2] Composite score increases following intervention have also been reported in NFL football players, firefighters, and mixed martial arts competitors.[14,15,81] These studies provide preliminary evidence that FMS scores may respond to interventions. Notwithstanding, only one longitudinal training study to date has incorporated a control group. In this investigation, the authors were unable to conclude that training could increase FMS scores.[13]

Table 1. Previous Findings on FMS & Performance

	Sprint	400	Shut.	Agil.	VJ	BJ	MB Throw	Plank	BSE	SitUp	Grip	Pullup	1RM Sqt	1Leg Sqt	Clean
DS															
Okada[26]				–			–							–	
Lockie[89]	–		–	–	.43	.46									
Frost[90]	–	–		–	–	–	–	–	–	–	-.30†	–	–	–	–
HS															
Okada[26]				-.52			.42								
Lockie[89]	–		–	–	–	–									
Frost[90]								–	–	–	–	–			
IL															
Okada[26]				-.46			–							–	
Lockie[89]	–		–	–	–	.45									
Frost[90]								.27			–				
McGill[24]											–	.25			
Hartigan[91]	–				–										
SM															
Okada[26]				.39†			-.39†							-.45†	
Frost[90]								–	–	–	–	–			
SLR															
Okada[26]				–			–							–	
Frost[90]								–	.34	–	–	–			
TSPU															
Okada[26]				–			.41							–	
Frost[90]								-.32†	–	–	–	–			
RS															
Okada[26]				–			.39							–	
Frost[90]								–	–	.42	–	–			
Total															
Crouse[82]*	-.20		-.37		.28	.28							.32		.37

FMS/performance correlations. Dashes indicate relationships that were tested and not found to be significant. DS = Deep Squat, HS = Hurdle Step, IL = Inline Lunge, SM = Shoulder Mobility, SLR = Active Straight Leg Raise, TSPU = Trunk Stability Push Up, RS = Rotary Stability, Total = FMS Composite Score. *Results from Crouse[82] are standardized coefficients from models which included other variables. †Indicates that the FMS score predicted poorer performance.

2.3.2 FMS & Physical Performance Prediction

The screen, or part(s) of the screen, is also used to assess physical performance capacity and identify talent. This association has been examined at both the composite score and item score levels. Performance outcomes in relation to FMS scores have been examined in collegiate soccer, volleyball,[92] basketball,[25] and golf;[27] law enforcement,[93] and public safety personnel.[24] Higher FMS scores are hypothesized to predict increased levels of physical performance through increased movement economy and stability, but have not been shown to do so consistently.

Several exploratory investigations have been conducted to identify relationships between FMS item scores and various performance outcomes. The relationships identified in these investigations have often been isolated. For example, Okada, Huxel, and Nesser[26] studied recreational athletes and found that higher Shoulder Mobility scores were predictive of better performance in the Backwards Overhead Medicine Ball Throw (r = -0.39, p = 0.042) and single leg squat endurance test (r = -.45, p = 0.017), but poorer performance in the agility T-test (r = 0.39, p = 0.039). The same study showed that better Backwards Overhead Medicine Ball Throw performance was also predicted by Trunk Stability Push Up (r = 0.41, p = 0.32), Hurdle Step (r = 0.42, p = 0.028), and Quadruped Rotary Stability (r = 0.39, p = 0.040) while better agility T-test times were predicted by higher scores in the Hurdle Step (r = -0.52, p = 0.005) and Inline Lunge (r = -0.46, p = 0.013). The association of Hurdle Step and Inline Lunge to T-test agility times may be explained by the shared demand for lateral stability among all three tests. Similarly, Trunk Stability Push Up and the BOMB test both rely on strength to some degree and therefore might be expected to correlate. It should be noted, however, that isolated testing of bivariate correlations can be misleading as this approach does not account for shared correlation among variables.

A shortened version of the T-test was used in another study conducted in recreational athletes by Lockie et al.[89] in which no relationship was observed with either the Inline Lunge or Hurdle Step. The latter study took a more focused approach in examining the first 3 FMS tests, which specifically address lower body function, and examined a variety of performance outcomes relating to multidirectional speed and power. They reported an association between the Deep Squat and broad jump (r = 0.43)/vertical jump (r = 0.46), but not between Deep Squat and sprint speed or agility.[89] Standing broad jump was also found to correlate positively with Inline Lunge on the left side (r = 0.45). This lack of association between vertical jump height is consistent with another study focusing specifically on the FMS Inline Lunge which found no relationship to drop jump height in a convenience sample of healthy 18-40 year old men and women.[91] While the protocols differ slightly (the Lockie et al. study used a countermovement jump), these findings may suggest planar specificity with regard to the performance implications of the Inline Lunge.

Cross-sectional studies similar to those just summarized have also been conducted in public safety personnel. In conjunction with FMS scores, Frost et al.[90] assessed performance outcomes including measures of trunk endurance (front and side plank time, Beiring-Sorenson extension test time), grip strength, and maximum number of pull up repetitions in a population of local law enforcement officers. Of the significant relationships observed, half were opposite the predicted direction. Trunk Stability Push Up predicted poorer plank endurance (r = -0.32 left side plank, r = -0.33 right side plank), while Deep Squat predicted poorer left grip strength (r = -0.30). The Inline Lunge was positively correlated with left side plank (r = 0.27) and Beiring-Sorensen Extension time (r = 0.34). An additional study from the same group, this time involving firefighters, reported a positive association between Inline Lunge scores and maximum number of pull up repetitions (r = 0.25 for Inline Lunge Left, r = 0.25 for Inline Lunge Right).[24]

While it would seem tenable that the relationships between FMS movement behaviors and physical performance are population-specific, the inconsistency among performance outcomes in the studies described thus far prevents drawing any firm conclusions in this regard. A greater degree of overlap can be found in studies using collegiate athletes from diverse sports. One study conducted in a sample of NCAA football players found relationships between the FMS composite score and squat strength, power clean strength, 40 yard dash time, shuttle run time, and vertical jump height.[82] While these findings are impressive, they may not be generalizeable to non-football athletes. Another study observed several similar performance measures—1RM squat, 10 meter sprint, 20 meter sprint, T-test agility, and vertical jump height—in a team of collegiate varsity golfers and found no relationships with the FMS composite score.[27]

In addition to those which focus on benchmark performance tests, some research has included event-specific proficiency outcomes. Prospectively, greater improvements in event performance were shown in a group of college track athletes scoring above 14 when compared to their teammates scoring 14 or below.[94] The approach of using a cutoff score is rare in performance-oriented investigations and therefore makes comparison to these results difficult. McGill, Andersen, & Horn[25] used FMS item scores prospectively to predict player efficiency statistics including points, rebounds, steals, assists, and blocks per game in male NCAA basketball players, though no relationships were found. One study has examined relationships between the FMS and tactical performance in law enforcement officers.[95] Performance measures included field tasks such as marksmanship and forcibly arresting assailants and were not found to associate with FMS composite scores.

2.3.3 Limitations of the FMS

Previous findings concerning the predictive validity of the FMS may have been limited by the analytical approaches taken by the authors. The original intention of the FMS creators was to arrive at a single metric—the composite score—which could be considered a comprehensive measure of an individual's movement quality. However, the utility of the composite score has been challenged by two recent factor analyses.[96,97] The first of these studies was conducted in a sample of 934 Marine Corps officer candidates.[96] Over 90% of the participants were male. The second study was conducted in a sample of 290 internationally competitive athletes and featured a much more even gender split (143 males, 147 females).[97] Both studies performed exploratory factor analyses on the polychoric correlation matrix of FMS item scores and both used varimax rotation in their factor solutions. The authors of the first study argued that the factor structure of the screen suggests a minimum of 2, and possibly up to 7, different underlying constructs. Further, their results showed that Rotary Stability is negatively correlated with the other component tests. The results of the second EFA study are more concretely indicative of a two-factor solution, the first featuring Rotary Stability on its own and the other featuring the three standing tests—Deep Squat, Hurdle Step, and Inline Lunge. In addition to showing refuting the screen's underlying unidimensionality, both studies reported low Cronbach's alphas (0.39 and 0.58, respectively). Considering these unfavorable psychometric properties, the better approach to analyzing FMS data may therefore be to interpret each item score separately. This might better preserve valuable information in cases where an equal composite score was achieved through different combinations of item scores.

As we have shown, analyzing item scores is the more common practice when the outcome of interest is performance. Even so, the implications of using item scores in place of the composite score have not been appropriately considered. The main challenges that arise when using item scores are 1) an increase in the number of predictors, and 2) the rank-order structure of the item-level scoring. Many of the studies summarized in Table 1 are critically underpowered for the number of statistical tests conducted. Further, all but one used Pearson or Spearman correlations. (The exception is Crouse 2014,[82] in which only the composite scores were analyzed.) While Spearman rank order correlations are appropriate for ordinal data, multiple regression would provide a better representation of predictive relationships with the outcome. In order to arrive at accurate regression models which account for rank order structure within the set of predictors, more robust methods are required. It should be noted that this problem is not limited to the FMS. Screening tools with similar scoring criteria are becoming increasingly popular. Of those pertaining strictly to rating movement quality, recently developed instruments with ordinal item scoring include the Resistance Training Skills Battery,[98] Return to Duty screen,[99] Frohm et al. Nine-Test Battery,[76] Movement Competency Screen,[10] JobFit,[100] 16-item Physical Performance Measure,[101] Athletic Ability Assessment,[102] and the Netball Movement Screening Tool.[103] This would suggest that revised analytical methods may be highly valuable for performance prediction in a variety of fields.

In an attempt to address the difficulties of analyzing ordinal item score data, we applied an advanced statistical technique known as penalized regression.[29] A more detailed description of penalized regression will be provided in the Methods chapter. Briefly, this technique uses a data-driven parameter to penalize unrealistic coefficient values. This serves to drive model selection away from solutions which are biased in favor of the unique error variance of a given sample. This can be especially useful for high-dimensional data or underpowered analyses. Further, special "smoothing" penalization methods have been developed for ordinally scaled predictor variables and can be applied both within and between predictor groups. Table 2 summarizes some of the findings our first investigation with and without coefficient smoothing for the weighted and unweighted FMS conditions. In this analysis, penalized regression proved to be a valuable tool for model selection and, following model selection, precise estimates of coefficient confidence intervals.

Table 2. Pilot Data on FMS & Performance

	Smoothed			Unsmoothed	
Test	Level	Coef	CI$_{95}$	Coef	CI$_{95}$
DS	2	-11.23	(-22.08, -1.73)*	-18.05	(-47.53, 4.08)
	3	-19.59	(-41.38, -8.79)*	-17.78	(-62.01, 3.77)
HS	2	-12.09	(-29.37, -1.86)*	-14.15	(-38.44, 31.94)
	3	-17.15	(-38.22, -4.96)*	-4.49	(-31.00, 84.50)
ILL	2	7.06	(-4.53, 18.68)	9.14	(-21.64, 33.42)
	3	19.3	(-1.84, 46.10)	29.99	(-3.86, 94.74)
SM	–	–	–	–	–
ASLR	–	–	–	–	–
TSPU	2	0.48	(-9.47, 13.30)	16.44	(-10.14, 105.53)
	3	-22.4	(-35.85, -5.46)*	-32.07	(-59.63, 12.37)
RS	2	-6.42	(-20.25, 3.78)	-11.56	(-43.49, 7.87)
	3	-6.42	(-20.25, 3.78)	-11.56	(-43.49, 7.87)

Previously published[104] unstandardized coefficients predicting Mobility for Battle course completion times from weighted and unweighted FMS scores. Penalized regression coefficients are shown alongside 95% confidence intervals. Note that corresponding FMS predictors in the unweighted condition were not retained. For the factors retained in the weighted FMS condition, confidence intervals are considerably narrower after accounting for the ordinal structure within item scores via the smoothing algorithm. Reference category coefficients (corresponding to FMS score = 1) are not shown. DS = Deep Squat, HS = Hurdle Step, ILL = Inline Lunge, SM = Shoulder Mobility, ASLR = Active Straight Leg Raise, TSPU = Trunk Stability Push Up, RS = Rotary Stability. *Statistically significant at $p < 0.05$.

2.3.4 Summary

The Functional Movement Screen™ is a simple, field-expedient instrument which has shown fair validity in predicting injury prospectively. The test comprises 7 scored items which are graded 0-3 based on predetermined criteria. Past research has used both the composite score and individual item scores for the prediction of injury and physical performance. Particularly with respect to predicting physical performance, previous findings have been equivocal. Recent factor analyses suggest that analysis of the composite score is inappropriate and that future work should focus on item score data.[96,97] Doing so introduces certain analytical challenges, but these challenges have promising solutions in the domain of penalized regression.

2.4 Complexity as a Window to the Effect of Constraints on the Movement System

As mentioned previously, using screens like the FMS to judge movement behaviors in relation to a theoretical norm is a potentially inappropriate application. The principles of dynamics tell us that movement behaviors will always be optimized relative to the constraints facing the movement system.[105] A more fitting analysis of performance during screens like the FMS might therefore focus on the constraints influencing movement behaviors rather than whether an observed movement pattern is normal. In addition to accounting for constraints with concurrent measures of variables theorized to influence movements, we can also observe their effects through analysis of the dynamic variability of the behaviors under investigation.

Systems which are more adaptable and resilient are characterized by complex variability relating to the dynamic interaction of that system's underlying components.[106] In this context it is the structure, rather than the magnitude, of the variability that defines the outcome of interest.[107] Whereas a sine wave and a white noise signal can be equivalent in terms of variation and central tendency, the structure of variability within the two signals is very different. Specifically, the sine wave is predictable while the white noise signal is random. Adaptable biological systems are characterized by a balance between these two extremes. It is in this middle ground that complexity is maximized and the system is able to adapt to continually changing constraints.[108] Deviations from this healthy pattern of variability, whether toward the extreme of predictability or the extreme of randomness, are thought to be indicative of aging, illness, and disease.[108,109]

The use of center of pressure time series data as an indicator of biological complexity has a rich history in the scientific literature. Maintaining balance even in quiet standing tasks involves complex interactions between sensory receptors and motor effectors. The nervous system must integrate information coming from visual, vestibular, and somatosensory pathways and use this information to coordinate motor responses involving many degrees of freedom.[110] The extent to which these components of the balance system are able to interact and coordinate movement behaviors in response to constantly changing conditions determines the complexity of the balance system's output.[111]

Of the many methods used to classify dynamic complexity in standing balance, Sample Entropy is among the most common.[112] A more detailed discussion of the algorithm will be provided in the Methods chapter. In short, the Sample Entropy statistic ranges from 0-2 and is inversely related to the regularity of the time series on which it is calculated. Thus, with higher values of Sample Entropy, the coordinative processes underlying the signal are said to be more complex. Sample Entropy calculated on center of pressure data has been shown to distinguish between experimental conditions as well as between healthy and clinical populations. Compared to a control condition with full visual information, Ramdani et al. observed a reduction in Sample Entropy of the differenced center of pressure time series (anteroposterior, mediolateral, and resultant) during a quiet double leg standing task in young adults when the subjects closed their eyes.[113] We compared chronic ankle instability subjects to healthy controls.[114] Our analyses revealed lower

Sample Entropy in the chronic ankle instability group for double leg and single leg resultant center of pressure velocity and single leg mediolateral center of pressure velocity. Cavanaugh et al. observed a reduction in Approximate Entropy—a statistic closely related to Sample Entropy—of the mediolateral center of pressure time series following cerebral concussion in a sample of collegiate athletes.[115] This reduction Approximate Entropy persisted 3-4 days following the injury irrespective of whether other indicators of postural stability had returned to baseline.

Thus, entropy values calculated on center of pressure time series data are sensitive to different constraints which impact the dynamics of postural control. The effects of constraints which limit physical performance may similarly be observable through data sampled during clinically accessible movement screens. Until recently, the application of entropy analyses to center of pressure data was limited by the technical requirements of the algorithms. Specifically, metrics like Sample Entropy are designed to be derived from stationary signals.[111] While Sample Entropy may be robust to the degree of nonstationarity in center of pressure profiles of quiet standing tasks, movement behaviors similar to the FMS tests would require a different approach. Additional processing methods are now available which can reliably approximate dynamic complexity over multiple timescales while accounting for nonstationarities that might be expected during moving tasks or transitions between experimental conditions.[111,116] These methods include empirical mode decomposition, Multiscale Sample Entropy, and multivariate extensions of the same[116] and will be discussed further in Chapter III.

2.5 Constraints on Military Performance & FMS Movement Behaviors

In discussing motor coordination, Newell considers constraints arising from 3 sources—task, environment, and organism.[17] Our investigation sought to highlight organismal constraints through modifying the task such that deficits in performance adaptations would have more profound consequences for the assessment.[29] The approach of manipulating constraints in this way is perhaps too reductionist to account for the dynamic nature of their interaction. Because the relationship between system constraints and system output is not additive,[16,17,111] we must acknowledge that our ability to generalize conclusions arising from specific experimental manipulations may be limited.[117] A more complete solution might take the approach suggested by Newell and in examining movement behaviors as constraints are adjusted throughout a broad range rather than under relatively few discrete configurations.[117] Because of the potential of increasing fatigue, such an involved approach is not possible with FMS-like experimental tasks. Even so, since this project proposes to modify constraints in a way that is directly relevant to soldier occupational tasks, we are confident that our results will be applicable in screening and training military performance.

Another limitation of the proposed experimental task concerns the standardization of external load. The term "standardized" is used to indicate that the same absolute mass will be used for external load condition for each subject. An alternative approach would be to normalize the mass to subject characteristics of interest. Without normalizing the load, we must assume that the manipulation of constraints will vary, in relative terms, from subject to subject. While this may introduce uncontrolled variance into some of the outcome measures, our reasons for proposing this specific design are two-fold. First, military occupational demands are not scaled to individual characteristics. Thus, normalizing the load would reduce the external validity of our design. Second, the appropriate method of normalization in this context is unclear. Performance measures do not scale geometrically with anthropometric characteristics.[118] Several "allometric" normalization models have been offered, but the scaling exponents vary. Given the methodological uncertainty in normalizing load, and the lack of external validity, the better approach is to standardize the task and account for subject-specific covariates in our statistical models.

2.6 Constraints on Physical Performance Outcomes

Having acknowledged these limitations, our approach to developing an understanding of the mechanisms by which the predictive relationship in our first investigation improved was to identify those constraints which are mutually relevant to 1) the observed performance outcomes, and 2) FMS movement behaviors during the loaded condition. In discussing Army Physical Readiness Training, Knapik et al. 2009[64] define the following subcomponents of fitness: 1) cardiorespiratory endurance, 2) muscular strength, 3) muscular endurance, 4) power, 5) flexibility, 6) balance, 7) speed, 8) agility, and 9) coordination. Each of these fitness subcomponents is likely to have contributed to performance scores in our previous investigation,[29] which included 27.43 meter sprints, a 400 meter run, the Mobility for Battle Assessment,[43] and a simulated partner rescue drag. We will further argue that muscular strength, flexibility (which Knapik et al. alternately refer to as ROM), and balance are also determinants of the quality of FMS movement behaviors performed with external load.

The relationship between strength and performance is well established. McBride et al. normalized 1-reptition maximum back squat measures to subjects' body mass and demonstrated that this quantity was predictive of 40-meter sprint speeds in collegiate football players.[119] Similar results have been observed using allometrically normalized back squat strength measures in internationally competitive soccer players.[120] Peterson, Alvar, & Rhea argue collected similar measures in collegiate athletes from a variety sports. Based on their data, they argue that 1-repetition maximum back squat strength, whether normalized to body mass or not, is a determinant of sprint and agility outcomes.[31] Although a 400 meter run would more appropriately be classified as a middle distance event, strength training and plyometrics have both been shown to improve performance in distance running.[121]

Balance is thought to be an important factor in supporting athleticism and preventing injury.[30] Balance measures have been shown to differentiate between athletes and controls or between competition levels within an athlete group. For example, Davlin showed that elite gymnasts, soccer players, and swimmers all outperformed non-competitive controls in a balance board stabilization task.[122] Likewise, Paillard et al. showed that soccer players competing at the national level exhibit smaller COP surface area and velocity than do regionally competitive players in static single leg standing.[123] Few cross sectional investigations have observed balance as a predictor of timed performance outcomes. However, training studies have demonstrated prospectively that isolated balance interventions can positively impact athletic performance measures. For example, wobble board training was shown to increase vertical jump height[124] while a BOSU™ ("Both

Sides Up") intervention resulted in decreased shuttle run times,[125] both in recreationally active adults.

Restricted range of motion in one area may limit the force that can be generated and delivered to the environment by that segment. Further, such restrictions can initiate a chain of compensatory responses in other areas in an attempt to accommodate any limitations that arise. Perhaps the most common example in this regard concerns the relationship between the hip and lumbar spine. Restricted hip motion is associated with low back pain and may play a mechanistic role by inducing compensatory loading in the spine[126] These mechanisms of compensation may also have consequences for physical performance as the latter has been associated with range of motion as well. For example, sit-and-reach has been identified as a significant predictor of shuttle run speed in D-1 football players.[127]

The relationship between range of motion and performance outcomes may not be linear and may depend on the body site under investigation. Both hypermobility and hypomobility have been observed in relatively low performing groups. One possible explanation is that higher performance is associated with a combination of stiffness in proximal body regions and greater range of motion in distal areas.[128] Alternatively, there may be a kind of "Goldie Locks" zone of ROM where the movement system is minimally constrained. We acknowledged that nonlinear relationships or interaction effects between ROM and specific body sites could make it difficult to identify differences. As a precaution, any subjects with clinical hypermobility will be excluded from participation. Additionally, the measures used in this project included tests of both proximal and distal range of motion to account for the possibility of unique relationships.

2.7 Constraints on FMS Movement Behaviors & Movement Behaviors Under Load

The FMS is purported to assess each of the traits identified in the previous section—strength, balance, and ROM.[8,9,22,23] When considering the additional challenges imposed by carrying an external load, it is likely that our weight vest treatment placed a greater premium on high levels of strength, balance, and range of motion as enablers of high-quality movement.[29] Much of the research on load carriage has been conducted in military populations. Using a sample of young military members, DeMaio et al. observed increases in both anteroposterior and mediolateral center of pressure displacement during quiet standing with personal protective equipment weighing an average of approximately 10 kg.[34] While the DeMaio et al. sample consisted primarily of men, similar results have been observed in military women. Heller et al. compared measures derived during quiet double legged standing with and without an 18.1 kg backpack and found that the backpack condition was associated with higher center of pressure excursion (both anteroposterior and mediolateral) and area.[35]

Schiffman et al. 2006 studied the effect of a series of external loads on the quiet standing postural control of 14 male Army soldiers during 30-second trials of double-leg stance.[129] Their analyses included linear summary measures of center of pressure motion as well as parameters derived from a technique called stabilogram diffusion analysis. With stabilogram diffusion analysis, quiet standing is typically discussed in terms open-loop/no-feedback and closed-loop/feedback mechanisms contributing to postural control over different timescales. As the external loads increased, the authors observed linear increases in center of pressure area and path length as well as a tendency toward less random center of pressure fluctuations over the timescales corresponding to closed-loop control mechanisms. The latter finding is interpreted to reflect increased control requirements to accommodate the added weight.[129]

Joint ROM changes associated with external loading during gait are more nuanced. Park et al. investigated the effects of an 8.16 kg protective vest on kinematic parameters of gait in 7 young adult males and found increases in peak knee flexion and plantar flexion, but decreased transverse motion of the pelvis, relative to a control condition.[130] Birell and Haslam conducted a similar study with loads up to 32 kg. Their findings agree with those of Park et al. concerning decreased transverse plane pelvic ROM. Unlike Park et al., however, the latter study found that sagittal plane knee ROM was reduced and observed no ROM effects at the ankle.[33] Both studies found that the loaded conditions were associated with greater mean anterior pelvic tilt. That ROM changes vary by body region may relate to the aforementioned tradeoffs in movement between different body sites.

While load carriage does not directly affect the force with which muscle tissue can contract, increasing weight does bring an individual proportionally closer to his or her theoretical 1-repetition maximum. Literature on resistance training suggests that technique for certain lifts changes as the load is progressively increased. For example, Walsh et al. found that greater back squat loads (defined as a percentage of the subject's 1-repetition maximum) were associated with greater lumbar hyperextension in young adults with a competitive athletic background.[131] Incremental soldier-relevant loading has also been shown to elicit biomechanical changes during a single-leg cutting task, with decreased knee flexion, hip flexion, and hip adduction of the stance leg resulting from increases in load.[132] Thus, it is conceivable that load increases will impact clinical tests of movement quality.

Because strength, balance, and ROM are relevant to the performance outcomes we propose to study with this investigation, and because they are each readily testable, we will include measures of these factors in our design. For strength, the gold standard assessment is repetition maximum testing. Concerns have been noted with regard to the safety and reliability of repetition maximum testing in untrained populations. We therefore propose to measure upper and lower body strength through validated alternatives with strong association to their one-repetition maximum analogs. Specifically, upper body strength will be assessed with a modified YMCA bench press test while lower body strength will be assessed via countermovement jump peak power. Balance will be evaluated through the resultant center of pressure velocity in quiet single leg stance as previous investigators have noted that double leg postural control is likely not sufficiently sensitive to detect meaningful differences in our population of interest. Finally, three clinical tests of joint ROM will be included. These are the weight-bearing lunge test, the Apley scratch test, and the sit-and-reach test. Respectively, these tests evaluate restrictions related to dorsiflexion, shoulder and thoracic mobility, and hip/trunk flexibility. As a final note, we point out that each of the three factors we have discussed—strength, balance, and range of motion—are modifiable through training and effective intervention programs have been implemented in military settings.[52]

2.8 General Summary

In the context of human movement, complexity arises from dynamic interactions among components of the movement system. Healthy, adaptable behaviors are characterized by complexity in the output of the system. In contrast, changes in complexity may be evident where the system is constrained by aging, injury, disease, or other factors. Constraints related to an individual's strength, balance, and ROM are likely a common influence on physical performance outcomes and externally loaded FMS movement behaviors. Concurrent evaluation of all of these characteristics is necessary in order to verify that this is the case.

As a complementary approach, analytical tools from dynamical systems theory enable us to assess the relative degree to which a system underlying a given signal is constrained. One such tool which is commonly applied in postural control research is sample entropy. Extensions of the sample entropy algorithm which facilitate analysis over a signal's intrinsic timescales make it possible to obtain entropy estimates for multiscale nonstationary time series such as would be produced over brief movement behaviors like FMS tests.

CHAPTER III

METHODS

The purpose of this dissertation project was to identify the impact of balance, strength, and ROM in protecting against 1) decreases in FMS item scores and 2) decreases in the dynamic complexity of movement associated with bearing a standardized military-relevant external load. Additionally, we sought to confirm the validity of externally weighted FMS testing in predicting soldier relevant physical performance outcomes in comparison with conventionally derived FMS scores. We hypothesized that higher decreases in FMS item scores and dynamic complexity of movement would be greater in subjects exhibiting low baseline levels of strength, balance, and ROM.

3.1 Participants

Twenty-five male and twenty-five female recreationally active adults (22.98 ± 3.09 years, 171.95 ± 11.46 cm, 71.77 ± 14.03 kg) participated in this research project. The targeted age demographic, 18-34 years, was intended to reflect the age range of the recruitment population for first time military accession. As shown in Table 3, data from our pilot work suggested that this sample size would provide sufficient power for detecting the effect of the external load condition on FMS item scores with the exception of the Hurdle Step test.[29] Identifying significance in specific regression terms required regularization, which is discussed further in the Statistical Plan subsection. Participation was limited to individuals who exercised a minimum of 90 minutes per week and did not suffer from clinical conditions which could have affected the outcome measurements. Among others, these conditions included chronic instability of the joints of the lower extremity, any recent history of injury ($\leq$ 6 months prior to data collection), Ehlers-Danlos or joint hypermobility syndrome, uncorrected visual impairments not including astigmatism, vestibular disorders, peripheral sensory disorders, or any musculoskeletal condition requiring ongoing care from a licensed healthcare provider.

Table 3. External Loading Effect Sizes for FMS Item Scores

Test	Effect Size	N
Deep Squat	0.59	21
Hurdle Step	0.07	1323
Inline Lunge	0.58	21
Shoulder Mobility	0.76	13
Active Straight Leg Raise	0.60	20
Trunk Stability Push Up	0.68	16
Rotary Stability	0.37	49

Sample sizes required for power = .8 for each FMS test item based on Wilcoxon signed rank tests for matched pairs. Effect sizes are based on pilot data (n = 20) from our laboratory.

3.2 Procedures

After having any questions addressed and providing written informed consent, subjects completed a Physical Activity Readiness Questionnaire (PARQ), demographic information sheet, and had their weight and height measured. Subjects then completed the remaining measurements in a single test session in the following order:

(1) Balance

- Quiet Single Leg Stance

(2) ROM

- Apley Scratch Test
- Weight-Bearing Lunge Test
- Sit-and-Reach Test

(3) FMS Testing (Condition Counterbalanced By Subject)

- FMS Condition 1
- FMS Condition 2
- Cyclic Deep Squat, Hurdle Step, and Inline Lunge Condition 1
- Cyclic Deep Squat Hurdle Step, and Inline Lunge Condition 2

(4) Strength

- YMCA Bench Press Test
- Countermovement Jump

(5) Break and Cycle Ergometer Warm Up

(6) Soldier Performance Outcomes

- 27.43 meter sprints
- 400 meter run
- Mobility for Battle Assessment

- Agility T-Test
- Simulated Partner Rescue Drag

Subjects were familiarized with all measures prior to data collection. The ordering of the tests was designed to minimize the effects of fatigue on the more sensitive measures where possible. Within the FMS Testing cluster, the order of conditions was counterbalanced so as to distribute bias related to learning and practice effects evenly across conditions. Altogether, the protocol averaged just under 3 hours in duration.

3.2.1 Survey & Demographic Data

Subjects completed a computerized physical activity readiness questionnaire and health history. In addition to screening for conditions which might contraindicate participation in the project, these surveys collected data related to physical activity, exercise, and injury history. After completing the surveys, subjects had their height and weight measured.

3.2.2 Single Leg Balance

Balance was tested in quiet single leg stance for a period of 20 seconds, a commonly used trial duration for assessment of postural control, of which only the first 10 seconds were analyzed.[51] Single leg stance was chosen as previous investigations have concluded that double leg standing may not be sensitive enough to detect meaningful differences in a young, healthy populations.[29] All balance testing was conducted barefoot with eyes closed and hands on hips. Participants were required to maintain the test position for the entire 20 second trial period. Any of the following errors constituted a mistrial: removing the hands from the hips, touching the force plate or ground with the non-stance limb, touching the stance limb with the non-stance limb, flexion/extension/abduction of the non-stance hip in excess of 30°, lifting or turning of the stance foot, and opening the eyes. After completing a preliminary practice trial, data were recorded until three successful test trials had been completed.

Ground reaction force data during balance testing was sampled at 100 Hz using an AMTI Accusway force plate (AMTI, Watertown, MA). Unfiltered force data was used to calculate center of pressure in the anteroposterior and mediolateral directions in the Balance Clinic software package (AMTI, Watertown, MA). All testing was conducted using the non-dominant limb with the dominant limb defined as the leg the subject would use to kick a ball for maximum distance. Anteroposterior and mediolateral center of pressure data were combined to yield a resultant time series using a custom LabVIEW program (National Instruments, Austin, TX).

3.2.3 Range of Motion

Range of motion was quantified using three validated clinical measures. The Apley scratch test quantifies range of motion in the shoulders and thoracic spine, the sit-and-reach test measures hip and trunk flexibility, and the weight-bearing lunge test measures dorsiflexion range of motion.

3.2.3.1 Apley scratch test

This test[41] closely mirrors the FMS Shoulder Mobility test. The test begins with participants standing with arms at their sides. When directed, the participant attempts to touch the hands together behind his/her back. With one hand, the subject reaches behind his/her head and down the back. The other hand reached behind his/her lower back and up the spine. The distance between the participant's hands is measured with tape and recorded as the score. In the present study, the average score of the left and right sides was used for analysis.

3.2.3.2 Sit-and-Reach

The sit-and-reach test was conducted using a 30.5 cm wooden box in accordance with the procedures outlined in Ayala et al. 2012.[40] Participants sat on the floor will their legs together and fully extended. For each participant, the examiner positioned the wooden box so that it was touching the soles of the participant's feet, which were aligned with the 22cm mark. Participants were instructed to place one hand on top of the other with palms facing down and to keep the knees and elbows extended. They were then instructed to reach forward along the measuring tape as far as possible and to hold the terminal position for 6 seconds. Subjects repeated the testing procedures until their scores stabilized to within 1cm for 3 successive efforts.

3.2.3.3 Weight-bearing Lunge Test

The weight-bearing lunge test was conducted according to the methods of Hoch et al. 2011.[39] This test began with the subject facing a wall and standing with the test foot aligned with a strip of measuring tape placed perpendicularly to the wall. The non-test foot was stepped back 12-18" for support. While keeping the heel of the test foot firmly on the ground, the subject was instructed to bend at the knee until his/her knee contacts the wall. After being familiarized with the task, subjects moved progressively further away from the wall and repeated the procedure until they were unable to move any further away without lifting the heel of the test foot during the lunge. The distance between the wall and the great toe was recorded and the test was then repeated on the other side. The average distance of both feet was used for analysis.

3.2.4 Strength

Strength testing procedures were selected on the basis of reliability and safety for the population of interest. One-repetition maximum testing is the gold-standard for strength assessment; however, a reliable estimate cannot be obtained during a single session in untrained populations[133] and may additionally be unsafe for these individuals.[38] We therefore used alternative methods which are feasible in untrained populations and are strongly associated with their repetition-maximum analogs.

3.2.4.1 Modified YMCA Bench Press Test

This test was conducted using a gender-specific standardized weight—36.4 kg (80 lbs.) for men and 15.9 kg (35 lbs.) for women.[38] The test began with subjects positioned on a standard weight bench grasping the bar at a comfortable position. A metronome was then set to 60 beats/minute and subjects were instructed to perform bench presses at 30 repetitions/minute such that each beat of the metronome coincided with the bar reaching the up (fully extended) or down (bar on chest) position. The number of repetitions at which the subject was no longer able to maintain the 30 repetitions/minute cadence, or at which the subject could no longer continue, was recorded as the final score. This score has been shown to be a strong predictor of one-repetition maximum bench press loads.[38] A truncated familiarization trial was performed so as to allow subjects to become accustomed to the weight and cadence of the test. However, in order to limit fatigue each participant was permitted only one trial.

3.2.4.2 Countermovement Jump Peak Power

Countermovement jump peak power has been shown to estimate one-repetition maximum back squat with high fidelity.[37,134] Each jump test requires maximal effort on behalf of the participant. Subjects were allotted one practice trial and three test trials with approximately 1 minute of rest between efforts. Subjects began standing on a Bertec force plate (4060-NC, Bertec Corporation, Columbus, OH) with hands on hips. When instructed, they crouched to a preferred depth (countermovement), immediately jumped as high as possible, and finally landed on the force plate. Vertical ground reaction force was sampled at 1000 Hz and low-pass filtered at 40 Hz using The Motion Monitor software (Innovative Sports Training, Inc., Chicago, IL). Data were recorded from the sampling buffer starting one second prior to the activation of a threshold trigger which marked the initiation of the countermovement. Because the first 1 second of data corresponded to quiet standing, it was assumed that initial center of mass velocity was zero. Instantaneous velocity was then calculated using the forward dynamics approach with the following equation.[135]

$$(Force * .001)/bodymass + v_{(i)} = v_{(i+1)} \tag{3.1}$$

Next, a power time series was calculated as the product of the force and velocity curves. The peak of the power time series during the concentric phase of the countermovement jump was then used for analysis.

3.2.5 Functional Movement Screen^{TM}

Following a familiarization round, the Functional Movement Screen™[8,9] was administered both under conventional conditions and while wearing an adjustable vest weighing 18.1 kg (MiR Vest Inc., San Jose, CA). This is comparable to loads used in other studies involving military personnel,[35,42] but may be less than the average combat loads in recent conflicts.[136] It was determined in pilot testing that greater weight vest loads would impose excessive mechanical restriction for several of the FMS tests. We elected to use a standardized 18.1 kg load as this load is sufficient to challenge FMS performance (Manuscript 1) and has a basis in previous research.[42] The weighted and unweighted conditions were randomized.

The FMS was administered by the primary investigator who is experienced and has established measurement reliability with the instrument (see Table 4). The tests administered are listed below and were scored according to the following criteria: 0) Subject experienced pain during any portion of the movement; 1) Subject was unable to complete the movement; 2) Subject was able to complete the task with errors noted; 3) Subject was able to complete the task without error.

(1) Deep Squat (DS): First FMS test item. Subjects began with feet shoulder width apart and arms holding a dowel pressed overhead. When cued, subjects squatted as deeply as possible while attempting to keep the spine straight and then return to the starting position.

(2) Hurdle Step (HS): Second FMS test item. Subjects rested the dowel across their shoulders behind the head. When cued, they raised one leg over an obstacle placed at the height of the tibial tuberosity, touched their heel on the opposite side of the obstacle, and returned to the starting position.

(3) Inline Lunge (ILL): Third FMS test item. Subjects held the dowel in place vertically behind their backs and stood with feet inline on a 2x6" board with a distance equal to the height of the tibial tuberosity separating the toe of the back foot from the heel of the front foot. When cued, subjects dropped down into a lunge position and lightly touched the back knee to the board before returning to the starting position.

(4) Shoulder Mobility (SM): Fourth FMS test item. Subjects were instructed to make fists with their thumbs on the inside. When cued, they attempted to touch their fists together behind their backs by reaching overhead/down the back on one side and up the back on the other side.

(5) Active Straight Leg Raise (ASLR): Fifth FMS test item. Subjects lay on their backs with hips and knees fully extended. Maintaining a straight knee in both legs, they flexed one hip as much as possible, held briefly, and returned to the starting position.

(6) Trunk Stability Push Up (TSPU): Sixth FMS test item. This test required subjects to perform a push up while maintaining a rigid torso and legs (plank) with hands placed in a position slightly more superior than that of a conventional push up. For males, hands were placed on the ground such that a line connecting the thumbs would cross the middle of the forehead when in the face-down position. For females, hands were placed such that a line connecting the thumbs would cross the chin.

(7) Quadruped Rotary Stability (RS): Seventh FMS test item. Subjects were positioned "on all fours," i.e. with hands and knees on the ground. A 2x6"

board was placed length-wise between the hands, knees, and feet, all of which were required to be in contact with the board on both sides. When cued, subjects attempted to reach forward with one hand while reaching backward with the ipsilateral foot, then touched the knee to the elbow, then reached out a second time, and finally returned to the starting position all while avoiding any contact with the floor on the working side.

The FMS also includes three categorical "clearing" exams which are scored as positive or negative based on whether or not the subjects feels pain. The three clearing tests, summarized below, are linked to specific scored tests which were assigned a zero if the associated clearing test was positive.

(1) Impingement Clearing Exam: The subject placed one hand on the opposite shoulder. When instructed, he/she lifted the elbow up and away from the torso until it was at least level with the shoulders. A pain response was considered positive and required that the Shoulder Mobility test be assigned a score of zero. This test was performed on both sides.

(2) Spinal Extension Clearing Exam: The subject lay prone with hands in the push up position. When instructed, the subject pressed the head and shoulders up from the ground until the elbows were fully extended while leaving the pelvis as close to the ground as possible. A pain response was considered positive and required that the Trunk Stability Push Up test be assigned a score of zero.

(3) Spinal Flexion Clearing Exam: The subject began on hands and knees. When instructed, he/she moved the hips backward while allowing the knees to bend

until the hips were directly over the heels. A pain response was considered positive and required that the Rotary Stability test be assigned a score of zero.

It was not anticipated that pain would be a significant factor in the population under investigation. However, because painful movements are automatically assigned a "0", we followed previous investigators in conducting separate analyses in which the effects of pain were considered.[96]

After completing the screen in each condition, participants performed the first three FMS tests (Deep Squat, Hurdle Step, and Inline Lunge) for 5 continuous repetitions on the Bertec 4060-NC force plate. Ground reaction force data were sampled at 100 Hz and used to calculate center of pressure in The Motion Monitor (Innovative Sports Training, Inc., Chicago, IL). Anteroposterior and mediolateral center of pressure displacement time series were then used to calculate the multivariate multiscale sample entropy of the subject's movement behaviors during each of these tasks.

3.2.6 Entropy

Of the entropy estimating algorithms that can be applied to short time series data, Approximate Entropy and Sample Entropy are likely the most popular in the analysis of quiet standing data.[112,137] Sample Entropy provides an index of irregularity within a time series. To begin, a window (i.e. "template") of a predetermined length m is incremented point-by-point throughout the remainder of the time series and compared to subsequent windows of the same length. Each time the template lies within a given radius r of the window to which it is being compared, a match is counted. Once the original template has been compared to the entire time series, a new template of equal length is defined beginning at the next data point. This is repeated until all windows of length m within the time series have been used as the template, all the while adding to the count every time a match is encountered. The template length is then incremented to m+1 and the entire process is completed a second time. The final entropy outcome (Sample Entropy) is the negative natural logarithm of the conditional probability that a match from the first iteration will remain a match at the incremented template length.[112] This number usually ranges from 0-2 with lower numbers reflecting relatively more regular time series and higher numbers reflecting more irregular or complex time series. A key distinction of the Sample Entropy algorithm, as opposed to Approximate Entropy, is that it does not count matches when a template is compared with itself. The authors of the Sample Entropy algorithm showed that this reduces bias,[112] which should recommend its use over Approximate Entropy.

While entropy metrics are very common in postural control research, two additional challenges had to be addressed before they could be applied in this project. First, the sample entropy algorithm operates over a single timescale and will not fully characterize the complexity of cyclic movement behaviors. Second, it assumes that the signal being analyzed is stationary. An extension of sample entropy known as multiscale sample entropy calculates sample entropy over progressively coarse-grained copies of a given time series.[138] This method begins to address the problem of classifying signal complexity over multiple timescales, but can neither accommodate short datasets nor identify timescales which are most salient for the signal under investigation. Because the oscillatory behavior of a signal will be specific to the task and individual being measured, obtaining meaningful entropy estimates requires that the intrinsic timescales characterizing that signal first be identified. The creators of the multiscale adaptation of sample entropy suggest the use of data-driven processing to address nonstationarities which may be present on different scales.[111] Ahmed et al. (2012) propose multivariate empirical mode decomposition to identify intrinsic mode functions inherent to the original signal and use these functions to calculate multivariate multiscale sample entropy.[116] This not only has the advantage of preventing data loss associated with coarse-graining, but also better characterizes complexity across multiple channels (in this case, the anteroposterior and mediolateral center of pressure series) over quasi-stationary intrinsic mode functions.[116]

3.2.7 Soldier Performance Battery

The final portion of the data collection consisted of 5 physical performance tests. This test battery was adapted from previous investigations on tactical performance.[42,43] The dependent variable for each test was completion time as recorded by a photoelectric timing gates (Brower Timing Systems, Draper, UT) or a handheld stopwatch.

(1) Agility T-test: Participants began with one foot depressing the timing gate start-on-release trigger. At the count of "3-2-1-Go," they ran forward 10 yards, shuffled right 5 yards, shuffled left 10 yards, shuffled right 5 yards, and back peddled 10 yards through a timing gate placed at the finish line, all as quickly as possible. The average completion time over 2 trials will be used for analysis.

(2) 27.43 meter (30 yard) Sprints: Participants began with one foot depressing the timing gate start-on-release trigger. They were then instructed to run as

quickly as possible through a timing gate following the examiner's countdown of "3-2-1-Go." The average completion time over 5 trials was used for analysis.

(3) 400 meter run: As in the sprint trials, subjects began with one foot depressing the timing gate trigger. The examiner counted down "3-2-1-Go," after which the participant began a 4.5 lap effort around the periphery of the Coleman Research Gym. As participants entered their final lap, the examiner reminded them to run through a timing gate placed at the 400 meter mark. Subjects were allowed 1 effort for this test.

(4) Mobility for Battle Assessment: This test was designed to provide an evaluation of the physical attributes required for combat[43] and incorporates a range of tasks including shuttle runs, pushups, bear crawls, broad jumps, and water can carries. Because of the detailed nature of the course, participants received a thorough description and demonstration prior to beginning the test, as well as real-time reminders of the tasks as they approached each station. Subjects will be allowed 1 effort for this test.

(5) Partner Rescue Simulation: This test is intended to simulate rescuing an injured soldier. Sandbags were fastened together with a flexible frame constructed from carpeting and wood to create a 68.05 kg (150 lbs.) dummy. Participants were instructed to drag the load for 50 yards as quickly as possible after the examiner counted down "3-2-1-Go." Completion time was recorded when the dummy had crossed the finish line entirely. Subjects were allowed 1 effort for this test.

3.3 Statistical Plan

Hypotheses for each of the three objectives were tested in penalized regression models. Penalized regression facilitates model selection and comparison of new results with those presented in our pilot work.[29] Our justification for using penalized regression was based on the test properties of the FMS. As was discussed in Chapter II, two independent factor analyses were recently published which question the psychometric validity of the FMS composite score.[96,97] Analyzing the item scores as independent variables requires that we find ways to accommodate 1) a much greater number of predictors, and 2) a grouped predictor structure in which the within-group levels are ordinally ranked. Standard regression approaches could be used in which the predictor variables are classified as interval or categorical level data. These approaches include dummy coding, which may lead to overfitting, as well as linear models, in which metric scaling is artificial.[139–141] Penalized regression offers an alternative to stepwise model selection which can be particularly useful for cases involving a large number of predictors relative to the sample size. Further, extensions of penalized regression have been developed which can account for ordinal scaling within the independent variables.[140]

Penalized regression uses a data-driven regularization parameter (Λ) to control the number of variables in the final model. The effect of minimizing the penalized sum of squares term is to drive model selection away from solutions which are biased in favor of the unique error variance of a given sample. These approaches have been shown in simulation studies to outperform conventional regression methods with respect to computational efficiency and model fitting. Further, when applied to smooth differences between adjacent levels of the predictors, penalized regression performs better still.[140,141] Thus, penalized regression methods can be helpful for instruments like the FMS or the other clinical tools identified in Chapter II which make use of ordinal level scoring among the independent variables.[10,76,98–103]

Our previous work suggests that the combination of external loading and regularization is sufficient for detecting a relationship between FMS scores and performance outcomes in a relatively small-sample model (n = 19, models selected from 21 predictors).[29] Based on our pilot data, it was determined that a sample of size 49 would provide enough power to reveal the effect of an external load condition on the item scores of all but one FMS test, the Hurdle Step (see Table 3). Because the regressions proposed in the present investigation will have a more favorable ratio of observations to predictors, we were confident that the condition effects revealed in a sample of n = 50 would be sufficient to highlight the most important model features.

Objectives 1 and 2 were designed to evaluate mediators of the load condition effect on FMS item scores and dynamic complexity, respectively. This was approached through penalized regression modeling techniques designed to evaluate interaction effects with repeated measures.[142] While it may be possible to increase power for these objectives by using separate models for each mediator tested, the advantage of testing all mediators in a single model was that relative importance could be derived based on the order in which factors were discarded from the model. Objective 3 was intended to compare the validity of item scores from the two FMS conditions (standardized external load, control) in predicting physical performance outcomes relevant to the soldier athlete. This was a confirmatory investigation of the results presented in Manuscript 1 with one additional test—the agility T-test.

The following R packages were be used to complete all analyses at an a priori significance level of .05: ordPens[140,143] (version 0.2-.1), grpreg[144] (version 2.6-0), boot[145] (version 1.3-11), and base.

Objective 1) Determine the extent to which balance, strength, and range of motion protect against decreases in quality of FMS movement behaviors associated with standardized external loading.

3.3.1 Objective 1 Hypotheses

(1) Lower center of pressure resultant velocity during single leg stance with eyes closed will be associated with fewer decrements in FMS item scores during a standardized external load condition.

(2) Greater countermovement jump peak power and YMCA bench press tests will be associated with fewer decrements in FMS item scores during a standardized external load condition.

(3) Greater range of motion in the sit-and-reach, weight-bearing lunge, and Apley scratch tests will be associated with fewer decrements in FMS item scores during a standardized external load condition.

Hypotheses for objective 1 were tested using separate penalized regression models for each FMS item score. In order to allow for analysis with the software described above, the log transform of the FMS item score was used as the outcome. We hypothesized that external loading would have a smaller effect on FMS item scores in subjects with higher levels of strength, balance, and range of motion. Such a relationship would be visible through an increase in the magnitudes of the associated coefficients in the external load condition. The models for Objective 1 hypotheses took the following form where "C" is an abbreviation for "Condition":

$$
\begin{aligned}
logScore = &\beta_0 + \beta_1 C + \beta_2 Age + \beta_3 Sex + \beta_4 Height + \beta_5 Weight + \\
&\beta_6 C * Age + \beta_7 C * Sex + \beta_8 C * Height + \beta_9 C * Weight + \\
&\beta_{10} CMJ_{PP} + \beta_{11} YMCA + \beta_{12} SR + \beta_{13} WBLT + \beta_{14} Apley + \\
&\beta_{15} RCOP_V + \beta_{16} C * CMJ_{PP} + \beta_{17} C * YMCA + \beta_{18} C * SR + \\
&\beta_{19} C * WBLT + \beta_{20} C * Apley + \beta_{21} C * RCOP_V
\end{aligned}
\tag{3.2}
$$

Objective 2) Determine the extent to which balance, strength, and range of motion protect against decreased complexity of movement associated with standardized external loading in dynamic postural tasks.

3.3.2 Objective 2 Hypotheses

(1) Lower center of pressure resultant velocity will be associated with smaller decrements in sample entropy during Deep Squat, Hurdle Step, and Inline Lunge performed with a standardized external load.

(2) Greater countermovement jump peak power and YMCA bench press test times will be associated with smaller decrements in sample entropy during Deep Squat, Hurdle Step, and Inline Lunge performed with a standardized external load.

(3) Greater range of motion in the sit-and-reach, weight-bearing lunge, and Apley scratch tests will be associated with smaller decrements in sample entropy during Deep Squat, Hurdle Step, and Inline Lunge performed with a standardized external load.

Hypotheses for objective 2 were tested using separate penalized regression models for each dynamic complexity outcome. We hypothesized that external loading would have a smaller effect on dynamic complexity in subjects with higher levels of strength, balance, and ROM. Such a relationship would be visible through an increase in the magnitudes of the associated coefficients in the weight vest condition. Significance of retained dummy coefficients was tested using bias-corrected and accelerated 95% bootstrap confidence intervals. The models took the following form:

$$\begin{aligned} MMSE = &\beta_0 + \beta_1 C + \beta_2 Age + \beta_3 Sex + \beta_4 Height + \beta_5 Weight + \\ &\beta_6 C * Age + \beta_7 C * Sex + \beta_8 C * Height + \beta_9 C * Weight + \\ &\beta_{10} CMJ_{PP} + \beta_{11} YMCA + \beta_{12} SR + \beta_{13} WBLT + \beta_{14} Apley + \\ &\beta_{15} RCOP_V + \beta_{16} C * CMJ_{PP} + \beta_{17} C * YMCA + \beta_{18} C * SR + \\ &\beta_{19} C * WBLT + \beta_{20} C * Apley + \beta_{21} C * RCOP_V \end{aligned} \tag{3.3}$$

Objective 3) Determine the extent to which soldier-relevant physical performance outcomes are predicted by FMS item scores obtained during a standardized external loading condition in comparison with item scores from a conventionally administered FMS.

3.3.3 Objective 3 Hypotheses

(1) FMS item scores from the standardized external load condition will be associated with shorter 27.4 meter sprint times.

(2) FMS item scores from the standardized external load condition will be associated with shorter 400 meter run times.

(3) FMS item scores from the standardized external load condition will be associated with shorter completion times in the Mobility for Battle Assessment.

(4) FMS item scores from the standardized external load condition will be associated with shorter completion times in the Partner Rescue Drag task.

(5) FMS item scores from the standardized external load condition will be associated with shorter agility T-test times.

(6) With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter 27.4 meter sprint times.

(7) With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter 400 meter run times.

(8) With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter completion times in the Mobility for Battle Assessment.

(9) With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter completion times in the Partner Rescue Drag task.

(10) With the exception of Trunk Stability Push Up, FMS item scores from the control condition will not be associated with shorter agility T-test times.

Hypotheses for objective 3 were tested using separate penalized linear regression models with smoothing of ordinal predictors. We hypothesized that physical performance outcomes would be predicted by FMS item scores in the external load condition. As in Manuscript 1, a group lasso was first be applied using the penalty parameter which minimized cross validation error. The penalty parameter was then be applied to smooth across adjacent levels (i.e. possible item scores ranging 0-3) within the retained groups. As in dummy coded regression, one of the levels must be used as a referent category. In each of our analyses, we designated the lowest score level to serve as this referent category. The models took the following form:

$$\begin{aligned}
Time = {} & \beta_0 + \beta_1 DeepSquat_2 + \beta_2 DeepSquat_3 + \\
& \beta_4 HurdleStep_2 + \beta_5 HurdleStep_3 + \\
& \beta_6 InlineLunge_2 + \beta_7 InlineLunge_3 + \\
& \beta_8 ShoulderMobility_2 + \beta_9 ShoulderMobility_3 + \\
& \beta_{10} ActiveStraightLegRaise_2 + \beta_{11} ActiveStraightLegRaise_3 + \\
& \beta_{12} TrunkStabilityPushUp_2 + \beta_{13} TrunkStabilityPushUp_3 + \\
& \beta_{14} RotaryStability_2 + \beta_{15} RotaryStability_3
\end{aligned} \tag{3.4}$$

In the models which account for pain, this corresponded to an FMS score of “0”, which was represented by a zero dummy coefficient. In the models which ignored pain, the lowest possible FMS item score was “1”. Therefore, in these latter models the zero dummy coefficient represented an item score of “1”. Significance of retained dummy coefficients was then tested using bias-corrected and accelerated 95% bootstrap confidence intervals.

In addition to these regression analyses, item scores from the weighted and control conditions were compared directly using the Wilcoxon signed rank test for matched pairs.

CHAPTER IV

MANUSCRIPT I

4.1 Introduction

Predicting and promoting physical performance is a perennial interest of the US military. Human Performance Optimization[6] (HPO) is an evolving initiative within the defense community which takes a multifaceted approach to addressing performance deficits. In the military, performance related attrition and washback have historically been responsible for substantial budget losses. For example, the estimated cost of attrition from Army basic combat training was over $57,000[56] per individual. Of the discharges occurring less than 6 months into the first term of service, as much as 29%—over 7,000 cases—may be attributable to substandard physical performance.[44]

As costly as performance deficits can be, the feasibility of wide scale pre-accession screening in this area is limited. Any such program in this vein must be valid and unbiased, must require a minimum of time and equipment, and must confer substantial benefit to warrant the effort associated with its implementation. The high recruiting volume of recent years largely allowed for the once active discussion[45,46] of pre-accession screening programs to be tabled. However, cuts in defense spending and a recovering economy are predicted to make future recruiting efforts much more challenging.[7,55] A strong economy provides potential service members with alternative opportunities while shrinking budgets leave less room for recruiting and training expenditures which do not yield a return. This may therefore be an opportune time to revisit strategies for identifying and developing high-performing tactical athletes.

Movement screening is a field-expedient clinical assessment methodology used by many military and paramilitary organizations as a tool for predicting injury risk and performance potential.[93,146] The popularity of movement screening has increased dramatically in recent years as evidenced by the number of screens which have been developed. Examples include the Functional Movement Screen,[8,9] Resistance Training Skills Battery,[98] Return to Duty screen,[99] Frohm et al. Nine-Test Battery,[76] Movement Competency Screen,[10] JobFit,[100] 16-item Physical Performance Measure,[101] Athletic Ability Assessment,[102] and the Netball Movement Screening Tool.[103] Of these, the Functional Movement Screen (FMS) is likely the most popular and well researched with extensive application in military populations.[1,3,147] Despite its popularity, the FMS is a poor predictor of physical performance outcomes.[24]

It has recently been shown that FMS scores decrease when the screen is administered using a standardized external load.[29] Further, item scores from a weighted FMS are better predictors of tactical performance than are conventional FMS item scores.[29] This latter finding might suggest that a load carriage treatment preferentially taxes individuals with low levels of traits which promote tactical athleticism. It remains unclear, however, what these traits might be and the extent to which they can be evaluated using a modified movement screening methodology such as a weighted FMS. Understanding the mechanisms underlying weight-related changes in movement quality, and the associated changes in the relationship between movement quality and tactical performance outcomes, may enhance our ability to identify and focus training on the most salient factors impacting tactical performance.

Performance is a multidimensional construct with several underlying factors. Classical models of human performance identify components such as strength, speed, power, agility, balance, flexibility, coordination, and endurance.[61,64] Several of these traits are also suggested to influence performance of FMS movement tests.[8,9,22,23] It may then be the case that the factors which mutually affect physical performance and clinical movement screens—particularly when such screens are modified to incorporate external load carriage—mediate the observed changes in movement quality and the resulting increase in its association to tactical performance. In other words, a load carriage treatment may highlight the effects of movement deficits which impact performance outcomes.

Of those factors which are said to be assessed by clinical movement quality screens, strength, balance, and ROM are potential mediators of the improved association between physical performance and movement quality under load. Each of these factors has a role in promoting athleticism.[31,125,127] Additionally, each has been shown to interact with external loading. Load carriage has the effect of increasing postural sway[34,35] and also elicits changes in lower body joint range of motion during gait.[130] Furthermore, for reasons which may seem intuitive, stronger individuals are likely to be more robust to the impact of a given absolute load on movement. Thus, studies of weight lifting behaviors often model the response to relative (e.g. percent repetition maximum) loads rather than absolute loads.[131,148] The effects of load may therefore be magnified in individuals with baseline deficits in any or all of these three qualities.

Therefore, the purpose of this investigation was to examine the role of these 3 factors—strength, balance, and range of motion—in mediating the effect of weight on FMS rated movement quality in the military's target recruitment population for initial entry. We hypothesized that 1) the main effect of load would be a decrease in FMS item scores, and 2) these decreases would be smaller in individuals with high levels of strength, balance, and range of motion.

4.2 Methods

This study used a randomized crossover trial to quantify the mediating effect of strength, balance, and range of motion on within-subject differences in movement quality related to external loading. Approval was obtained from the Institutional Review Board at UNC-Greensboro. Data were collected in a laboratory setting by a single investigator experienced in the required measurement techniques. Twenty-five male and twenty-five female recreationally active adults (22.98 $\pm$ 3.09 years, 171.95 $\pm$ 11.46 cm, 71.77 $\pm$ 14.03 kg) participated in the project. Participation was limited to individuals between 18-34 years of age in order to reflect the recruitment pool for military and tactical occupations. Subjects were additionally required to report a habit of accumulating 90 minutes/week of physical activity. All subjects provided written consent to participate and completed a physical activity readiness questionnaire (PAR-Q) before data collection.

4.2.1 Procedures

Participants reported to the laboratory for a single data collection lasting approximately 3 hours. The data presented in this manuscript pertain to the first half of the 3 hour session, which included additional measures as part of a larger investigation. Following consent and completion of the PAR-Q, participants proceeded through the data collection in the following order: 1) Balance, 2) Range of motion, 3) FMS testing, 4) Strength testing.

4.2.2 Balance

Balance was assessed in quiet, single-leg stance using a portable AMTI Accusway force plate and Balance Clinic software (AMTI Inc., Watertown, MA). Similar to previous work,[51] subjects stood barefoot for three trials of single-leg stance during which they were instructed to remain as motionless as possible. Testing was conducted using the nondominant limb with hands on hips and eyes closed. Here, the dominant limb was defined as the preferred side used for kicking a ball for maximum distance. Mediolateral and anteroposterior center of pressure (COP) coordinates were calculated from the raw force data sampled at 100Hz. These data were used to create a resultant displacement time series which was then differenced and divided by the sampling interval to yield a resultant center of pressure velocity (CP_V) series. The mean of this velocity series was recorded for each subject. Only the first 10 seconds of the first acceptable trial was used for analysis in this investigation.

4.2.3 Range of Motion

Range of motion was quantified using three validated clinical measures. The Apley scratch test quantifies range of motion in the shoulders and thoracic spine, the sit-and-reach test measures hip and trunk flexibility, and the weight-bearing lunge test measures dorsiflexion range of motion.

4.2.3.1 Apley scratch test

This test closely mirrors the FMS Shoulder Mobility test.[41] The test begins with participants standing with arms at their sides. When directed, the participant attempts to touch the hands together behind his/her back. With one hand, the subject reaches behind his/her head and down the back. The other hand reached behind his/her lower back and up the spine. The distance between the participant's hands is measured with tape and recorded as the score. In the present study, the average score of the left and right sides was used for analysis.

4.2.3.2 Sit-and-Reach

This sit-and-reach (S&R) test was conducted using a 30.5 cm wooden box in accordance with the procedures outlined in Ayala et al. 2012.[40] Participants sat on the floor will their legs together and fully extended. For each participant, the examiner positioned the wooden box so that it was touching the soles of the participant's feet, which were aligned with the 22cm mark. Participants were instructed to place one hand on top of the other with palms facing down and to keep the knees and elbows extended. They were then be instructed to reach forward along the measuring tape as far as possible and to hold the terminal position for 6 seconds. Subjects repeated the testing procedures until their scores stabilized to within 1cm for 3 successive efforts.

4.2.3.3 Weight-bearing Lunge Test

The weight-bearing lunge test (WBLT) was conducted according to the methods of Hoch et al. 2011.[39] This test began with the subject facing a wall and standing with the test foot aligned with a strip of measuring tape placed perpendicularly to the wall. The non-test foot was stepped back 12-18" for support.

While keeping the heel of the test foot firmly on the ground, the subject was instructed to bend at the knee until his/her knee contacts the wall. After being familiarized with the task, subjects moved progressively further away from the wall and repeated the procedure until they were unable to move any further away without lifting the heel of the test foot during the lunge. The distance between the wall and the great toe was recorded and the test was then repeated on the other side. The average distance of both feet was used for analysis.

4.2.4 Strength

Strength testing procedures were selected on the basis of reliability and safety for the population of interest. One-repetition maximum testing is the gold-standard for strength assessment; however, a reliable estimate cannot be obtained during a single session in untrained populations[133] and may additionally be unsafe for these individuals.[38] We therefore used alternative methods which are feasible in untrained populations and are strongly associated with their repetition-maximum analogs.

4.2.4.1 Modified YMCA Bench Press Test

This test was conducted using a gender-specific standardized weight—36.4 kg (80 lbs.) for men and 15.9 kg (35 lbs.) for women.[38] The test began with subjects positioned on a standard weight bench grasping the bar at a comfortable position. A metronome was then set to 60 beats/minute and subjects were instructed to perform bench presses at 30 repetitions/minute such that each beat of the metronome coincided with the bar reaching the up (fully extended) or down (bar on chest) position. The number of repetitions at which the subject was no longer able to maintain the 30 repetitions/minute cadence, or at which the subject could no longer continue, was recorded as the final score. This score has been shown to be a strong predictor of one-repetition maximum bench press loads.[38] A truncated familiarization trial was performed so as to allow subjects to become accustomed to the weight and cadence of the test. However, in order to limit fatigue each participant was permitted only one trial.

4.2.4.2 Countermovement Jump Peak Power

Countermovement jump peak power has been shown to estimate one-repetition maximum back squat with high fidelity.[37,134] Each jump test requires maximal effort on behalf of the participant. Subjects were allotted one practice trial and three test trials with approximately 1 minute of rest between efforts. Subjects began standing on a Bertec force plate (4060-NC, Bertec Corporation, Columbus, OH) with hands on hips. When instructed, they crouched to a preferred depth (countermovement), immediately jumped as high as possible, and finally landed on the force plate. Vertical ground reaction force was sampled at 1000 Hz and low-pass filtered at 40 Hz using The Motion Monitor software (Innovative Sports Training, Inc., Chicago, IL). Data were be recorded from the sampling buffer starting one second prior to the activation of a threshold trigger which marked the initiation of the countermovement. Because the first 1 second of data corresponded to quiet standing, it was assumed that initial center of mass velocity was zero. Instantaneous velocity was then calculated using the forward dynamics approach with the following equation.[135]

$$(Force * .001)/bodymass + v_{(i)} = v_{(i+1)} \tag{4.1}$$

Next, a power time series was calculated as the product of the force and velocity curves. The peak of the power time series during the concentric phase of the countermovement jump was then used for analysis.

4.2.5 Functional Movement Screen™

Following a familiarization round, the Functional Movement Screen™ [8,9] was administered both under conventional conditions (FMS_C) and while wearing an adjustable vest weighing 18.1 kg (MiR Vest Inc., San Jose, CA) (FMS_W). This is comparable to loads used in other studies involving military personnel,[35,42] but may be less than the average combat loads in recent conflicts.[136] We elected to use a standardized 18.1 kg load as this load is sufficient to challenge FMS performance[29] and has a basis in previous research.[42] The weighted and unweighted conditions were randomized.

The FMS was administered by the primary investigator who is experienced and has established measurement reliability with the instrument (see Table 4). The tests administered are listed below and were scored according to the following criteria: 0) Subject experienced pain during any portion of the movement; 1) Subject was unable to complete the movement; 2) Subject was able to complete the task with errors noted; 3) Subject was able to complete the task without error.

Table 4. Test-Retest Reliability for FMS Item Scores

	Kappa	z	p	Reliability
Deep Squat	0.67	2.88	<0.01	Good
Hurdle Step	0.78	2.54	0.01	Good
Inline Lunge	1.00	3.16	<0.01	Very Good
Shoulder Mobility	0.78	3.00	<0.01	Good
Active Straight Leg Raise	1.00	3.16	<0.01	Very Good
Trunk Stability Push Up	1.00	3.16	<0.01	Very Good
Rotary Stability	0.74	2.42	0.02	Good

Cohen's kappa for test-retest reliability with ordinal data.

4.2.6 Statistics

In order to compare our results to previous work, decreases in FMS item scores related to the weight vest condition were tested with directional Wilcoxon signed rank tests for matched pairs. We then tested our hypotheses concerning effect modifiers using separate regression models with the log-transform of each FMS test item serving as a dependent variable. (The log-transform was used to facilitate analysis with the existing options available in the relevant software packages, detailed below.)

We refer to interaction effects in our models, but it should be noted that a varying coefficients structure was used to account for the differential covariate effects in the two testing conditions.[142] While this type of model is traditionally used to analyze effects which vary over time, it can be applied similarly to analyze effects which vary over condition.[142] Regardless of the order in which the tests were administered, the design matrix was specified such that data from the unweighted condition is modeled as the first of two coefficients for each variable. The second coefficient, corresponding to the weighted condition, represents the change in the effect of the covariate relative to the unweighted condition. Thus, this coefficient can be interpreted as a covariate*condition interaction term. Note that, unlike the examples offered in Hess et al.,[142] the modifying factor in our model is not time, but rather weight vest condition. Because all data were collected on the same day, our set of independent variables is the same for each model. This is reasonable because we do not expect intrinsic performance attributes to vary within the span of a few minutes and any bias associated with condition order is addressed by randomization.

Recall that we hypothesized the decrease in weighted FMS item scores relative to the unweighted condition would be smaller for those subjects showing greater levels of strength, balance, and range of motion. In our models, this would be visible as a positive relationship between the item score and our three mediators in time point two. Because some of our measures are inversely related to their respective constructs, the predicted sign of their coefficients in the weighted condition is negative. Table 5 summarizes the hypothesized sign of the coefficients corresponding the weighted condition.

Table 5. Hypothesized Directions of Effect for Mediator Variables

Variable	H1 coeff sign in FMS_W
CP_V	-
WBLT	+
Apley Scratch Test	-
Sit and Reach	+
YMCA Bench Press	+
CMJ PP	+

A number of nuisance variables, as well as their interactions with the weight vest condition, are also accounted for in our models. These include height, weight, age, and sex. It is apparent from our list of independent variables that the models of interest in our study are likely substantially underpowered for conventional regression techniques. This is especially true for the detection of interaction effects. Problems associated with model selection and lack of power in complex regression analyses such as ours can be addressed through penalization. Penalized regression methods minimize an error term just as more familiar forms of regression, but are subject to additional constraints on the magnitude of the coefficients. These constraints are incorporated using a data-driven tuning parameter, here denoted lambda (Λ), which is usually selected on the basis of some information criterion or cross-validation procedure. The effect of employing such a penalty is to prevent overfitting a model to the variance that is unique to a given sample. Once the models have been selected, standard methods of estimation and significance testing can be applied.

In this investigation, the tuning parameter (Λ) associated with minimum cross validation error (CVE) was first determined using a 5-fold cross validation routine.

Model selection was then performed using the group lasso at the identified Λ value. Because we are interested in explaining variance after accounting for differences attributable to the nuisance covariates (age, sex, height, and weight), these variables were not penalized during tuning parameter identification or model selection. This ensures that they will be included in the final multiple linear regression model. All analyses in the present study were conducted using R (The R Foundation) with add-on packages grpreg[144] and boot.[145] A significance level of $\alpha = .05$ was specified a priori.

4.3 Results

Consistent with results from previous research,[29] the weight vest condition was associated with a decrease in item scores for each FMS test except the Hurdle Step. Item score differences are summarized in Table 6. Model summary statistics are shown in Table 9. With the exception of the Rotary Stability test, each model is significant at the .05 level and accounts for a moderate to large proportion of variance (adjusted R^2 = 0.21 - 0.77). Exponentiated coefficients for individual predictors are presented in Table 11 and Table 12. These coefficients may be interpreted as the factor by which the outcome score is expected change in response to a 1-unit increase in the associated predictor. In this context, a value of “1” corresponds to no effect, whereas values greater or less than “1” correspond to positive and negative effects, respectively. For a given model, relative importance of the various predictors after accounting for the nuisance parameters can be seen in Table 10. This table shows the order in which predictors are retained in the model as the penalty parameter lambda is progressively relaxed from a point at which all coefficients are equal to zero. This same method was used in Hess et al. 2013.[142]

Table 6. Summary of Paired Differences in FMS Item Scores

Outcome	V	p
Deep Squat	71.5	<0.01*
Hurdle Step	36	0.40
Inline Lunge	120	<0.01*
Shld. Mobility	1081	<0.01*
Active Leg Raise	40	0.01*
Push Up	378	<0.01*
Rotary Stability	44	0.03*

Table 7. FMS Item Scores for Weighted & Unweighted Conditions

Test	Unweighted				Weighted			
Score	0	1	2	3	0	1	2	3
DS	1	5	31	13	1	12	27	10
HS	0	1	29	20	0	1	30	19
IL	0	0	18	32	0	3	25	22
SM	2	2	15	31	2	24	23	1
ASLR	0	2	16	32	0	4	19	27
TSPU	0	11	2	37	0	23	15	12
RS	1	0	41	8	0	1	46	3

4.3.1 Nuisance Parameters

In general, weight was the most influential nuisance covariate, having the effect of reducing test performance in the Deep Squat, Active Straight Leg Raise, and Trunk Stability Push Up. Height was predictive of poorer Trunk Stability Push Up performance in the weight vest condition specifically, whereas Sex had differential effects depending on the test. Male sex was associated with poorer performance in the Inline Lunge and better performance in the Trunk Stability Push Up, each of these being relatively strong effects.

Table 8. Descriptive Statistics for Mediator Variables (Manuscript I)

	Mean	SD
Age (years)	22.98	3.13
Height (cm)	172.11	11.18
Weight (kg)	71.82	14.18
WBLT (cm)	9.83	3.83
Apley (cm)	13.12	7.43
SR (cm)	27.22	10.18
YMCA (repetitions)	27.80	16.43
CMJ (Watts)	3308.63	1025.04
CP_V (cm/s)	4.96	1.46

4.3.2 Strength

Predictors related to strength were retained for the Deep Squat, Shoulder Mobility, and Trunk Stability Push Up tests. In each of these tests, higher scores on the YMCA Bench Press test were associated with higher item scores.

4.3.3 Balance

Greater mean CP_V was a significant predictor of better Hurdle Step performance.

4.3.4 Range of Motion

High scores on the weight bearing lunge test predicted better performance in the Deep Squat and Inline Lunge. In the weight vest condition specifically, the same variable was predictive of greater Shoulder Mobility test performance. Higher scores on the sit-and-reach test were associated with better performance in the Active Straight Leg Raise, but poorer performance in the Shoulder Mobility test. Finally, lower (i.e. better) Apley Scratch test scores were predictive of better performance in the Shoulder Mobility test.

Table 9. Penalization and Final Model Summary Statistics

Outcome	Λ	Features	F	p	R^2
DS	0.018	11	$F_{(11,88)} = 7.58$	<0.01	0.42
HS	0.022	10	$F_{(10,89)} = 3.74$	<0.01	0.22
IL	0.013	13	$F_{(13,86)} = 2.98$	<0.01	0.21
SM	0.010	13	$F_{(13,86)} = 26.36$	<0.01	0.77
ASLR	0.017	11	$F_{(11,88)} = 6.71$	<0.01	0.39
TSPU	0.033	9	$F_{(9,90)} = 12.75$	<0.01	0.52
RS	0.017	9	$F_{(9,90)} = 1.09$	0.38	0.01

Table 10. Variable Selection Order as a Function of Λ (Manuscript I)

Rank	DS	HS	IL	SM	ASLR	TSPU	RS
1	Age	Age	Age	Age	Age	Age	Age
1	Age*L	Age*L	Age*L	Age*L	Age*L	Age*L	Age*L
1	Ht	Ht	Ht	Ht	Ht	Ht	Ht
1	Ht*L	Ht*L	Ht*L	Ht*L	Ht*L	Ht*L	Ht*L
1	Sex	Sex	Sex	Sex	Sex	Sex	Sex
1	Sex*L	Sex*L	Sex*L	Sex*L	Sex*L	Sex*L	Sex*L
1	Wt	Wt	Wt	Wt	Wt	Wt	Wt
1	Wt*L	Wt*L	Wt*L	Wt*L	Wt*L	Wt*L	Wt*L
9	WL	CP_V	WL	Ap	SR	Y	Y
10	Y	Ap	$CP_V{}^{\dagger}$	Y	WL	$CP_V{}^{\dagger}$	SR
11	Ap	SR	$Y^{\dagger}$	WL*L	Y	$SR^{\dagger}$	WL
12	CP_V	CJ	Y*L	CP_V	CP_V	Y*L	SR*L
13	Y*L	WL	CJ	SR	Ap*L	WL	$Ap{*}L^{\dagger}$
14	Ap*L	SR*L	Ap	Y*L	Ap	CJ	$Y{*}L^{\dagger}$
15	CJ	Y	SR*L	CJ	Y*L	Ap*L	WL*L
16	SR	$Ap{*}L^{\dagger}$	CP_V*L	SR*L	WL*L	$SR{*}L^{\dagger\dagger}$	Ap
17	WL*L	$Y{*}L^{\dagger}$	WL*L	WL	$CJ{*}L^{\dagger}$	$WL{*}L^{\dagger\dagger}$	CJ
18	L	CP_V*L	CJ*L	Ap*L	$CP_V{*}L^{\dagger}$	Ap	CP_V
19	SR*L	CJ*L	SR	$CJ{*}L^{\dagger}$	SR*L	CJ*L	L
20	CJ*L	WL*L	Ap*L	$CP_V{*}L^{\dagger}$	CJ	CP_V*L	CP_V*L
21	CP_V*L	L	L	L	L	L	CJ*L

Relative importance of covariates shown as a function of increasing the penalty parameter (Λ). The nuisance variables (age, sex, height, and weight), along with their interaction effects, were not penalized and are therefore present in all models. Subsequent variables which share a superscript (†,††) were selected in the same iteration. L = Load-Bearing/Weighted Vest Condition, DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, WL = Weight-Bearing Lunge Test, CP_V = Resultant Center of Pressure Velocity, Y = YMCA Bench Press Test, SR = Sit-and-Reach Test, Ap = Apley Scratch Test, CJ = Countermovement Jump.

Table 11. Coefficients for Deep Squat, Hurdle Step, & Inline Lunge Models

Outcome	Var.	Coef.	t	p	Outcome	Var.	Coef.	t	p
DS	Sex	1.23	1.94	0.06	IL	Sex	0.78	-2.04	0.04*
	Sex*L	1.10	0.69	0.49		Sex*L	1.17	1.31	0.20
	Age	1.00	0.23	0.82		Age	1.02	1.31	0.19
	Age*L	0.98	-0.88	0.38		Age*L	0.99	-0.37	0.71
	Ht	1.00	1.11	0.27		Ht	1.00	0.04	0.96
	Ht*L	1.00	-0.44	0.66		Ht*L	1.00	-0.15	0.88
	Wt	0.99	-2.36	0.02*		Wt	0.99	-1.75	0.08
	Wt*L	1.01	0.96	0.34		Wt*L	1.00	-0.92	0.36
	WL	1.06	7.07	0.00*		WL	1.02	3.33	0.00*
	Ap	0.99	-1.79	0.08		Y	1.00	-0.22	0.83
	Y	1.01	3.07	0.00*		Y*L	1.01	1.89	0.06
						CJ	1.00	1.78	0.08
HS	Sex	0.91	-1.21	0.23		CP_V	1.03	1.82	0.07
	Sex*L	1.11	1.04	0.30					
	Age	1.00	-0.36	0.72					
	Age*L	0.99	-0.52	0.61					
	Ht	1.01	3.79	0.00*					
	Ht*L	1.00	0.12	0.90					
	Wt	0.99	-1.90	0.06					
	Wt*L	1.00	-0.12	0.90					
	Ap	0.99	-1.58	0.12					
	CP_V	1.04	2.18	0.03*					

Coefficients and t-statistics for the Deep Squat, Hurdle Step, and Inline Lunge models. DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, WL = Weight-Bearing Lunge Test, CP_V = Resultant Center of Pressure Velocity, Y = YMCA Bench Press Test, Ap = Apley Scratch Test, CJ = Countermovement Jump.

Table 12. Coefficients for Shoulder Mobility, Active Straight Leg Raise, Trunk Stability Push Up, & Rotary Stability Models

Outcome	Var.	Coef.	t	p	Outcome	Var.	Coef.	t	p
SM	Int	3.14	2.67	0.01*	TSPU	Int	1.54	0.69	0.49
	Sex	0.87	-1.59	0.12		Sex	1.78	4.54	0.00*
	Sex*L	0.93	-0.71	0.48		Sex*L	1.34	1.76	0.08
	Age	0.99	-1.17	0.24		Age	1.00	-0.06	0.95
	Age*L	0.98	-1.61	0.11		Age*L	0.99	-0.41	0.69
	Ht	1.00	0.88	0.38		Ht	1.00	0.38	0.70
	Ht*L	1.00	-0.64	0.52		Ht*L	0.99	-2.07	0.04*
	Wt	1.00	1.36	0.18		Wt	0.99	-2.91	0.00*
	Wt*L	1.00	0.51	0.61		Wt*L	1.01	1.50	0.14
	WL*L	1.02	2.32	0.02*		Y	1.01	4.41	0.00*
	Ap	0.97	-6.86	0.00*					
	SR	0.99	-1.96	0.05*	RS	Int	2.28	3.09	0.00*
	Y	1.00	-2.28	0.03*		Sex	0.97	-0.53	0.60
	CP_V	1.03	1.79	0.08		Sex*L	1.10	1.26	0.21
						Age	1.00	-0.23	0.82
ASLR	Int	3.16	2.35	0.02*		Age*L	1.00	-0.24	0.81
	Sex	1.09	0.95	0.35		Ht	1.00	0.43	0.67
	Sex*L	0.99	-0.12	0.91		Ht*L	1.00	-1.20	0.23
	Age	1.00	-0.02	0.98		Wt	1.00	-1.15	0.26
	Age*L	0.98	-0.95	0.34		Wt*L	1.00	0.93	0.36
	Ht	1.00	-0.33	0.74		Y	1.00	1.44	0.15
	Ht*L	1.00	0.47	0.64					
	Wt	0.99	-2.05	0.04*					
	Wt*L	1.00	0.39	0.70					
	WL	0.99	-1.82	0.07					
	SR	1.02	5.60	0.00*					
	Y	1.00	0.87	0.39					

Coefficients and t-statistics for the remaining models. SM = Shoulder Mobility, ASLR = Active Straight Leg Raise, TSPU = Trunk Stability Push Up, RS = Rotary Stability, SR = Sit-and-Reach Test, WL = Weight-Bearing Lunge Test, CP_V = Resultant Center of Pressure Velocity, Y = YMCA Bench Press Test, Ap = Apley Scratch Test, CJ = Countermovement Jump.

4.4 Discussion

Paired differences in FMS item scores closely mirror previously reported changes.[29] The models depicted in Table 12 contain variables which serve as important benchmarks in the present study. The Shoulder Mobility test is very similar to the Apley Scratch test, which would lead us to expect the Apley Scratch test predictors to be the most important for this model. This expectation is confirmed. (Recall that lower Apley Scratch test scores are indicative of greater range of motion; therefore, the directionality of the expected change in Shoulder Mobility scores for a unit increase in Apley Scratch test scores is negative.) For similar reasons, we would expect sit-and-reach and YMCA Bench Press test coefficients to feature prominently in the Active Straight Leg Raise and Trunk Stability Push Up models, respectively. These expectations are confirmed as well, as is the directionality of the predicted changes. These observations provide an indication that the statistical algorithms employed in our analyses are selecting appropriate models and parameters.

Previous work has shown that FMS item scores, with the exception of Trunk Stability Push Up, were not related to tactical athlete criterion performance tasks unless the screen was performed with an external load.[29] Considering the possibility that the additional load highlighted performance-relevant attributes during the screening process, we hypothesized that the vest treatment preferentially taxed those subjects with relatively low levels of strength, balance, and ROM. We therefore hypothesized that high levels of strength, balance, and range of motion would be associated with smaller decreases in FMS item scores when comparing $\mathrm{FMS_W}$ to $\mathrm{FMS_C}$. This hypothesis would be supported by coefficients corresponding to the weight vest condition ("*L") being selected early and exhibiting a relatively large effect in the appropriate direction, which was generally not the case. Most of the condition-specific covariates were among the last to be retained in the models and were usually not selected at the optimized value of the tuning parameter (Λ). The only condition-specific covariate which was weight-bearing lunge test in the Shoulder Mobility model, in which a positive relationship was observed.

In contrast, several of our mediator variables show noteworthy global effects. WL was the most important predictor of Deep Squat, which supporting what has been suggested by previous authors,[22,73] as well as the Inline Lunge. Deep Squat performance was also promoted by higher levels of upper body strength as measured via the YMCA Bench Press test, possibly indicating the importance of strength throughout the kinetic chain as this outcome focuses more closely on the lower body. Other mediators may have had non-zero effects which simply failed to meet the significance threshold in this study. While this cannot be stated conclusively, the number of non-significant p-values falling under 0.1 could be taken to warrant further study with additional observations. Related to the three lower body tests, these included the Ap coefficient for Deep Squat and the Y*L, CJ, and CP_V coefficients for Inline Lunge. Each of these effects is supported by a plausible theoretical mechanism. The Deep Squat task requires upper body range of motion to maintain the position of the dowel. The Inline Lunge requires balance to accommodate its difficult stance position, as well as strength in both the lower and upper body. Lower body strength facilitates the return to a standing position without "cheating" with trunk extension while upper body strength is essential in transferring the weight of an external load to the lower extremity, similar to the effect of upper body strength on ruck marching capacity.[149] Whether a function of the scoring resolution of the FMS, the effect sizes of our covariates, sample size, or some combination, it may be the case that power to detect multiple effects after controlling for nuisance factors was lacking in our analyses.

Our balance variable, mean CP_V, was retained as the most important variable in the Hurdle Step model. Traditionally, greater COP velocity would be interpreted to reflect poorer balance control.[52,150] However, in this case it was predictive of higher Hurdle Step scores. While balance is one of the attributes purported to be assessed during clinical movement screens, data relating balance and FMS component tests is very limited. One investigation found no relation between Inline Lunge scores and COP excursion[91] while another observed an inverse relationship between Hurdle Step Performance and COP standard deviation.[151] Previous data from our lab indicated that anteroposterior CP_V, albeit from double leg standing, was associated with higher scores in the Deep Squat, weighted Deep Squat, and weighted Inline Lunge.[29] This was initially interpreted as a spurious result potentially related to small sample size and/or the use of a double leg standing protocol, which may lack discriminatory ability in young, healthy populations. While it is difficult to compare the presents results with the previous findings directly, the pattern suggests at least two possibilities. First, it could be the case that the variance in CP_V which is predictive of lower FMS component scores is actually related to a confounder variable such and height or weight. This possibility seems unlikely based on our control and model selection procedures, though it should be noted that other regularization algorithms may be more effective at handling multicollinearity. Alternatively, higher CP_V in this dataset may actually be a reflection of better postural control. It is possible that balance limitations can result in compensatory decreases in postural motion where individuals are not confident to explore their postural control space. It may be the case that lower CP_V in our sample is indicative of a more constrained postural control strategy which limits performance of dynamic tasks.[152]

While the effects of certain covariates are consistent across our models, the variation in predictors and directionality might suggest that the FMS items do not load on a general movement quality capacity. This would be consistent with large scale factor analyses which have concluded that the underlying structure of the FMS composite score is not unidimensional.[96,97] The FMS creators consider movement quality to be a separate component of functional performance. In this sense, it might form a separate category in classic multidimensional fitness/performance models like that of Fleishman,[61] making a unique contribution to an individual's performance capacity. In contrast, the hypotheses in the present investigation consider clinical movement screens to be a convenient, feasible method of observing previously identified performance domains.

4.5 Conclusion

Our findings confirm that a moderate to large portion of the variance (average R^2 = 0.21 - 0.77) in FMS scores is explained through models which include strength, balance, and range of motion predictors. These attributes may therefore be important constituents of performance on clinical movement screens after accounting for the influence of age, sex, height, and weight. At the same time, our analyses failed to show that strength, balance, and range of motion prevent movement quality decreases related to external loading. This may suggest the existence of other factors which are responsible for the differential abilities of FMS_C and FMS_W in predicting tactical performance outcomes such as sprinting, obstacle course completion, and simulated partner rescue tasks. In conclusion, clinically scoring of movement behaviors may be a viable means of predicting physical performance; however, further research is needed to understand the complex relationships between movement quality and performance attributes.

CHAPTER V

MANUSCRIPT II

5.1 Introduction

Performance-related discharge historically accounts for a large portion, if not a majority, of basic combat training attritions.[44] Notwithstanding the associated costs, recruitment efforts over the last decade have been sufficiently successful as to permit the discontinuation of pre-accession fitness screening programs such as the Assessment of Recruit Motivation and Strength (ARMS).[46] Recent reports suggest that this favorable recruiting environment has come to an end. At the same time as defense budgets are being curtailed, a recovering economy is presenting the nation's recruit population with attractive employment alternatives.[7,55] Further compounding the issues, an increasing proportion of America's youth is physically unfit for enlistment.[153] It is critical therefore that the defense department minimize performance and injury-related attrition, which carry a considerable economic burden[44,56] and compromise defense readiness. Efforts to do so might target screening, intervention, or both.

A logical solution might be to allocate more time to physical training and conditioning. However, whereas training is required for enhancing human performance, it also increases exposure. With excessive volume or intensity, the benefits of physical training reach a plateau while risk of injury continues to increase.[57] Injury rates among service members are already unacceptably high. Of the 600,000 soldiers who report musculoskeletal injuries on an annual basis, the majority result from overtraining or overuse.[7] Effectively promoting human performance while preventing musculoskeletal injuries therefore requires a measured approach.[7,59] Accordingly, a number of programs have been developed in recent years which incorporate the principles of so-called functional training.[50–53,154] A staple of these programs is the focus on promotion of movement quality, which has been used to rate physical ability, classify injury risk, track training progress, and assist with return to duty decisions. The Functional Movement Screen™ is arguably the most popular assessment tool used to this end, and its adoption in all branches was officially recommended in a 2011 Directive of the Joint Chiefs of Staff outlining the "Total Force Fitness Framework."[54]

When used as prediction or screening tool, the popularity of movement quality assessments has merit. However, as is often the case with tests of any kind, there has been a growing emphasis on identifying intervention programs designed to increase performance on the test itself.[15,80] This approach is potentially problematic as it assumes the existence of generic, optimal patterns of movement behavior.[77,155] In contrast to this perspective, variability signatures indicate that physiological systems are engaged in a constant attempt to adjust to a set of constraints which evolve continuously over time.[16,156] Given this continuous variation over time, invariant patterns of movement should not necessarily be encouraged.

Whereas clinical literature often describes movement itself as functional or dysfunctional, it may be more accurate to discuss movement variability in those terms. Here, it is the structure of the variability that is most relevant. Those individuals unencumbered by intrinsic constraints on movement are able to adapt more seamlessly to changing demands related to the task or environment.[156] In these cases, movement behaviors are complex in the sense that their output signals are rich in information content,[78] which can be viewed as a reflection of effective adaptation to a changing profile of constraints. To define a benchmark movement strategy as the optimal would be to discount the relevance of the underlying constraints, along with the functional variability that follows in a healthy, adaptable system. Therefore, rather than training to achieve such a benchmark in the movement itself, the more appropriate focus should be on identifying and modifying intrinsic constraints which limit the adaptability of the movement system.

Strength, balance, and range of motion are suggested to be important determinants of clinically rated movement quality.[8,9,22,23] It is additionally possible that these factors moderate decreases in movement quality related to external load carriage, a nearly ubiquitous task for tactical athletes. These findings offer clinical utility in that field-expedient screening methods can be used to gain insight into correlates of physical performance. Importantly, however, they may also be invoked to recommend movement pattern training as a method of promoting the associated qualities. The motivation for the present study was to complement previous findings which analyzed clinically scored movement tasks[29] with data which demonstrates the implications of intrinsic constraints on the dynamics of discrete, fundamental movement behaviors. In addition to providing evidence of complexity in fundamental movement strategies, quantifying the relevance of specific constraints in this way will support the design of constraint-based performance intervention programs. Therefore, the purpose of this investigation was to quantify the role of strength, balance, and range of motion in promoting dynamic complexity during discrete, fundamental movement tasks. Experimental manipulation of task constraints adds an important within-subjects dimension and additionally helps contextualize complexity metrics, which may be sensitive to data processing techniques.[157] As such, we also sought to determine the extent to which these factors mitigate loss of complexity related to external load carriage, an ecologically valid treatment for the population of interest, during the same tasks. We hypothesized that 1) higher levels of strength, balance, and range of motion will predict greater complexity, and 2) these same attributes would dampen loss of complexity associated with external loading.

5.2 Methods

This investigation was approved by the Institutional Review Board at The University of North Carolina at Greensboro. All participants provided written informed consent prior to beginning data collection.

5.2.1 Participants

Fifty recreationally active adults (25 male, 25 female; 22.98 $\pm$ 3.09 years, 171.95 $\pm$ 11.46 cm, 71.77 $\pm$ 14.03 kg) 18-34 years of age were recruited to participate in this investigation. This population was chosen to reflect the target demographic for first time military accessions. Participation was limited to those who exercise at least 90 minutes per week and who do not suffer from clinical conditions which may affect the outcome measurements. Such conditions included chronic instability of the joints of the lower extremity, any recent history of injury ($\leq$ 6 months prior to data collection), Ehlers-Danlos or joint hypermobility syndrome, uncorrected visual impairments not including astigmatism, vestibular disorders, peripheral sensory disorders, or any musculoskeletal condition requiring ongoing care from a licensed healthcare provider.

5.2.2 Procedures

Data collection proceeded as follows. Subjects first completed a Physical Activity Readiness Questionnaire (PARQ) and demographic information survey, and had their height and weight measured. Subjects then completed all clinical and laboratory assessments during a single data collection session in the following order:

(1) Balance
- Quiet Single Leg Stance

(2) ROM
- Apley Scratch Test
- Weight-Bearing Lunge Test
- Sit-and-Reach Test

(3) Cyclic Deep Squat, Hurdle Step, and Inline Lunge
- Conditions Counterbalanced by Subject

(4) Strength
- YMCA Bench Press Test
- Countermovement Jump

NB: The “Weight-Bearing Lunge Test” is so-named because it is a test of ankle range of motion in which dorsiflexion is assisted by the participant’s own bodyweight. No external loads were used for this test.

5.2.3 Single Leg Balance

Balance was assessed in quiet single leg stance for a period of 10 seconds.[51] Single leg stance was chosen as previous investigations have concluded that double leg standing may not be sensitive enough to detect meaningful differences in young, healthy populations.[29] Subjects completed all tests barefoot with eyes closed and hands on hips with a slight bend in hip and knee of the non-stance leg. Any of the following errors constituted a mistrial and was discarded: removal of hands from hips, touching the force plate or ground with the non-stance limb, touching the stance limb with the non-stance limb, flexion/extension/abduction of the non-stance hip in excess of 30°, lifting or turning of the stance foot, and opening the eyes. After completing a preliminary practice trial, subjects will continue balance testing until three successful test trials have been completed. Only data from the first completed trial was used for analysis.

Ground reaction force data during balance testing were sampled at 100 Hz using an AMTI Accusway force plate (AMTI, Watertown, MA). Center of pressure (COP) coordinates were then calculated in the anteroposterior and mediolateral directions in the Balance Clinic software package (AMTI, Watertown, MA). All testing was conducted using the non-dominant limb with the dominant limb defined as the leg the subject would use to kick a ball for maximum distance. Anteroposterior and mediolateral center of pressure data were combined to yield a resultant time series using a custom LabVIEW program (National Instruments, Austin, TX). Mean velocity of the resultant time series (CP_V) was entered as a predictor in our models. We used the raw signal so as to avoid the influence of filtering on velocity outcomes.[157]

5.2.4 Range of Motion

Range of motion was quantified using three validated clinical measures. The Apley scratch test quantifies range of motion in the shoulders and thoracic spine, the sit-and-reach test measures hip and trunk flexibility, and the weight-bearing lunge test measures dorsiflexion range of motion.

5.2.4.1 Apley scratch test

This test requires participants to attempt to touch fists behind their backs[41] One hand reaches down behind the neck while the other reaches up from behind the lower back. The average of the distances between attempts with the right and left hands on top was recorded as the final score.

5.2.4.2 Sit-and-Reach

A 30 .5 cm box was used to administer the sit-and-reach test following previously described guidelines.[40] Participants sat with the soles of their feet flush against the surface of the box. Keeping their knees straight and hands together, they reached as far along the box as possible and held that position until the examiner counted down from 6 seconds. Testing was continued until subjects achieved the same score on three consecutive trials.

5.2.4.3 Weight-bearing Lunge Test

Weight-bearing lunge test procedures followed those described in Hoch et al. 2011.[39] A piece of measuring tape was placed perpendicularly to a wall. The subject positioned his/her test foot such that it was lined up with the tape with the big toe touching the wall. The subject then incrementally moved the test foot backward from the wall and attempted to touch the wall with the ipsilateral knee while keeping the heel firmly planted. The maximum distance at which this task was successfully executed was recorded for both sides. The final score was an average of the left and right sides.

5.2.5 Strength

Participants were given the opportunity to complete additional warm up trials of the strength measures before data were recorded. While one-repetition maximum lifts are the gold standard for strength measurement, such tests were not appropriate in this investigation for reasons related to safety[38] and reliability.[133] Accordingly, we administered tests which closely estimate one-repetition maximum lifting capacity and can be administered safely and reliably in untrained populations.

5.2.5.1 Modified YMCA Bench Press Test

The YMCA bench press test is a paced, maximal-repetitions bench press effort with a fixed load assigned according to the subject's sex. Males performed the test with a load of 36.4 kg (80 lbs.) and women with a load of 15.9 kg (35 lbs.).[38] Subjects synchronized their repetitions to a metronome set to 60 beats per minute. Each beat corresponded to one half of a repetition (either lifting or lowering), resulting in a cadence of 30 presses per minute. The test was terminated when the subject was unable to maintain this pace. The number of repetitions performed was recorded as the final score.[38]

5.2.5.2 Countermovement Jump Peak Power

Peak countermovement jump power, derived from a vertical ground reaction force signal, is an excellent predictor of one-repetition maximum back squat capacity.[37,134] After familiarization, subjects completed three test trials. Each test trial was a distinct effort to jump as high as possible, followed by approximately 1 minute of rest. Subjects were allotted one practice trial and two test trials, each with a minimum of 60 seconds of rest between efforts. Participants held their hands on their hips while standing quietly on a Bertec force plate (4060-NC, Bertec Corporation, Columbus, OH). They were instructed to maintain the hands-on-hips position and complete a quick countermovement followed by maximum-height vertical leap. Vertical ground reaction force data were collected beginning 1 second prior to the countermovement, identified online by a falling threshold trigger. The signal was captured at 1000 Hz and low-pass filtered at 40 Hz. A vertical center of mass velocity time series was calculated via forward dynamics with the assumption that initial velocity was zero.[135]

The force and velocity time series data were multiplied to create a power time series. The peak of the pre-flight concentric movement phase was then identified and used for analysis.

5.2.6 Cyclic Deep Squat, Hurdle Step, and Inline Lunge

Dynamic postural control was assessed using the Functional Movement Screen™ Deep Squat, Hurdle Step, and Inline Lunge. Following familiarization, each test was performed for 5 continuous repetitions in both the weighted and control condition. The conditions were administered in randomized order. Subjects were instructed to move through a complete range of motion with each repetition of the test while adhering to standard Functional Movement Screen™ verbal cues. Aside from completing each trial in less than 30 seconds, the only instructions provided regarding cadence were to complete the repetitions at a comfortable, self-selected pace. The purpose of this approach was to avoid introducing artificial time domain constraints, which can substantially influence entropy estimates.[116]

Upon completing the fifth repetition, participants held their finishing position until 30 seconds of data had been collected. Ground reaction force data were sampled at 1000 Hz using a Bertec 4060-NC force plate and used to calculate center of pressure (COP) time series for each trial in The Motion Monitor software (Innovative Sports Training, Inc., Chicago, IL).

For the Deep Squat, participants were positioned horizontally such that both feet were entirely within the boundaries of the force plate. Feet were placed shoulder width apart and parallel. Holding a dowel directly overhead with arms fully extended, the participant proceeded to squat as deeply as possible while maintaining heel contact with the force plate.

The Hurdle Step was performed with an elastic hurdle set to the height of the participant's tibial tuberosity and placed across the center of the plate along its short axis. The base of the hurdle was slightly raised to prevent any contact with the plate. Subjects were instructed to begin with their feet together and toes just shy of touching the hurdle. Space on the plate was sufficient to contain each participant's entire base of support in double leg stance while still allowing room for the heel of the working leg to touch down within the borders of the plate on the far side of the hurdle.

Lastly, the Inline Lunge was performed with the front foot positioned in the center of the plate. The toes of the subject's back foot were placed in line behind the front heel by a distance equal to the height of the tibial tuberosity when standing. With each repetition, subjects were instructed to contact the plate gently with their back knee on a towel which had been placed just behind their front heel.

Each signal was programmatically truncated based on vertical ground reaction force thresholds which were used to identify the completion of five repetitions. Figure 1 depicts representative plots for each task, both ground reaction force and center of pressure, along with the threshold identifying where the signal was truncated.

The two-dimensional center of pressure time series were downsampled to 100 Hz and low-pass filtered at 12 Hz. All trials were centered such that the mean anteroposterior and mediolateral displacements were 0. The coordinate system of the Deep Squat trials was then rotated so that AP and ML directions were the same for all three tasks. The COP displacement time series were then separated into nine data-driven scales represented by intrinsic mode functions (IMFs; see details in next section) using multivariate empirical mode decomposition (MEMD). Finally, multivariate multiscale sample entropy (MMSE) was calculated for each cumulative IMF and summed to yield a composite MMSE index.

5.2.7 Sample Entropy, MMSE, and MEMD Enhanced MMSE

While entropy metrics are very common in postural control research,[158] special concerns had to be addressed for the present study. First, the sample entropy algorithm operates over a single timescale and will not fully characterize the complexity of cyclic movement behaviors over relatively few cycles. Second, it assumes that the signal being analyzed is stationary. An extension of sample entropy known as multiscale sample entropy calculates sample entropy over progressively coarse-grained copies of a given time series.[138] This method begins to address the problem of classifying signal complexity over multiple timescales, but can only be applied to relatively lengthy time series data as an increasing number of adjacent data points is averaged with each coarse-graining iteration. Further, the timescales created with this method are arbitrary and likely do not represent the most salient frequencies of the signal under investigation. The creators of the multiscale adaptation of sample entropy suggest the use of data-driven processing to address nonstationarities which may be present on different scales.[111] Ahmed et al. 2012 propose multivariate empirical mode decomposition to identify oscillatory scales, represented by intrinsic mode functions (IMFs), inherent to the original signal and use these functions to calculate multivariate multiscale sample entropy.[116] This has the advantage of preventing data loss associated with coarse-graining as the IMF length is equivalent to that of the original signal. Additionally, this method better characterizes complexity across multiple channels (in this case, the anteroposterior and mediolateral center of pressure series) over intrinsic scales which are quasi-stationary.[116]

5.2.8 Statistical Plan

One-tailed dependent t-tests were used to confirm the effect of condition on MMSE complexity index (MMSECI), represented by the summation of MMSE values across all IMF scales derived from the same trial. MMSECI for each outcome was then modeled using multiple linear regression. Our hypotheses call for tests of both global effects and effects specific to the weight vest condition. In order to account for these effects, a panel data structure with varying coefficients was used with the control condition specified as the baseline.[142] The coefficients specific to the weight vest condition are indicated by interaction terms ("*L").

While each of the variables of interest in this study is hypothesized to have beneficial/increasing effect on MMSE, two of these variables will have a negative coefficient as they are inversely related to the underlying constructs they measure. These are the Apley Scratch test and mean CP_V, in which lower scores are interpreted to be better.

Group lasso penalization was used to address the potential for bias in our relatively high-dimensional models. We followed the guidelines of Hess et al. 2013 in applying the group lasso algorithm with a varying-coefficients structure and[142] and interpret the resulting models to reflect the interaction of weight vest condition with the mediators outline above. The penalty parameter (Λmin) was determined based on minimum cross validation error (CVE). For each outcome, the appropriate Λmin was used to identify the group lasso solution and the selected models were subsequently analyzed as multiple linear regressions using the general linear model. Height, weight, sex, and age were included as nuisance variables, as were their interaction effects. In order to ensure their inclusion in the final models analyzed, these variables were excluded from penalization during the preceding steps. Statistical computations for this investigation were performed in R (The R Foundation) using the base and grpreg[144] packages.

5.3 Results

Table 13 shows results for one-sided tests of mean differences (H1: Control - Weight Vest > 0) for cyclic movement behavior tasks. For all tasks, complexity index was lower in the weight vest condition whereas coefficient of variation did not change. The relationship between MMSE and IMF scale for both conditions can be seen in Figures 2-6. Note that the EMD algorithm iterates until a predefined stoppage criterion is reached. In our analysis, 9 IMFs were generally retained by the EMD process and, therefore, used to calculate the complexity index. Summary statistics for each regression model are presented in Table 14. Table 15 shows the order in which variables were retained in the model as the penalty parameter was adjusted from a maximum value, in which all coefficients were shrunk to zero, to the value at which the final variable was retained. Finally, Table 16 shows coefficient values and their respective significance tests.

Table 13. Descriptives and Paired Differences for Cyclic Movement Task Complexity Index & Coefficient of Variation

	Complexity Index					Coefficient of Variation				
Outcome	W	C	Diff	t_{49}	p	W	C	Diff	t_{49}	p
DS	11.76	10.94	0.82	3.17	<0.01	3.18	2.84	0.34	0.37	0.36
HDom	11.71	11.00	0.77	5.21	<0.01	13.13	20.67	-7.54	-0.84	0.80
HNon	11.76	10.99	0.72	4.68	<0.01	2.05	1.07	0.98	1.47	0.07
IDom	14.23	13.42	0.81	3.68	<0.01	1.88	3.88	-2.00	-1.27	0.89
INon	14.96	14.14	0.82	3.33	<0.01	8.43	4.77	3.67	1.21	0.12

Means and paired t-tests for mean differences of Complexity Index and Coefficient of Variation (H1: Control - Weight Vest > 0). DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, Dom = Dominant Side, Non = Non-Dominant Side, C = Control Condition, W = Weighted Condition.

Each regression model was significant at the .05 level. Proportions of variance accounted for ranged from 0.19 to 0.36. Of the nuisance covariates, only age and weight were predictive of complexity outcomes. Weight was associated with lower complexity indices in all models while age had the same effect for the Deep Squat and Hurdle Step (Dominant) tasks. Specific to the external loading condition, weight was predictive of increased complexity in the Deep Squat task.

Table 14. Penalization and Final Model Summary Statistics (Manuscript II)

Outcome	Λ	Features	F	p	R^2
DS_{CI}	0.143	11	$F_{(11,88)} = 3.94$	<0.01	0.25
HNon_{CI}	0.135	9	$F_{(9,90)} = 7.30$	<0.01	0.36
HDom_{CI}	0.143	9	$F_{(9,90)} = 7.21$	<0.01	0.36
INon_{CI}	0.162	9	$F_{(9,90)} = 6.01$	<0.01	0.31
IDom_{CI}	0.199	9	$F_{(9,90)} = 3.66$	<0.01	0.19

Penalized and Unpenalized Model Summaries. DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, Dom = Dominant Side, Non = Non-Dominant Side, CI = Complexity Index.

Predictors related to range of motion were significant in the Deep Squat task only, but with differential effects. Greater range of motion in the weight-bearing lunge test was associated with higher complexity whereas greater sit-and-reach range of motion was associated with lower complexity.

Table 15. Variable Selection Order as a Function of Λ (Manuscript II)

Rank	DS_{CI}	$HNon_{CI}$	$HDom_{CI}$	$INon_{CI}$	$IDom_{CI}$
1	Age	Age	Age	Age	Age
2	Age*L	Age*L	Age*L	Age*L	Age*L
3	Ht	Ht	Ht	Ht	Ht
4	Ht*L	Ht*L	Ht*L	Ht*L	Ht*L
5	Sex	Sex	Sex	Sex	Sex
6	Sex*L	Sex*L	Sex*L	Sex*L	Sex*L
7	Wt	Wt	Wt	Wt	Wt
8	Wt*L	Wt*L	Wt*L	Wt*L	Wt*L
9	WL	RCOP	WL	WL	RCOP
10	RCOP	WL	SR††	SR	Y
11	SR	SR	Y††	Ap*L†	SR
12	SR*L	CJ††	RCOP	RCOP††	WL
13	Ap	Y*L†	CJ	Ap	Ap*L†
14	Y	Ap*L	Ap*L	WL*L	CJ††
15	Y*L	RCOP*L	CJ*L	Y*L	WL*L
16	CJ	Ap	Ap	RCOP*L	SR*L
17	WL*L	WL*L	Y*L	CJ	Ap
18	RCOP*L	CJ*L	RCOP*L††	SR*L	CJ*L††
19	Ap*L	SR*L	WL*L††	Y	L††
20	CJ*L	L	SR*L	CJ*L	RCOP*L
21	L	Y	L	L	Y*L

Relative importance of covariates shown as a function of increasing the penalty parameter (Λ). The nuisance variables (age, sex, height, and weight), along with their interaction effects, were not penalized and are therefore present in all models. Subsequent variables which share a superscript (†,††) were selected in the same iteration. L = Load-Bearing/Weighted Vest Condition, DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, Dom = Dominant Side, Non = Non-Dominant Side, CI = Complexity Index, WL = Weight-Bearing Lunge Test, CP_V = Resultant Center of Pressure Velocity, Y = YMCA Bench Press Test, SR = Sit-and-Reach Test, Ap = Apley Scratch Test, CJ = Countermovement Jump.

5.4 Discussion

As expected, MEMD-enhanced MMSE associated with dynamic postural control tasks decreased when those tasks were performed with an external load. While not the explicit purpose of this study, this is a novel finding which may have implications for load carriage in tactical occupations. The observed differences in complexity index did not coincide with changes in coefficient of variation, suggesting that the more sophisticated analyses allow us to capture more information related to movement behavior in the two conditions. Relatively few effects related to our covariates of interest reached the threshold for statistical significance, and none of the condition-specific effects was selected by the group lasso. Therefore, with the exception of the Deep Squat, our results do not provide sufficient evidence to support our hypotheses.

While our a priori hypotheses were not supported, there does appear to be a pattern suggesting that certain covariates related to balance and range of motion may be relevant. In each model, either weight-bearing lunge scores or mean CP_V was the first non-nuisance predictor retained. For four of the outcomes, the corresponding p-values—which are two-tailed—fall below 0.10. Availability of data in clinical research involving human subjects often limits an investigator's ability to find small, but meaningful effects. Thus, despite a lack of significance at the $\alpha = 0.05$ level, our results could be interpreted to justify further research in this vein.

The decreases in MMSECI in this dataset may be interpreted differently, depending on one's interpretation of complexity. The notion of "loss of complexity" commonly associated with decreases in entropy statistics like SampEn or ApEn[159] has a somewhat longer tradition in the literature. Aging, injury, or disease may

Table 16. Coefficients for MMSE in Cyclic Movement Task Models

DV	Var.	Coef.	t	p	DV	Var.	Coef.	t	p
DS_{CI}	Int	14.49	4.89	0.00	$HDom_{CI}$	Int	14.25	8.35	0.00
	Sex	0.21	0.38	0.70		Sex	0.44	1.29	0.20
	Sex*L	-0.90	-1.31	0.20		Sex*L	0.19	0.42	0.68
	Age	-0.14	-1.84	0.07		Age	-0.11	-2.20	0.03
	Age*L	-0.05	-0.47	0.64		Age*L	0.07	0.96	0.34
	Ht	0.02	0.81	0.42		Ht	0.01	0.55	0.58
	Ht*L	-0.01	-0.99	0.32		Ht*L	-0.01	-0.52	0.60
	Wt	-0.05	-2.65	0.01		Wt	-0.03	-2.67	0.01
	Wt*L	0.06	2.16	0.03		Wt*L	-0.02	-1.28	0.20
	WL	0.10	2.47	0.02		WL	0.05	1.90	0.06
	SR	-0.03	-1.95	0.05					
	CP_V	0.20	1.70	0.09	$IDom_{CI}$	Int	13.79	5.50	0.00
						Sex	0.10	0.20	0.85
$HNon_{CI}$	Int	11.52	6.61	0.00		Sex*L	-0.33	-0.49	0.63
	Sex	0.23	0.65	0.52		Age	0.13	1.85	0.07
	Sex*L	-0.03	-0.07	0.94		Age*L	-0.12	-1.25	0.22
	Age	-0.03	-0.68	0.50		Ht	0.03	1.36	0.18
	Age*L	-0.07	-1.05	0.30		Ht*L	0.00	0.00	1.00
	Ht	0.02	1.57	0.12		Wt	-0.09	-4.97	0.00
	Ht*L	0.00	0.28	0.78		Wt*L	0.04	1.35	0.18
	Wt	-0.05	-3.69	0.00		WL	-0.05	-1.27	0.21
	Wt*L	0.01	0.36	0.72					
	RCOP	0.14	1.86	0.07	$INon_{CI}$	Int	14.40	5.04	0.00
						Sex	-0.59	-0.99	0.33
						Sex*L	0.23	0.31	0.76
						Age	0.08	0.99	0.33
						Age*L	0.06	0.55	0.58
						Ht	0.02	0.90	0.37
						Ht*L	-0.02	-1.45	0.15
						Wt	-0.07	-3.09	0.00
						Wt*L	0.02	0.69	0.49
						RCOP	0.22	1.79	0.08

Predictor coefficients and t-statistics for each model. CI = Complexity Index, DS = Deep Squat, H = Hurdle Step, I = Inline Lunge, Dom = Dominant Side, Non = Non-Dominant Side, CI = Complexity Index, WL = Weight-Bearing Lunge Test, RCOP = Resultant Center of Pressure Velocity, SR = Sit-and-Reach Test.

constrain the adaptability of physiological systems such that behavioral responses to the environment are limited. In this scenario we may expect to observe an overly regular or predictable pattern of variability in the system's output, such as those exhibited by Parkinson's disease patients during postural control tasks, which would lead to a lower complexity index value relative to less constrained populations. Complexity can also refer more specifically to the information content of a behavioral signal, which is maximized not at the extreme of irregularity, but somewhere in between pure determinism and pure randomness.[78,109] When used in this sense, increased complexity is not necessarily synonymous with increased irregularity (i.e. higher entropy).

If complexity follows an inverted-U pattern as a function of MMSE, understanding the implications of a given change can be difficult. Further compounding this issue, dynamical systems metrics are often sensitive to equipment and processing techniques,[157] making comparison across studies challenging. Researchers commonly address this by including a control group or condition in their experiments to serve as a point of comparison. This point of comparison most often indicates that injury, illness, or increased difficulty of the experimental task are associated with lower entropy values.[114] Our data appear to follow a similar trend in which MMSECI is lower during the more challenging experimental conditions. Unexpectedly, however, the traits which we hypothesized would promote behavioral complexity, both in general and with regard to the weight vest condition specifically, were not consistent in the pattern of observed effects. For example, our two significant range of motion variables had diverging effects in the Deep Squat task with higher sit-and-reach range of motion predicting lower MMSECI. Our quiet standing balance variable, though not significant, also appears to have a negative influence on complexity for the models in which it was selected. The possibility that our presupposed meaning of a decrease in MMSECI—specifically, that such a decrease reflects less adaptable movement behaviors—is mistaken should therefore be entertained.

Much of the software used for our analysis was developed by Ahmed et al.,[116] who report higher quiet standing COP complexity indices in young subjects compared to elderly subjects. They additionally show that complexity indices for stride interval are lowered by experimental constraints, in this case a pacing metronome. In another investigation which analyzed COP displacement data over few repetitions of a cyclical task, RQA entropy was shown to increase as a function of task difficulty.[159] (Note: RQA entropy and Sample Entropy move in opposite directions.[158]) Combining these observations with the generally negative influences of age and weight in our data, it is reasonable to conclude that the load-related decreases in MMSECI are most likely a reflection of a less complex, less adaptable pattern of postural control.

It is unclear then why separate indicators of what are traditionally considered to be “good” traits would point in opposite directions with respect to their effect on MMSECI. It may be the case that the covariates which tended to limit complexity of cyclic postural control tasks in our study are not universally supportive of healthy, adaptable movement behaviors. This could be a function of the requirements of the task, competitive relationships between adaptations, or potentially both, and may indicate that movement behaviors are best considered separately rather than in relation to a single latent movement quality trait.

One important limitation that should be kept in mind regarding the present study is that our methods manipulated constraints through a limited range of possibilities. Because the behavior of a complex system is not confined to linear changes as a function of its inputs, a more complete topological map requires a rigorous approach in which constraints are manipulated through a wide range of configurations.[16,160] Practical limitations—fatigue, most importantly—prevented us from testing our hypotheses under a comprehensive range of conditions. Another notable limitation relates to our comparison of conditions. Specifically, while the participants were intentionally not constrained in the time domain, the weight vest had non-negligible effects on trial duration in some cases. In the extreme, this could imply that the conditions are representative of different behaviors entirely as opposed to similar behaviors with varying constraint profiles. This seems unlikely in our dataset. At the very least, however, the complexity of the signal could be affected by variations in trial length. While MEMD does not ensure that the scales being analyzed are the same within or between individuals, it does allow us to analyze the most relevant oscillatory features of a given signal, and should therefore be informative nonetheless.

With the exception of the Deep Squat, we cannot draw firm conclusions regarding the role of strength, balance, and range of motion in promoting complexity during cyclic postural control behaviors based on our data. In all models, inclusive of the Deep Squat, no evidence was found to support the conclusion that these variables modify the effects of external loading on movement complexity during the same tasks. Notwithstanding, further research in this area is warranted as nonsignificant trends may be driven by clinically meaningful relationships. If confirmed, such relationships could have important implications for current approaches to assessment and training of movement quality, in which functional norms are most often defined based on the medical model. In this context, overreliance on the medical model could lead to interventions which are ineffective or potentially contraindicated. For example, recommending a movement pattern exercise that depends on a high range of motion to a relatively inflexible individual may be more likely to encourage a novel, unintended motor solution than it is to "teach" that individual an increased range of motion. The result could be a set of undesirable training adaptations or even a training-related injury.

5.5 Conclusion

The findings of this study indicate that standardized external loading is associated with decreased multiscale complexity during commonly used clinical assessments of functional movement. These changes may occur without corresponding differences in linear summary measures of variability. Both age and weight are generally associated with lower movement complexity during these tasks. In the Deep Squat task, there is evidence to support an association between movement complexity and range of motion; however, the direction of this range of motion effect depends on the body site. Specifically, dorsiflexion range of motion is associated with increased complexity whereas the opposite is true of sit-and-reach range of motion.

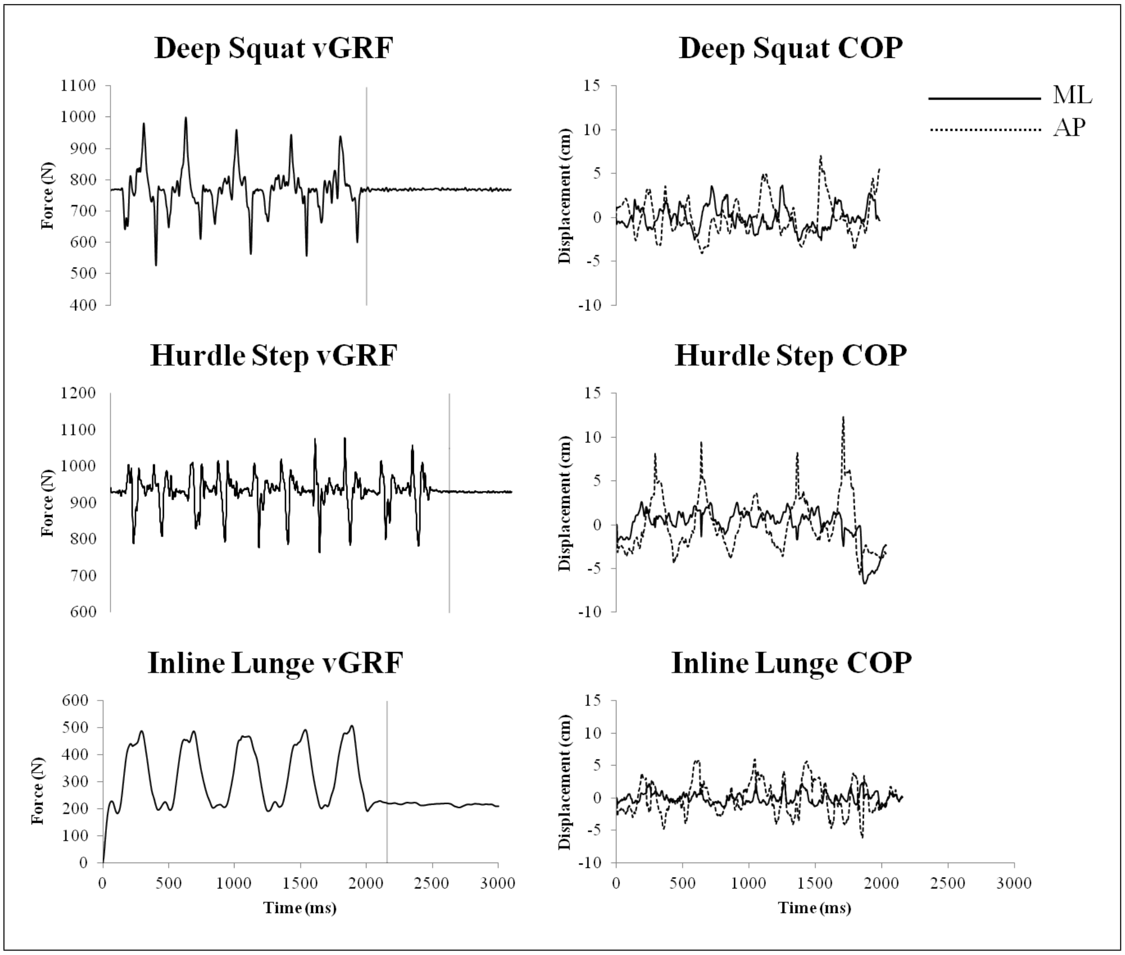

Figure 1. Representative Plots of Vertical Ground Reaction Force and Center of Pressure Time Series for Cyclic Movement Tasks. In the COP plots, dashed lines represent anteroposterior COP displacement while solid lines represent mediolateral COP displacement. The following threshold definitions were used to define the end of the task and the point at which the time series was truncated. Deep Squat End Point = 1 second after the point at which the vGRF first falls below 6 SDs of the baseline mean for 20 or more consecutive data points. Baseline defined as the last .5 seconds of the vGRF time series after low-pass filtering at 5 Hz. Hurdle Step End Point = 1 second after signal first exceeds 5 SDs of the baseline mean for 10 or more consecutive data points. Baseline defined as the last .5 seconds of the vGRF time series after low-pass filtering at 2 Hz. Inline Lunge End Point = 2 seconds after the index at which signal first exceeds baseline mean by 20 SDs for more than 10 consecutive points. Baseline defined as the last 3.5 seconds of the vGRF time series after low-pass filtering at 1 Hz.

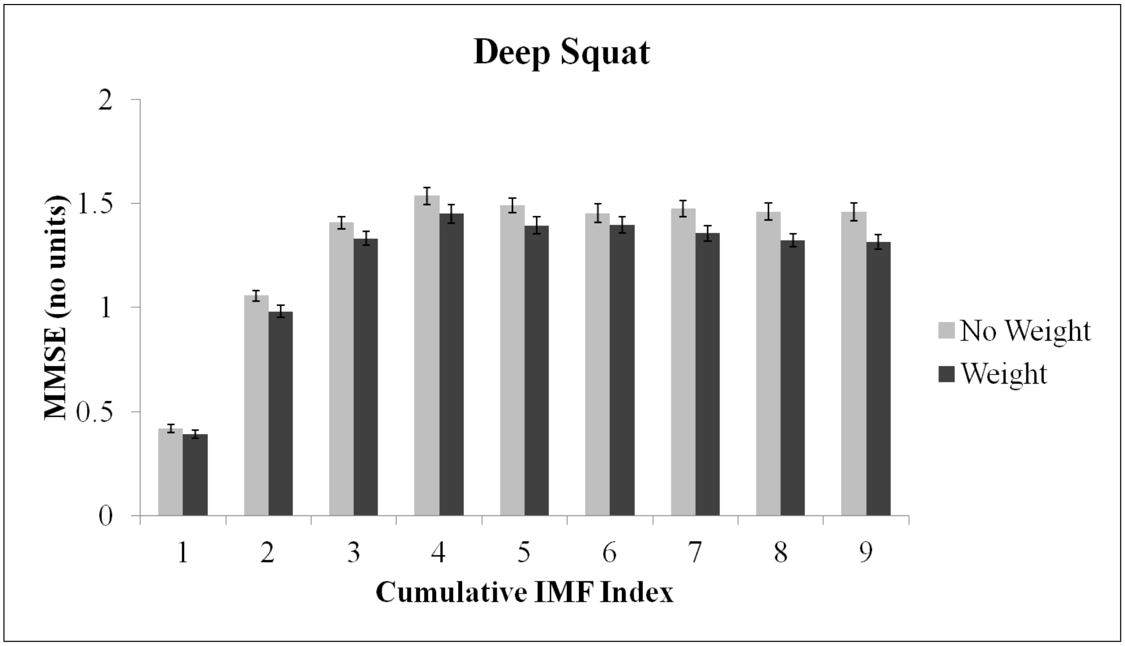

Figure 2. Deep Squat MMSE Value for Each Scale, Represented by Cumulative IMFs

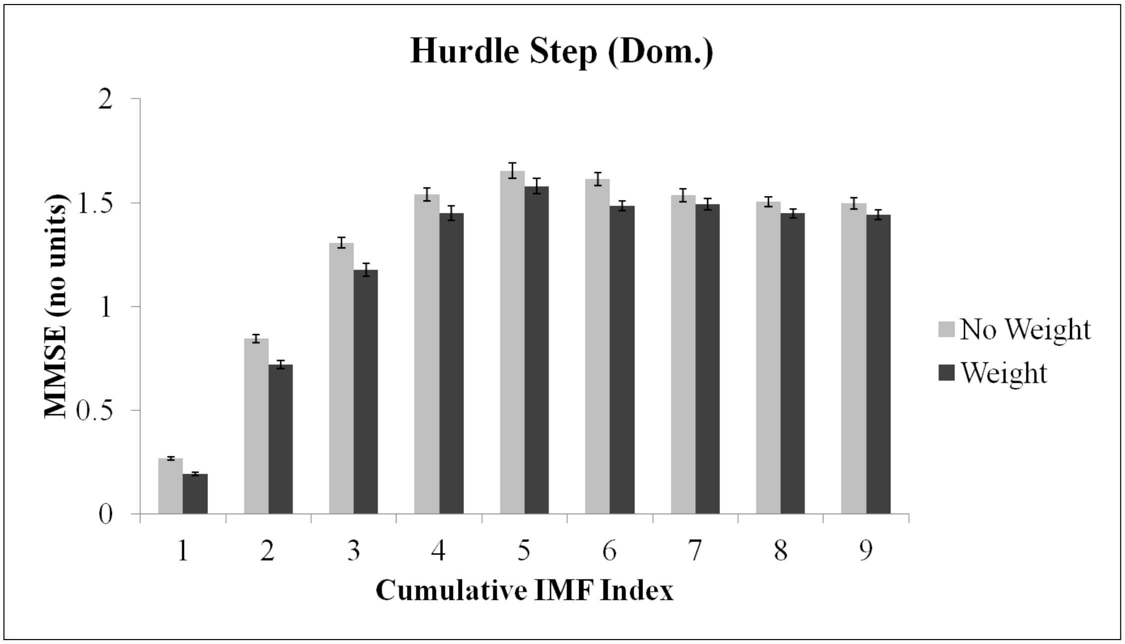

Figure 3. Dominant Side Hurdle Step MMSE Value for Each Scale, Represented by Cumulative IMFs

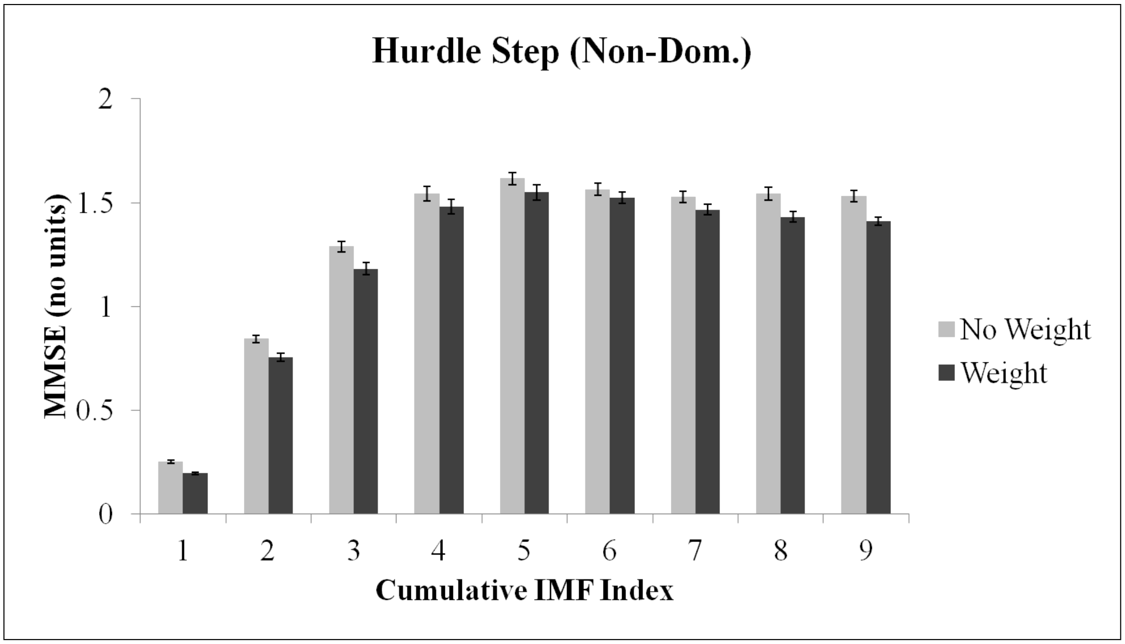

Figure 4. Non-dominant Side Hurdle Step MMSE Value for Each Scale, Represented by Cumulative IMFs

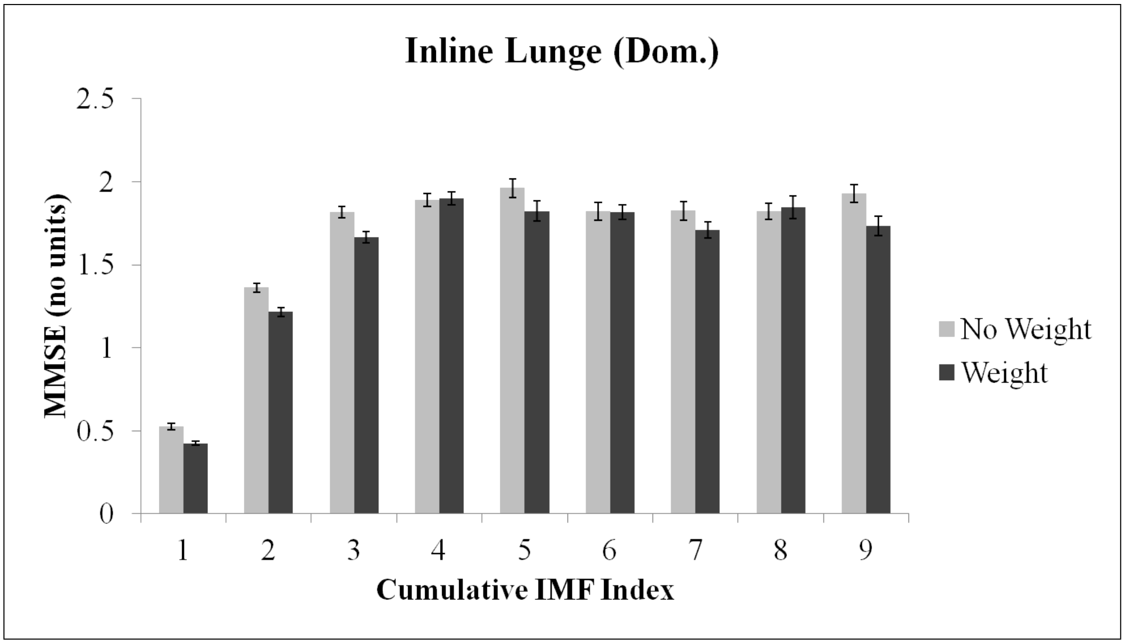

Figure 5. Dominant Side Inline Lunge MMSE Value for Each Scale, Represented by Cumulative IMFs

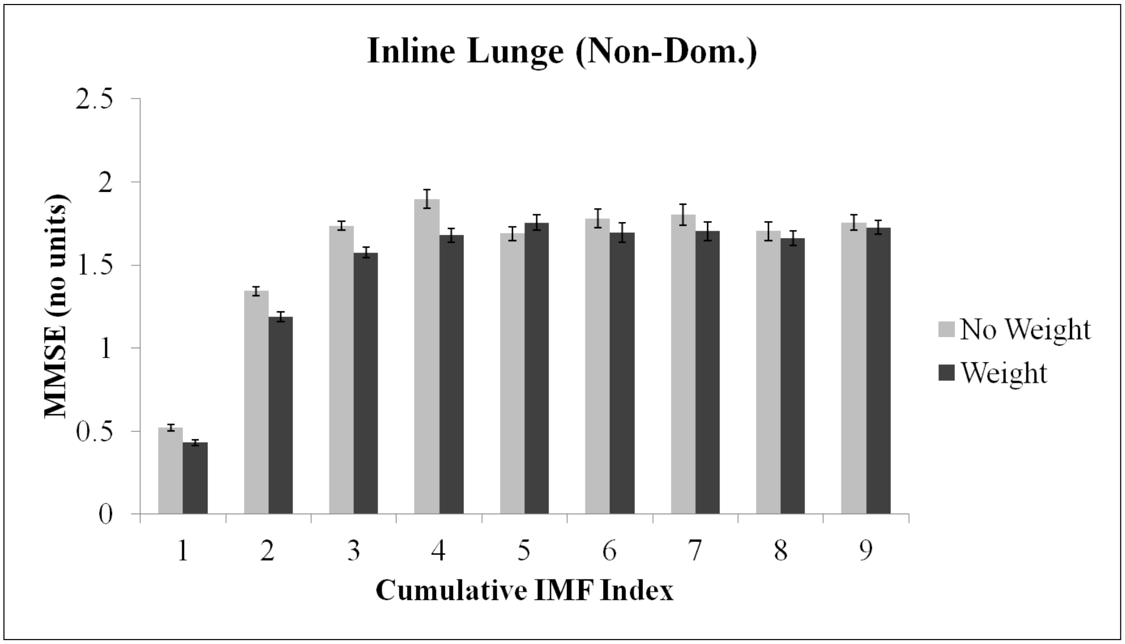

Figure 6. Non-dominant Side Inline Lunge MMSE Value for Each Scale, Represented by Cumulative IMFs

CHAPTER VI

MANUSCRIPT III

6.1 Introduction

Recent recruiting cycles have been extraordinarily successful for our nation's military.[161] Accession goals were often exceeded and strong candidates required to wait for a vacancy. Such a favorable recruiting environment dampened the need for any pre-accession performance screening system. It is anticipated, however, that the recruiting climate of the near future will present greater challenges.[55] The combination of defense budget cuts and economic alternatives for the military's recruitment population place a renewed emphasis on minimizing preventable attrition due to substandard fitness or injury.

Clinical movement screens have seen a tremendous growth in use during the past decade, in large part because they are cost effective and field-expedient.[8–10,98] In addition to classifying individuals by injury risk, movement screens have also been applied to predict performance in tactical athlete populations.[95] Most research thus far has failed to show a relationship between clinically rated movement scores and performance outcomes, a lack of association which likely stems two sources. First, relatively undemanding movement tests do not present a challenge sufficient to highlight deficiencies relevant to athletic performance. Accordingly, it has been suggested that adjusting screening practices to increase specificity or difficulty may increase the likelihood of detecting deficiencies clinically.[24,29,162] Another limitation of these instruments relates to methods for scoring and analyzing data. Item scores are often rated on ordinal scales which are summed into a total. While such a parsimonious representation of test performance has its appeal, this practice is appropriate only if the construct underlying the total score is unidimensional. With respect to clinical movement screens, there is strong evidence to suggest this is not the case.[96,97] More detailed information can be found in the item scores themselves, although certain considerations must be addressed concerning their analysis. In addition to a rank order structure which is difficult to accommodate in linear or logistic regression,[140] direct analysis of component score data in existing clinical movement screens would substantially increase the dimensionality of a prediction model.

A recent investigation in young, recreationally active non-servicemembers showed that Functional Movement Screen™ tests under load show increased predictive validity with respect to criterion performance measures specific to the tactical athlete.[29] In the same context, this study also demonstrated the utility of regularization techniques designed to accommodate high-dimensional regression problems with ordered predictors. Combining these two modifications in approach may move use closer to a field-expedient, feasible means of conducting pre-accession screening for injury risk and performance deficits.

While the study in question showed promising results, certain limitations affect our ability to generalize the findings. First, it was conducted using a relatively small sample size (n = 19). Second, this sample did not contain an even balance of men and women. The purpose of the present investigation was to replicate these findings while addressing the noted limitations. These efforts will underscore the previously observed increase in the predictive validity of a convenient, field-expedient method of evaluating performance potential.

6.2 Methods

Data were collected in a laboratory setting by a single investigator experienced in the required measurement techniques. Participation was limited to individuals between 18-34 years of age in order to reflect the recruitment pool for military and tactical occupations. Subjects were additionally required to accumulate a minimum of 90 minutes/week of physical activity. All subjects provided written consent to participate and completed a physical activity readiness questionnaire (PAR-Q) before data collection.

6.2.1 Procedures

This study was approved by the Institutional Review Board at UNC-Greensboro. A total of fifty recreationally active adults, 25 male and 25 female, participated in the study (22.98 ± 3.09 years, 171.95 ± 11.46 cm, 71.77 ± 14.03 kg). Participation was limited to adults 18-34 years of age without recent (< 6 months) injury and who accumulated a minimum of 90 minutes of physical activity per week. Subjects reported to the laboratory for a single data collection session. Following consent and completion of the PAR-Q, the Functional Movement Screen™ was administered under two conditions in randomized order. Finally, participants completed a battery of physical performance tests.

The data presented here were collected in conjunction with other measures as part of a larger project. In order to control for the effect of fatigue, subjects reported perceived rate of exertion using a standard Borg scale (6-18) at predetermined intervals throughout the data collection. This scale was designed to provide a valid method for comparing subjective exertion levels between individuals.[163]

6.2.2 Functional Movement Screen^{TM}

Following a familiarization round, the Functional Movement Screen^{TM}[8,9] was administered both with and without an 18.10 kg weight vest (MiR Vest Inc., San Jose, CA). This is comparable to loads used in previous investigations on the topic of tactical athleticism,[35,42] as well as those used in clinical screens designed to predict physical performance.[102] All testing was conducted by an experienced investigator with established reliability in each of the FMS component tests (see Table 17). Scores for component tests were assigned based on a 1-3 scale according to the following criteria outlined as part of the FMS protocol[8,9]: 1, subject was unable to complete the movement; 2, subject was able to complete the task with errors noted; 3, subject was able to complete the task without error.

Table 17. Reliability Data

	Kappa	z	p	Reliability
Deep Squat	0.67	2.88	<0.01	Good
Hurdle Step	0.78	2.54	0.01	Good
Inline Lunge	1.00	3.16	<0.01	Very Good
Shoulder Mobility	0.78	3.00	<0.01	Good
Active Straight Leg Raise	1.00	3.16	<0.01	Very Good
Trunk Stability Push Up	1.00	3.16	<0.01	Very Good
Rotary Stability	0.74	2.42	0.02	Good

6.2.3 Physical Performance Tests

Following completion of the FMS in both testing conditions, participants performed a 10-minute cycle ergometer warm up during which they instructed to target an RPE of 13 (“Somewhat Hard”). Instructions were to complete each individual test as quickly as possible. All tests were administered in the same order:

(1) Agility T-Test

(2) 5 x 27.43 meter (30 yard) sprints

(3) 400 meter run

(4) Mobility for Battle Assessment (MOB)

(5) Partner Rescue Simulation Task

Completion time for both the Agility T-Test and sprints was recorded using an infrared timing gate (Brower Timing Systems, Draper, UT). The Agility T-Test was administered according to previously described methods. Subjects began on the starting mark with one foot positioned on a start-on-release trigger. When directed, subjects performed the following sequence: forward sprint 9.14 m (10 yds), right side-shuffle 4.57 m (5 yds), left side-shuffle 9.14 m (10 yds), right side shuffle 4.57 m (5 yds), back peddle 9.14 m (10 yds). The timing gates were applied similarly in the 27.43 m sprints, each of which was separated by approximately 60 seconds of rest.

Owing to logistical restrictions, completion time for the remaining tests was recorded using a handheld stopwatch. Courses for the 400 m run and Mobility for Battle[43] were mapped with cones in an indoor gymnasium. The 400 m run was administered as a series of 4.5 laps around the periphery of the gym space. The

Mobility for Battle (MOB), designed as a multifaceted test incorporating several soldier-relevant field maneuvers, was organized in stations according to the methods described in Crowder et al.[43] Participants were allotted up to 5 minutes of recovery time upon finishing each of the 400 m and MOB tests.

The final test was a simulated partner rescue, in which subjects were required to drag a 68.04 kg (150 lbs.) dummy across a distance of 45.72 m (50 yds). The dummy was fashioned from sandbags wrapped in carpet with a handle attached to one end. Completion time was recorded after the final bag crossed the finishing line.

6.2.4 Statistics

Several researchers have noted the limitations of analyzing the FMS composite score.[29,96,97] The item scores themselves are likely the better source of information, though extra care must be taken to select appropriate models in a high-dimensional predictor space. Further, more of this information can be preserved by using methods which account for the ordinal structure of the scores. Each of these challenges can be addressed via penalization. Application of regression penalization algorithms is common in, for example, genome-wide association studies (GWAS), in which the number of predictors often greatly exceeds the number of observations. The effect of penalization is to shrink large coefficients and thereby reduce bias toward data characteristics which are unique to a given sample. Additional penalization can be applied to smooth the differences between successive levels of a predictor.[140] Thus, these techniques offer an attractive solution to the problems that arise when analyzing FMS item score data.

In our analyses, a group lasso penalty was first applied to select an appropriate model. Differences between neighboring levels within the retained predictors were then smoothed using a second penalization algorithm. The same penalty parameter (Λ) was used in each step, identified as the value of Λ which minimized cross-validation error in the group lasso. The final step after model selection and smoothing was to construct bootstrap 95% confidence intervals of the estimated coefficients using the bias-corrected and accelerated method. Each of these steps was completed using R v3.1.0 with ordPens 0.3-[140] and grpreg 2.8-[164] packages.

6.3 Results

Model summaries are presented in Table 18 while smoothed coefficients and their bootstrap confidence intervals are presented in Table 19 and Table 20. A positive R^2 was observed in only two of the models corresponding to the unweighted condition. These were the Sprint and MOB models, in which FMS item scores respectively accounted for 11% and 19% of the variance. In contrast, positive R^2 values were observed in all models featuring scores from the weighted condition, with variance explained ranging from 11% - 29%.

In the unweighted condition, higher Trunk Stability Push Up scores were predictive of faster completion times for the Sprint and MOB tests. A similar influence was observed for the remaining 3 performance outcomes, though variation in scores was not explained at the model level. Higher Trunk Stability Push Up scores from the weighted condition were predictive of faster completion times for all measures. Additionally in the weight vest condition, a Hurdle Step score of 3 was predictive of faster Agility T-Test times while a score of 2 or 3 was predictive of faster 400 m times. Higher weighted Inline Lunge scores were also associated with performance, with a score of 3 predicting faster Sprint times and a score of 2 or 3 predicting faster MOB times. Interestingly, a weighted Inline Lunge score of 3 was also predictive of slower time to completion on the partner rescue simulation. A similar inverse relationship was observed between 400 m times and scores of 2 or 3 in the weighted Shoulder Mobility test. Finally, a weighted Deep Sqaut score of 2 was predictive of faster sprint times while a score of 2 or 3 was predictive of faster partner rescue times.

Table 18. Penalization and Final Model Summary Statistics (Manuscript III)

	Unweighted Condition				Weighted Condition			
	Λ	CVE	R^2	Features	Λ	CVE	R^2	Features
Agility	0.31	2.48	0.00	1	0.16	1.84	0.22	3
	0.00	3.14	0.00	7	0.00	2.22	0.05	7
Sprint	0.09	0.25	0.11	2	0.02	0.20	0.29	7
	0.00	0.29	0.00	7	0.00	0.21	0.27	7
400m	4.12	285.68	0.00	1	1.01	223.15	0.17	6
	0.00	294.19	0.00	7	0.00	227.78	0.16	7
MOB	4.46	637.21	0.19	1	2.99	559.84	0.29	3
	0.00	720.13	0.09	7	0.00	606.77	0.23	7
RSQ	1.27	90.06	0.00	3	1.46	78.21	0.11	3
	0.00	93.32	0.00	7	0.00	95.79	0.00	7

Summary statistics for group lasso penalized models selected separately for the weighted and unweighted conditions. Unpenalized summaries are presented as well for comparison.

Table 19. Coefficient Summaries for Agility, Sprint, and 400m Outcomes

Unweighted Condition

		T-Agility		Sprint		400m	
Test	Level	Coef	CI_{95}	Coef	CI_{95}	Coef	CI_{95}
DS	2	–	–	–	–	–	–
	3	–	–	–	–	–	–
HS	2	–	–	–	–	–	–
	3	–	–	–	–	–	–
ILL	2	–	–	–	–	–	–
	3	–	–	–	–	–	–
SM	2	–	–	–	–	–	–
	3	–	–	–	–	–	–
ASLR	2	–	–	-0.01	(-1.63, 0.53)	–	–
	3	–	–	0.21	(-0.26, 0.94)	–	–
TSPU	2	0.41	(-0.91, 4.27)	0.00	(-0.63, 1.10)	-1.21	(-12.10, 8.57)
	3	-1.40	(-1.97, -1.78)*	-0.64	(-0.94, -0.94)*	-11.62	(-30.41, -12.86)*
RS	–	–	–	–	–	–	–

Weighted Condition

		T-Agility		Sprint		400m	
Test	Level	Coef	CI_{95}	Coef	CI_{95}	Coef	CI_{95}
DS	2	–	–	-0.23	(-1.12, -0.02)*	-1.95	(-16.19, 7.51)
	3	–	–	-0.15	(-0.95, 0.32)	-6.02	(-28.49, 3.61)
HS	2	-1.13	(-6.49, 0.00)*	-0.04	(-2.54, 0.34)	-28.84	(-50.70, -29.29)*
	3	-1.95	(-5.96, -1.51)*	-0.35	(-3.00, 0.08)	-37.03	(-49.89, -42.12)*
ILL	2	–	–	-0.29	(-2.53, 0.15)	–	–
	3	–	–	-0.43	(-2.69, -0.02)*	–	–
SM	2	0.43	(-0.26, 1.96)	0.25	(0.03, 1.16)*	6.74	(1.75, 19.52)*
	3	-0.20	(-2.90, 1.89)	0.21	(-0.28, 2.49)	16.89	(6.47, 51.08)*
ASLR	2	–	–	0.12	(-0.66, 1.12)	–	–
	3	–	–	0.15	(-0.63, 1.25)	–	–
TSPU	2	-1.62	(-2.08, -2.06)*	-0.43	(-0.95, -0.36)*	-12.67	(-24.17, -13.18)*
	3	-1.57	(-2.49, -1.79)*	-0.50	(-1.51, -0.46)*	-11.19	(-25.65, -7.67)*
RS	–	–	–	–	–	–	–

Summary statistics for group lasso penalized models selected separately for the weighted and unweighted conditions. Unpenalized summaries are presented as well for comparison.

Table 20. Coefficient Summaries for MOB & RSQ Outcomes

Unweighted Condition

		MOB Obstacle		RSQ	
Test	Level	Coef	CI_{95}	Coef	CI_{95}
DS	2	–	–	–	–
	3	–	–	–	–
HS	2	–	–	–	–
	3	–	–	–	–
ILL	2	–	–	0.00	(0.00, 0.00)
	3	–	–	3.54	(1.50, 9.01)*
SM	2	–	–	-1.36	(-34.51, 6.22)
	3	–	–	2.98	(-7.19, 12.92)
ASLR	2	–	–	–	–
	3	–	–	–	–
TSPU	2	-6.30	(-37.33, 3.88)	-1.39	(-10.34, 6.52)
	3	-31.05	(-52.60, -52.60)*	-6.01	(-14.65, -4.63)*
RS		–	–	–	–

Weighted Condition

		MOB Obstacle		RSQ	
Test	Level	Coef	CI_{95}	Coef	CI_{95}
DS	2	–	–	-4.40	(-16.25, -2.37)*
	3	–	–	-4.57	(-17.39, -1.03)*
HS	2	–	–	–	–
	3	–	–	–	–
ILL	2	-25.34	(-78.57, -14.32)*	–	–
	3	-29.59	(-80.03, -21.00)*	–	–
SM	2	–	–	–	–
	3	–	–	–	–
ASLR	2	–	–	–	–
	3	–	–		
TSPU	2	-28.81	(-28.88, -28.88)*	-9.39	(-9.90, -9.90)*
	3	-30.54	(-28.88, -28.88)*	-10.83	(-11.54, -11.54)*
RS		–	–	–	–

Summary statistics for group lasso penalized models selected separately for the weighted and unweighted conditions. Unpenalized summaries are presented as well for comparison.

6.4 Discussion

These findings parallel previously reported increases in predictive validity of FMS item scores related to testing under a standardized external load.[29] The combination of the present findings with similar results derived from an unrelated sample establishes strong evidence in support of this effect. While the expected improvement in prediction was observed in the present study, it is interesting to note that the test items driving that improvement were not necessarily the same as those observed previously. Among the items that overlap, weighted Trunk Stability Push Up predicted better performance in all outcomes, as did the unweighted Trunk Stability Push Up for the Sprint. Also common to both datasets, weighted Hurdle Step predicted faster 400 m run times. The previously published data contains several unique effects from the weighted condition. These includes relationships between the Deep Squat and 400 m run, the Deep Squat and MOB, and the Hurdle Step and MOB. Two unique performance-inhibiting effects were also noted. Specifically, a 3 on the Inline Lunge or Shoulder Mobility tests predicted slower times in the Partner Rescue task.

Several unique effects were also observed in the current data set. These include the relationships between weighted Inline Lunge and sprint speed, weighted Inline Lunge and MOB times, weighted Deep Squat and Partner Rescue times, and lastly the relationship between unweighted Trunk Stability Push Up and all outcomes other than sprint speed. Performance-inhibiting associations unique to the current dataset included those between weighted Shoulder Mobility and sprint speed, weighted Shoulder Mobility and 400 m times, and unweighted Inline Lunge and

Partner Rescue times. Because the current results are based on a larger sample size which is split evenly between men and women.

In most cases, the hypothesized relationship between clinically rated movement behaviors and physical performance has eluded investigators.[27,162] The difficulty in demonstrating this association may be rooted in the relatively low demand of most screening tools, low scoring resolution, or improper analysis. We have taken steps to address each of these concerns with the results suggesting that the relationship can be observed when movement is evaluated under load. These findings may have considerable implications for pre-accession screening strategies in a time when the cost of performance failure is unacceptably high. Specifically, a cheap and easily administered assessment could forestall attrition and washback where screening was previously too burdensome.

A natural follow up question might seek to explain the relationship between movement patterns and performance outcomes on a mechanistic level. Different interpretations of the present findings could be taken to support vastly different approaches to training. Proponents of movement screening consider movement behaviors themselves to be a kind of stand-alone functional criterion.[10,11] This understanding has recently inspired efforts to identify intervention protocols capable of improving screening scores.[13,15] An alternative interpretation we propose would suggest that the utility of clinical screens in this context is that they allow us to see evidence of performance-relevant attributes using a convenient, field-expedient test. Understanding which attributes mediate the relationship between movement and performance will be therefore be the more appropriate focus of training and is the subject of ongoing research in our laboratory.

In conclusion, the results of this study corroborate the previously observed increase in validity of clinically rated movement as a predictor of tactical athlete performance outcomes. The clinical implication is that the elusive relationship between movement behavior and physical performance exists can be observed provided that appropriate testing conditions are in place. Future research should focus on refining testing methods to increase feasibility and information gained, as well as identifying modifiable factors that best explain this relationship.

CHAPTER VII

DISCUSSION

The research presented in this dissertation was designed to address three related questions motivated primarily by the challenges associated with identifying candidates who are physically prepared for the rigors of tactical occupations. Specifically for the US military, recruiting and training an adequate number of service members is costly. As America's youth become progressively less fit and defense spending declines,[7,153] the practice of using basic training academies as a prolonged screening process is becoming increasingly untenable. The confluence of these factors has revitalized the discussion concerning methods which might be used to identify individuals likely to attrite for reasons related to performance or injury prior to accession. A closely related discussion concerns how best to train for high performance in the face of excessively rates of injury.

A number of recent studies have explored the use of clinical assessment of movement quality, and the functional training paradigm in general, to address these problems.[3,4,165] One focus of this dissertation was to reconcile the theoretical link between movement quality and physical performance with an empirical lack of association between the two. As part of a confirmatory analysis related to Objective 3, it was demonstrated that a simple experimental treatment—namely, adding a moderate external load—improved the ability of clinical movement scores to predict criterion performance measures. This replication of our previous work in an independent sample suggests that cost-effective, field-expedient screening tools can be used as correlates of performance and moves us a step closer to establishing acceptable methods for doing so. It is worth noting, however, that while predictive validity was improved in general, the variables retained for a given outcome were not necessarily the same as those retained in the original study. To give an example, the previous data set indicated that higher Deep Squat and Hurdle Step scores were predictive of faster completion of the Mobility for Battle assessment. In contrast, the present findings show that faster completion times for the same test are predicted by higher Inline Lunge scores without contribution from Deep Squat or Hurdle Step ability. While the trend of increased predictive validity is encouraging, the more current dataset is likely a better representation of the true relationships in the population as it contains a greater number of observations and is balanced with respect to sex.

The remainder of the dissertation was intended to test a mechanistic theory regarding this association and to provide evidence in support of an alternative understanding of the utility of clinical movement screens. The prevailing theory underlying commercially available movement screens suggests that quality of movement can be defined in relation to a set of optimal, "fundamental" movement behaviors.[9,71] This has motivated a line of intervention research in which success is gauged primarily as a function of how closely individuals can approximate these behaviors, represented by scores achieved on a given clinical index of movement quality.[14,15,80] The hypotheses tested in Objective 1 examined the role of a set of performance-relevant attributes in mediating the decreases in movement quality previously observed in our laboratory. The intent of this objective was to lay the groundwork for interpreting clinical movement behaviors not in relation to an objective norm, but rather as indicators of underlying deficits which may constrain movement.

The motivation for Objective 2 was to provide evidence of the role of movement variability in tasks which are commonly used to evaluate movement clinically. This objective tested the mediating role of the same attributes observed in Objective 1, this time as mitigators of load-related loss of complexity. Complexity, patterned variability, and constrained optimization are concepts consistent with the dynamical systems interpretation of motor control.[16] While it is difficult to disprove the existence of optimal movement techniques, demonstrating the impact of various constraints during standardized movement tests offers support for the notion that a single normative reference may not exist. Accordingly, training to replicate such a reference pattern may not be universally appropriate.

The mediators tested in Objectives 1 and 2 were strength, balance, and range of motion. Strength was quantified using the YMCA Bench Press test and countermovement jump peak power (CMJPP), which respectively indicate upper and lower body strength. Range of motion was also quantified using multiple measures. The Weight-Bearing Lunge Test (WBLT) quantifies dorsiflexion range of motion; the Apley Scratch test quantifies scapulothoracic range of motion; and the sit-and-reach test (S&R) quantifies range of motion in hip and trunk flexion. Finally, balance was quantified using a single force plate metric—mean resultant center of pressure velocity (CP_V). While we occasionally use the term "constraints" in reference to these three factors, our hypotheses assume that movement would be constrained by deficits in these attributes. Our results show a relationship between each factor and clinically rated movement quality. With respect to complexity of movement, range of motion was the only factor demonstrated to be influential. Aside from dorsiflexion range of motion during the Shoulder Mobility test, no evidence was found to support the hypothesized mediating effects of strength, balance, or range of motion.

Despite failing to show mediating effects which would explain the improved validity of weighted movement scores predicting performance outcomes, our data do demonstrate a general influence of strength, balance, and range of motion on both movement quality and movement complexity. With the exception of one movement quality test which was analyzed (the FMS Rotary Stability test), each of the regression models pertaining to movement quality outcomes featured significant effects related to strength, balance, and/or range of motion after controlling for the influence of potential confounders. A number of these models featured additional effects which fell just short of our a priori significance threshold of 0.05. For example, in the prediction equation for the FMS Deep Squat task, the p-value for Apley Scratch test scores was 0.08. Similarly, in the equation for the FMS Inline Lunge, p-values of 0.06, 0.07, and 0.08 were observed for the YMCA Bench Press (specific to the weight vest condition), CP_V, and CMJPP, respectively. The same phenomenon was observed in our Objective 2 models. The only outcome featuring significant non-nuisance effects in this case was the complexity index for the cyclic Deep Squat task. In the same model, however, the p-value for CP_V was 0.09. A low, non-significant p-value for CP_V was also observed in both unilateral cyclic tasks when testing the non-dominant side with $p = 0.08$ for the Inline Lunge and $p = 0.07$ for the Hurdle Step. For the dominant-side cyclic Hurdle Step, the WBLT p-value was 0.06.

It is worth noting here the suggestion that p-values have been overemphasized in clinical research.[166] Clinically meaningful effects can easily go undetected when relying on significance criteria alone. The patterns of near-significance may reflect nonzero effects with error components which were slightly too high for positive hypothesis tests. In each of the near-significant cases just described, a theoretical explanation for a true nonzero effect is readily available. Adequate dorsiflexion range of motion is required in the Hurdle Step to enable the subject to clear the obstacle without frontal or transverse plane rotation throughout the lower body. The Deep Squat task requires upper body range of motion to maintain the position of the dowel, and balance to load equally into each hip while resisting the tendency to fall forward. Balance also facilitates execution of the Hurdle Step and Inline Lunge owing to their respective single-leg and inline stance positions. Finally, lower body strength is required to rise from the bottom of the Inline Lunge while maintaining a rigid torso and upper body strength provides further assistance in transferring forces to the lower extremity in the weight vest condition as has been observed, for example, in ruck marching.[149]

It was further demonstrated in our analyses that both movement quality and movement complexity decrease with the addition of a standardized external load. Out of 12 paired comparisons, the only outcome for which a difference was not observed was the Hurdle Step ($V = 36$, $p = 0.40$). These observations, together with the finding of a general association between our dependent measures and variables related to strength, balance, and range of motion, may be interpreted to suggest that movement behaviors do not arise as a function of ingrained motor programs which are limited to classification as intact or broken. Rather, subtle and continuous differences may be observed between individuals as a function of underlying attributes which limit the range of motor solutions at one's disposal. For example, two otherwise comparable individuals with different ranges of dorsiflexion motion will also differ in their predicted Deep Squat FMS scores. In this case, the subject with less range of motion would likely score lower than the other. These same two individuals are also likely to exhibit varying degrees of movement complexity during the cyclic Deep Squat, again with lower range of motion predicting a lower complexity index. Instead of explaining these observations in terms of motor programs, a simpler conclusion to draw would be that restrictions in dorsiflexion range of motion confine an individual to a more limited repertoire of Deep Squat strategies. Using this limited repertoire, that individual will be less effective at adapting to an ever-changing set of constraints, yielding a behavioral signal which is relatively low in complexity. This individual may also fail to reach the criteria for a high clinical score, but it is not necessary to consider this a reflection of the status of a fundamental motor program which governs the Deep Squat.

Within-subject differences in movement behavior can similarly be observed as a function of constraints, in this case more likely those related to the task or environment. Our investigation introduced an experimental constraint in the form of a standardized external load. The observed effect was a decrease in both movement quality and movement complexity, presumably proportional to the degree of constraint induced by the load in a given individual. As simple as this treatment was, it presents a problem for those who would interpret movement quality ratings in relation to an optimal movement pattern technique. Specifically, if fundamental motor programs are the primary constituent of movement behaviors, what accounts for the decreases in movement quality resulting from a standardized external load? As soon as we make mention of strength, size, neuromuscular control, or any factor other than a motor program which influences movement behavior, we open our discussion to include the topic of constraints, a topic which is currently not possible to address comprehensively. It is quite possible that the current focus on fundamental movement patterns reflects a tendency toward false oversimplifications of a phenomenon which does not so easily lend itself to generalization. Therefore, despite failing to demonstrate that the specific factors we measured mediate the effect of load on movement quality and complexity, there is evidence to suggest that the movement behaviors we observed reflect a process of constrained optimization. This warrants further work designed to identify and quantify the influence of various constraints, as well as research to create more comprehensive topological maps documenting how these behaviors change in relation to varying task and environmental constraints.

Certain limitations regarding the current work need to be considered. Our data were collected in a sample of young, healthy men and women and therefore should not be generalized to populations with differing clinical or demographic characteristics. Further, our primary experimental tasks were selected from a limited range of tests which have been applied clinically and should not be interpreted as a comprehensive battery of functional movement assessments. Next, the models tested in each objective were not sufficiently powered for conventional analyses. While efforts were made to address this by using appropriate penalization algorithms, it is still possible that nonzero effects went undetected in the final models. Perhaps most importantly, our data are limited to repeated cross-sectional observations and therefore cannot be assumed to predict outcomes in a longitudinal design. Prospective studies will be required before any training recommendations or firm conclusions regarding future performance in tactical occupations can be made.

Notwithstanding these limitations, we can draw the following conclusions related to each of our objectives: 1A) Standardized external loading results in decreases in clinical movement quality scores, 1B) Strength, balance, and range of motion are associated with movement quality, 2A) Standardized external loading results in decreased complexity in commonly used assessments of functional movement, 2B) Range of motion, and potentially balance, are associated with movement complexity in these tasks; 3) For outcomes related to tactical athletic performance, standardized external loading increases the predictive validity of clinically rated movement quality.

REFERENCES

[1] Motte SJ, Lisman P, Beutler A, Deuster P, O'Connor F. Functional movement screen scores in pre-deploying marines are low. *Med Sci Sports Exerc.* 2014;46:116–117.

[2] Goss D, Christopher G, Faulk R, Moore J. Functional training program bridges rehabilitation and return to duty. *J Spec Oper Med.* 2009;9:29–48.

[3] Lisman P, O'Connor F, Deuster P, Knapik J. Functional movement screen and aerobic fitness predict injuries in military training. *Med Sci Sports Exerc.* 2013;45:636–43.

[4] O'Connor F, Deuster P, Davis J, Pappas C, Knapik J. Functional movement screening: predicting injuries in officer candidates. *Med Sci Sports Exerc.* 2011;43:2224–30.

[5] Peterson LQ, Evans EM, Kennedy-Ambruster C. Exercise training improves functional movement screen (fms) scores in 40+ year-old active duty military personnel. *Med Sci Sports Exerc.* 2014;46:118–118.

[6] Deuster PA, O'Connor FG, Henry KA, et al. Human performance optimization: an evolving charge to the department of defense. *Mil Med.* 2007;172:1133–7.

[7] Nindl BC. Strategies for enhancing military physical readiness in the 21st century. *US Army Med Dep J.* 2013:5–23.

[8] Cook G, Burton L, Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function: Part 1. *N Am J Sports Phys Ther.* 2006;1:62.

[9] Cook G, Burton L, Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function: Part 2. *N Am J Sports Phys Ther.* 2006;1:132.

[10] Kritz M. *Development, reliability and effectiveness of the Movement Competency Screen (MCS).* PhD thesisAuckland University of Technology 2012.

[11] Cook G. *Movement: Functional movement systems: Screening, assessment, and corrective strategies.* Aptos, California: On Target Publications 2010.

[12] Fischer K, David J, Paul K, Leonard R. The effect of functional movement screen-based circuit training on balance and postural stability. in *International Journal of Exercise Science: Conference Proceedings*;9:19 2013.

[13] Frost D, Beach TA, Callaghan JP, McGill SM. Using the functional movement screen to evaluate the effectiveness of training. *J Strength Cond Res.* 2012;26:1620–30.

[14] Kiesel K, Plisky PJ, Butler R. Functional movement test scores improve following a standardized off-season intervention program in professional football players. *Scand J Med Sci Sports.* 2011;21:287–92.

[15] Bodden JG, Needham RA, Chockalingam N. The effect of an intervention program on functional movement screen test scores in mixed martial arts athletes. *J Strength Cond Res.* 2013.

[16] Glazier PS, Davids K. Constraints on the complete optimization of human motion. *Sports Med.* 2009;39:15–28.

[17] Newell KM. Constraints on the development of coordination. *Motor development in children: Aspects of coordination and control.* 1986;34:341–360.

[18] Goldberger A, Peng C, Lipsitz L. What is physiologic complexity and how does it change with aging and disease? *Neurobiol Aging.* 2002;23.

[19] Lipsitz LA, Goldberger AL. Loss of 'complexity' and aging. potential applications of fractals and chaos theory to senescence. *JAMA.* 1992;267:1806–9.

[20] Hamill J, Palmer C, Van Emmerik R. Coordinative variability and overuse injury. *Sports Med Arthrosc Rehabil Ther Technol.* 2012;4:1–9.

[21] Glazier P. Game, set and match? substantive issues and future directions in performance analysis. *Sports Med.* 2010;40:625–634.

[22] Butler R, Plisky P, Southers C, Scoma C, Kiesel K. Biomechanical analysis of the different classifications of the functional movement screen deep squat test. *Sports Biomech.* 2010;9:270–9.

[23] Kiesel K, Plisky P, Voight M. Can serious injury in professional football be predicted by a preseason functional movement screen? *N Am J Sports Phys Ther.* 2007;2:147.

[24] McGill S, Frost D, Andersen J, Crosby I, Gardiner D. Movement quality and links to measures of fitness in firefighters. *Work.* 2013;45:357–66.

[25] McGill S, Andersen J, Horne A. Predicting performance and injury resilience from movement quality and fitness scores in a basketball team over 2 years. *J Strength Cond Res.* 2012;26:1731–9.

[26] Okada T, Huxel KC, Nesser TW. Relationship between core stability, functional movement, and performance. *J Strength Cond Res.* 2011;25:252–61.

[27] Parchmann C, McBride J. Relationship between functional movement screen and athletic performance. *J Strength Cond Res.* 2011;25:3378–84.

[28] Glass S, Ross S. Single-subject balance training effects on functional movement and postural control. *Unpublished pilot data.* 2013.

[29] Glass S, Ross S. Modified functional movement screening as a predictor of tactical performance in recreationally active adults. *Int J Sports Phys Ther.* 2015;10:612–621.

[30] Hrysomallis C. Balance ability and athletic performance. *Sports Med.* 2011;41:221–32.

[31] Peterson MD, Alvar BA, Rhea MR. The contribution of maximal force production to explosive movement among young collegiate athletes. *J Strength Cond Res.* 2006;20:867–73.

[32] Attwells RL, Birrell SA, Hooper RH, Mansfield NJ. Influence of carrying heavy loads on soldiers' posture, movements and gait. *Ergonomics.* 2006;49:1527–1537.

[33] Birrell SA, Haslam RA. The effect of military load carriage on 3-D lower limb kinematics and spatiotemporal parameters. *Ergonomics.* 2009;52:1298–1304.

[34] DeMaio M, Onate J, Swain D, Morrison S, Ringleb S, Naiak D. Physical performance decrements in military personnel wearing personal protective equipment (PPE). in *RTO-MP-HFM-181 Human Performance Enhancement for NATO Military Operations* 2009.

[35] Heller MF, Challis JH, Sharkey NA. Changes in postural sway as a consequence of wearing a military backpack. *Gait Posture.* 2009;30:115–117.

[36] Polcyn AF, Bensel CK, Harman EA, Obusek JP, Pandorf C. Effects of weight carried by soldiers: Combined analysis of four studies on maximal performance, physiology, and biomechanics. tech. rep.DTIC Document 2002.

[37] Carlock JM, Smith SL, Hartman MJ, et al. The relationship between vertical jump power estimates and weightlifting ability: a field-test approach. *J Strength Cond Res.* 2004;18:534–9.

[38] Kim PS, Mayhew JL, Peterson DF. A modified YMCA bench press test as a predictor of 1 repetition maximum bench press strength. *J Strength Cond Res.* 2002;16:440–445.

[39] Hoch MC, Staton GS, McKeon PO. Dorsiflexion range of motion significantly influences dynamic balance. *J Sci Med Sport.* 2011;14:90–92.

[40] Ayala F, Baranda P, De Ste Croix M, Santonja F. Reproducibility and criterion-related validity of the sit and reach test and toe touch test for estimating hamstring flexibility in recreationally active young adults. *Phys Ther Sport.* 2012;13:219–226.

[41] Ross M, Paterno G. Reliability of the apley's range of motion test and dominant and nondominant upper extremity differences. *Eur J Phys Rehabil Med.* 1999;8:166–169.

[42] Harman EA, Gutekunst DJ, Frykman PN, et al. Prediction of simulated battlefield physical performance from field-expedient tests. *Mil Med.* 2008;173:36–41.

[43] Crowder T, Ferrara AL, Levinbook MD. Creation of a criterion-referenced military optimal performance challenge. *Mil Med.* 2013;178:1085–1101.

[44] Gebicke M. Military attrition: Dod needs to better understand reasons for separation and improve recruiting systems. tech. rep.United States General Accounting Office 1998.

[45] Knapik JJ, Jones BH, Sharp MA, Darakjy S, Jones S. The case for pre-enlistment physical fitness testing: Research and recommendations. tech. rep.DTIC Document 2004.

[46] Niebuhr D, Scott C, Powers T, et al. Assessment of recruit motivation and strength study: preaccession physical fitness assessment predicts early attrition. *Mil Med.* 2008;173:555–562.

[47] Niebuhr DW, Page WF, Cowan DN, Urban N, Gubata ME, Richard P. Cost-effectiveness analysis of the us army assessment of recruit motivation and strength (ARMS) program. *Mil Med.* 2013;178:1102–1110.

[48] Niebuhr DW, Scott CT, Li Y, Bedno SA, Han W, Powers TE. Preaccession fitness and body composition as predictors of attrition in us army recruits. *Mil Med.* 2009;174:695–701.

[49] Loughran DS, Orvis BR. The effect of the assessment of recruit motivation and strength (ARMS) program on army accessions and attrition. tech. rep.DTIC Document 2011.

[50] Kelly T, Masi R, Walker B, Knapp S, Leuschner K. An assessment of the army's tactical human optimization, rapid rehabilitation and reconditioning program. tech. rep.DTIC Document 2013.

[51] Sell T, Abt J, Crawford K, et al. Warrior model for human performance and injury prevention: eagle tactical athlete program (ETAP) part i. *J Spec Oper Med.* 2010;10:2–21.

[52] Sell T, Abt J, Crawford K, et al. Warrior model for human performance and injury prevention: eagle tactical athlete program (ETAP) part ii. *J Spec Oper Med.* 2010;10:22–33.

[53] Stephenson M, Dawes J, Snyder S. Training the tactical athlete: special forces assessment and selection. 2007.

[54] Instruction C. Instruction 3405.01: Chairman's total force fitness framework. 2011.

[55] Lopez C. Recruiting force remains unchanged, despite shrinking goals. 2014.

[56] Molloy JM, Feltwell DN, Scott SJ, Niebuhr DW. Physical training injuries and interventions for military recruits. *Mil Med.* 2012;177:553–558.

[57] Jones BH, Cowan DN, Knapik JJ. Exercise, training and injuries. *Sports Med.* 1994;18:202–214.

[58] Hauret KG, Taylor BJ, Clemmons NS, Block SR, Jones BH. Frequency and causes of nonbattle injuries air evacuated from operations iraqi freedom and enduring freedom, U.S. army, 2001-2006. *Am J Prev Med.* 2010;38:S94–107.

[59] O'Connor FG, Deuster PA. Musculoskeletal health and injury prevention. 2008.

[60] Jonas WB, O'Connor FG, Deuster P, Peck J, Shake C, Frost SS. Why total force fitness? *Mil Med.* 2010;175:6–13.

[61] Fleishman EA. *The structure and measurement of physical fitness.* Prentice-Hall 1964.

[62] Hogan J. Structure of physical performance in occupational tasks. *J Appl Psychol.* 1991;76:495.

[63] Myers DC, Gebhardt DL, Crump CE, Fleishman EA. The dimensions of human physical performance: factor analysis of strength, stamina, flexibility, and body composition measures. *Hum Perform.* 1993;6:309–344.

[64] Knapik JJ, Rieger W, Palkoska F, Van Camp S, Darakjy S. United States army physical readiness training: rationale and evaluation of the physical training doctrine. *J Strength Cond Res.* 2009;23:1353–62.

[65] U.S. A. Army physical readiness training field manual FM 7-22. 2012.

[66] U.S. N. Opnav instruction 6110.1H. 2005.

[67] Force UA. Air force instruction 36-2905. tech. rep.Department of the Air Force 2013.

[68] Roy TC, Springer BA, McNulty V, Butler NL. Physical fitness. *Mil Med.* 2010;175:14–20.

[69] Hewett TE, Myer GD, Ford KR. Anterior cruciate ligament injuries in female athletes: Part 1, mechanisms and risk factors. *Am J Sports Med.* 2006;34:299–311.

[70] Myer GD, Kushner AM, Brent JL, et al. The back squat: A proposed assessment of functional deficits and technical factors that limit performance. *Strength Cond J.* 2014;36:4–27.

[71] Clark M, Russell A. Optimum performance training for the performance enhancement specialist. 2002.

[72] Hirth CJ. Clinical movement analysis to identify muscle imbalances and guide exercise. *Athl Ther Today.* 2007;12.

[73] Kritz M, Hume P, Cronin J. The bodyweight squat: A movement screen for the squat pattern. *Strength Cond J.* 2009;31:76–85.

[74] Bernstein NA. *The co-ordination and regulation of movements.* Oxford, New York: Pergamon Press 1967.

[75] Cook G. *Athletic body in balance.* Champaign, IL: Human Kinetics 2003.

[76] Frohm A, Heijne A, Kowalski J, Svensson P, Myklebust G. A nine-test screening battery for athletes: a reliability study. *Scand J Med Sci Sports.* 2012;22:306–15.

[77] Davids K, Glazier P, AraÃºjo D, Bartlett R. Movement systems as dynamical systems : The functional role of variability and its implications for sports medicine. *Sports Med.* 2003;33:245–260.

[78] West BJ. *Where medicine went wrong rediscovering the path to complexity.* Hackensack, NJ: World Scientific 2006.

[79] Brisson TA, Alain C. Should common optimal movement patterns be identified as the criterion to be achieved? *J Mot Behav.* 1996;28:211–223.

[80] Wright MD, Portas MD, Evans VJ, Weston M. The effectiveness of 4 weeks of fundamental movement training on functional movement screen and physiological performance in physically active children. *J Strength Cond Res.* 2015;29:254–61.

[81] Cowen VS. Functional fitness improvements after a worksite-based yoga initiative. *J Bodyw Mov Ther.* 2010;14:50–4.

[82] Crouse V. *The functional movement screen and its relationship to measures of athletic-related performance, body composition, and injury rates.* PhD thesisThe Pennsylvania State University 2014.

[83] Duncan M, Stanley M. Functional movement is negatively associated with weight status and positively associated with physical activity in british primary school children. *J Obes.* 2012;2012:697563.

[84] Perry F, Koehle M. Normative data for the functional movement screen in middle-aged adults. *J Strength Cond Res.* 2013;27:458–62.

[85] Hoover D, Killian C, Bourcier B, Lewis S, Thomas J, Willis R. Predictive validity of the functional movement screentm in a population of recreational runners training for a half marathon. *Med Sci Sports Exerc.* 2008;40:S219.

[86] Chorba RS, Chorba DJ, Bouillon LE, Overmyer CA, Landis JA. Use of a functional movement screening tool to determine injury risk in female collegiate athletes. *N Am J Sports Phys Ther.* 2010;5:47–54.

[87] Letafatkar A, Hadadnezhad M, Shojaedin S, Mohamadi E. Relationship between functional movement screening score and history of injury. *Int J Sports Phys Ther.* 2014;9:21.

[88] Sorenson EA. *Functional movement screen as a predictor of injury in high school basketball athletes.* PhD thesisUniversity of Oregon 2009.

[89] Lockie RG, Schultz AB, Jordan CA, Callaghan SJ, Jeffriess MD, Luczo TM. Can selected functional movement screen assessments be used to identify movement deficiencies that could affect multidirectional speed and jump performance? *J Strength Cond Res.* 2015;29:195–205.

[90] Frost D, Andersen J, Lam T, Finlay T, Darby K, McGill S. The relationship between general measures of fitness, passive range of motion and whole-body movement quality. *Ergonomics.* 2013;56:637–649.

[91] Hartigan EH, Lawrence M, Bisson BM, Torgerson E, Knight RC. Relationship of the functional movement screen in-line lunge to power, speed, and balance measures. *Sports Health.* 2014;6:197–202.

[92] Sprague PA, Mokha M, Gatens DR. Changes in functional movement screen scores over a season in collegiate soccer and volleyball athletes. *J Strength Cond Res.* 2014:1.

[93] McGill S, Frost D, Lam T, Finlay T, Darby K, Andersen J. Fitness and movement quality of emergency task force police officers: An age-grouped database with comparison to populations of emergency services personnel, athletes and the general public. *Int J Ind Ergon.* 2013;43:146–153.

[94] Chapman RF, Laymon AS, Arnold T. Functional movement scores and longitudinal performance outcomes in elite track and field athletes. *Int J Sports Physiol Perform.* 2014;9:203–11.

[95] Bock C, Orr R, Stierli M, Hinton B. The functional movement screen as a predictor of police occupational task performance. *J Sci Med Sport.* 2014;18:e79.

[96] Kazman JB, Galecki JM, Lisman P, Deuster PA, O'Connor FG. Factor structure of the functional movement screen in marine officer candidates. *J Strength Cond Res.* 2014;28:672–8.

[97] Li Y, Wang X, Chen X, Dai B. Exploratory factor analysis of the functional movement screen in elite athletes. *J Sports Sci.* 2014:1–7.

[98] Lubans DR, Smith JJ, Harries SK, Barnett LM, Faigenbaum AD. Development, test-retest reliability, and construct validity of the resistance training skills battery. *J Strength Cond Res.* 2014;28:1373–80.

[99] Thelen M. A novel return to duty screening tool for military clinicians. 2014.

[100] Legge J, Burgess-Limerick R. Reliability of the jobfit system pre-employment functional assessment tool. *Work.* 2007;28:299–312.

[101] Tarara DT, Hegedus EJ, Taylor JB. Real-time test-retest and interrater reliability of select physical performance measures in physically active college-aged students. *Int J Sports Phys Ther.* 2014;9:874–887.

[102] McKeown I, Taylor-McKeown K, Woods C, Ball N. Athletic ability assessment: a movement assessment protocol for athletes. *Int J Sports Phys Ther.* 2014;9:862.

[103] Reid DA, Vanweerd RJ, Larmer PJ, Kingstone R. The inter and intra rater reliability of the netball movement screening tool. *J Sci Med Sport.* 2014.

[104] Glass S, Ross S. Modified functional movement screening as a predictor of tactical performance potential in recreationally active adults. *Int J Sports Phys Ther.* 2015;10:612–21.

[105] Mazur JE. Optimization: a result or a mechanism? *Science.* 1983;221:977.

[106] Lipsitz LA. Dynamics of stability: the physiologic basis of functional health and frailty. *J Gerontol.* 2002;57:115–25.

[107] Slifkin AB, Newell KM. Is variability in human performance a reflection of system noise? *Curr Dir Psychol Sci.* 1998;7:170–177.

[108] Stergiou N, Decker LM. Human movement variability, nonlinear dynamics, and pathology: is there a connection? *Hum Mov Sci.* 2011;30:869–88.

[109] Vaillancourt D, Newell K. Changing complexity in human behavior and physiology through aging and disease. *Neurobiol Aging.* 2002;23:1–11.

[110] Winter DA. Human balance and posture control during standing and walking. *Gait Posture.* 1995;3:193–214.

[111] Peng CK, Costa M, Goldberger AL. Adaptive data analysis of complex fluctuations in physiologic time series. *Adv Adapt Data Anal.* 2009;1:61–70.

[112] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol.* 2000;278.

[113] Ramdani S, Seigle B, Lagarde J, Bouchara F, Bernard PL. On the use of sample entropy to analyze human postural sway data. *Med Eng Phys.* 2009;31:1023–1031.

[114] Glass S, Ross S, Arnold B, Rhea C. Center of pressure regularity with and without stochastic resonance stimulation in stable and unstable ankles. *Athletic Training & Sports Health Care.* 2014;6:170–178.

[115] Cavanaugh J, Guskiewicz K, Giuliani C, Marshall S, Mercer V, Stergiou N. Recovery of postural control after cerebral concussion: new insights using approximate entropy. *J Athl Train.* 2006;41.

[116] Ahmed MU, Rehman N, Looney D, Rutkowski TM, Mandic DP. Dynamical complexity of human responses: a multivariate data-adaptive framework. *Bulletin of the Polish Academy of Sciences: Technical Sciences.* 2012;60:433–445.

[117] Newell KM, Van Emmerik REA, McDonald PV. Biomechanical constraints and action theory. *Hum Mov Sci.* 1989;8:403–409.

[118] Jaric S, Mirkov D, Markovic G. Normalizing physical performance tests for body size: Aproposal for standardization. *J Strength Cond Res.* 2005;19:467–474.

[119] McBride JM, Blow D, Kirby TJ, Haines TL, Dayne AM, Triplett NT. Relationship between maximal squat strength and five, ten, and forty yard sprint times. *J Strength Cond Res.* 2009;23:1633–1636.

[120] Wisloff U, Castagna C, Helgerud J, Jones R, Hoff J. Strong correlation of maximal squat strength with sprint performance and vertical jump height in elite soccer players. *Br J Sports Med.* 2004;38:285–288.

[121] Saunders PU, Pyne DB, Telford RD, Hawley JA. Factors affecting running economy in trained distance runners. *Sports Med.* 2004;34:465–485.

[122] Davlin C. Dynamic balance in high level athletes. *Percept Mot Skills.* 2004;98:1171–6.

[123] Paillard T, Noe F, Riviere T, Marion V, Montoya R, Dupui P. Postural performance and strategy in the unipedal stance of soccer players at different levels of competition. *J Athl Train.* 2006;41:172.

[124] Kean CO, Behm DG, Young WB. Fixed foot balance training increases rectus femoris activation during landing and jump height in recreationally active women. *J Sports Sci Med.* 2006;5:138.

[125] Yaggie J, Campbell B. Effects of balance training on selected skills. *J Strength Cond Res.* 2006;20:422–8.

[126] Mellin G. Correlations of hip mobility with degree of back pain and lumbar spinal mobility in chronic low-back pain patients. *Spine.* 1988;13:668–670.

[127] Davis DS, Barnette BJ, Kiger JT, Mirasola JJ, Young SM. Physical characteristics that predict functional performance in division I college football players. *J Strength Cond Res.* 2004;18:115–20.

[128] Kibler WB, Press J, Sciascia A. The role of core stability in athletic function. *Sports Med.* 2006;36:189–198.

[129] Schiffman JM, Bensel CK, Hasselquist L, Gregorczyk KN, Piscitelle L. Effects of carried weight on random motion and traditional measures of postural sway. *Appl Ergon.* 2006;37:607–614.

[130] Park H, Nolli G, Branson D, Peksoz S, Petrova A, Goad C. Impact of wearing body armor on lower body mobility. *Clothing and Textiles Research Journal.* 2011;29:232–247.

[131] Walsh JC, Quinlan JF, Stapleton R, FitzPatrick DP, McCormack D. Three-dimensional motion analysis of the lumbar spine during "free squat" weight lift training. *Am J Sports Med.* 2007;35:927–932.

[132] Brown TN, O'Donovan M, Hasselquist L, Corner B, Schiffman JM. Soldier-relevant loads impact lower limb biomechanics during anticipated and unanticipated single-leg cutting movements. *J Biomech.* 2014;47:3494–3501.

[133] Ritti-Dias RM, Avelar A, Salvador EP, Cyrino ES. Influence of previous experience on resistance training on reliability of one-repetition maximum test. *J Strength Cond Res.* 2011;25:1418–1422.

[134] Nuzzo JL, McBride JM, Cormie P, McCaulley GO. Relationship between countermovement jump performance and multijoint isometric and dynamic tests of strength. *J Strength Cond Res.* 2008;22:699–707.

[135] Dugan EL, Doyle TL, Humphries B, Hasson CJ, Newton RU. Determining the optimal load for jump squats: a review of methods and calculations. *J Strength Cond Res.* 2004;18:668–74.

[136] Knapik J, Reynolds K, Santee W, Friedl K. Load carriage in military operations a review of historical, physiological, biomechanical, and medical aspects. tech. rep.Office of the Surgeon General and the Borden Institute 2010.

[137] Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci USA.* 1991;88:2297–2301.

[138] Costa M, Peng CK, Goldberger AL, Hausdorff JM. Multiscale entropy analysis of human gait dynamics. *Physica A.* 2003;330:53–60.

[139] Gertheiss J, Oehrlein F. Testing linearity and relevance of ordinal predictors. *Electron J Stat.* 2011;5:1935–1959.

[140] Gertheiss J, Tutz G. Penalized regression with ordinal predictors. *Int Stat Rev.* 2009;77:345–365.

[141] Tutz G, Gertheiss J. Rating scales as predictors: the old question of scale level and some answers. *Psychometrika.* 2013:1–20.

[142] Hess W, Persson M, Rubenbauer S, Gertheiss J. Using lasso-type penalties to model time-varying covariate effects in panel data regressions. tech. rep.Research Institute of Industrial Economics Working Paper No. 961 2013.

[143] Gertheiss J. ordpens: Selection and/or smoothing of ordinal predictors. r package version 0.2-1. 2010.

[144] Breheny P, Huang J. Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. *Stat Comput.* 2013:1–15.

[145] Canty A, Ripley B. boot: Bootstrap r (s-plus) functions. r package version 1.3-11. 2014.

[146] Knipscher BW. Thor3: Humans are more important than hardware. tech. rep.DTIC Document 2010.

[147] Teyhen D, Shaffer S, Lorenson C, et al. The functional movement screen: a reliability study. *J Orthop Sports Phys Ther.* 2012;42:530–40.

[148] Kim J. *Lower Body Kinematic Comparisons between Front and Back Squats in Response to Loads*. PhD thesisBridgewater State University 2014.

[149] Knapik JJ, Harman EA, Steelman RA, Graham BS. A systematic review of the effects of physical training on load carriage performance. *J Strength Cond Res.* 2012;26:585–97.

[150] Linens SW, Ross SE, Arnold BL, Gayle R, Pidcoe P. Postural-stability tests that identify individuals with chronic ankle instability. *J Athl Train.* 2014;49:15–23.

[151] Clifton D, Harrison B, Hertel J, Hart J. Relationship between functional assessments and exercise-related changes during static balance. *J Strength Cond Res.* 2013;27:966–72.

[152] Riccio GE, Stoffregen TA. Affordances as constraints on the control of stance. *Hum Mov Sci.* 1988;7:265–300.

[153] Christeson W, Taggart A, Messner-Zidelli S. Too fat to fight. tech. rep.Mission: Readiness 2010.

[154] McMillian D. Ranger athlete warrior systematic approach to conditioning. *Infantry.* 2007;96:5–8.

[155] Latash ML, Anson JG. What are "normal movements" in atypical populations? *Behav Brain Sci.* 1996;19:55.

[156] Bartlett R, Wheat J, Robins M. Is movement variability important for sports biomechanists? *Sports Biomech.* 2007;6:224–243.

[157] Rhea CK, Kiefer AW, Wright WG, Raisbeck LD, Haran FJ. Interpretation of postural control may change due to data processing techniques. *Gait Posture.* 2015;41:731–5.

[158] Rhea C, Silver T, Hong S, et al. Noise and complexity in human postural control: interpreting the different estimations of entropy. *PLoS One.* 2011;6.

[159] Haddad JM, Van Emmerik RE, Wheat JS, Hamill J. Developmental changes in the dynamical structure of postural sway during a precision fitting task. *Exp Brain Res.* 2008;190:431–41.

[160] McGinnis PM, Newell KM. Topological dynamics: A framework for describing movement and its constraints. *Hum Mov Sci.* 1982;1:289–305.

[161] Kapp L. *Recruiting and Retention: An Overview of FY2010 and FY2011 Results for Active and Reserve Component Enlisted Personnel.* Congressional Research Service 2012.

[162] Lockie R, Jalilvand F, Jordan C, Callaghan S, Jeffriess M. Research-grade scoring for the functional movement screen and relationships with athletic performance tests in team sport athletes. *J Athl Enhanc.* 2015;8:14–17.

[163] Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc.* 1982;14:377–81.

[164] Breheny P, Huang J. Penalized methods for bi-level variable selection. *Stat Interface.* 2009;2:369–380.

[165] Teyhen DS, Shaffer SW, Umlauf JA, et al. Automation to improve efficiency of field expedient injury prediction screening. *J Strength Cond Res.* 2012;26:S61-S72.

[166] Page P. Beyond statistical significance: Clinical interpretation of rehabilitation research literature. *Int J Sports Phys Ther.* 2014;9:726–736.

APPENDIX A
PHYSICAL ACTIVITY AND INJURY HISTORY

Subject # ______

Date ____________

Subject Demographics

- Sex:
- Age:
- Height (cm):
- Mass (kg):
- Preferred kicking leg:

Activity Readiness Questions

- Has a doctor ever said that you have a heart condition and that you should only perform physical activity recommended by a doctor?
  Yes______ No______
- Are you pregnant? Yes______ No______ NA______
- Do you smoke? Yes______ No______ If yes, how often?
- Do you drink alcohol? Yes______ No______ If yes, how often?
- Do you have any General Health Problems or Illnesses? (e.g. diabetes, respiratory disease) Yes______ No______
- Do you feel pain in your chest when you perform physical activity? Yes______ No______

- In the past month, have you had chest pain when you were not performing any physical activity? Yes______ No______
- Do you have any vestibular (inner ear) or balance disorders? Yes______ No______
- Do you lose your balance because of dizziness or do you ever lose consciousness? Yes______ No______
- Do you have a bone or joint problem that could be made worse by a change in your physical activity? Yes______ No______
- Do you have any history of connective tissue disease or disorders? (e.g. Ehlers-Danlos, Marfan's Syndrome, Rheumatoid Arthritis)
  Yes______ No______
- Is your doctor currently prescribing any medication for your blood pressure or for a heart condition? Yes______ No______
- Do you know of any other reason why you should not engage in physical activity? Yes______ No______
- Please list any medications you take regularly:

Please list any previous injuries to your lower extremities. Please include a description of the injury (e.g. ligament sprain, muscle strain), severity of the injury, approximate date of the injury, and whether it was on the left or right side

	Body Part	Description	Severity	Date of Injury	L or R
Hip					
Thigh					
Knee					
Lower Leg					
Ankle					
Foot					

Please list any previous surgery to your lower extremities (Include a description of the surgery, the date of the surgery, and whether it was on the left or right side):

Body Part	Description	Date of Surgery	L or R

Please list all physical activities that you are currently engaged in. For each activity, please indicate how much time you spend each week in this activity, the intensity of the activity (i.e. competitive or recreational) and for how long you have been regularly participating in the activity:

Body Part	Body Part	Days/week	Minutes/Day	Intensity	When Began?
Hip					
Thigh					
Knee					
Lower Leg					
Ankle					
Foot					

What time of day do you generally engage in the above activities?

Please list other conditions / concerns that you feel we should be aware of:

Investigator Comments: