

Potential Mediators of Load-Related Decreases in Movement Quality in Young, Healthy Adults

Stephen M. Glass, PhD*; Randy J. Schmitz, PhD, LAT, ATC†; Christopher K. Rhea, PhD‡; Scott E. Ross, PhD, LAT, ATC, FNATA‡

*Department of Otolaryngology, The Ohio State University, Columbus; †Applied Neuromechanics Research Group, Department of Kinesiology, The University of North Carolina at Greensboro

Context: Predicting and promoting physical performance are important goals within the tactical professional community. Movement screens are frequently used in this capacity but are poor predictors of performance outcomes. It has recently been shown that prediction improved when movement quality was evaluated under load, but the mechanisms underlying this improvement remain unclear. Because balance, range of motion, and strength are mutually relevant to physical performance and movement quality, these attributes may mediate load-related decreases in movement quality and account for the resulting increase in performance prediction.

Objective: To quantify the roles of balance, range of motion, and strength in mediating load-related decreases in clinical movement-screen scores.

Design: Crossover study.

Setting: Research laboratory.

Patients or Other Participants: Twenty-five male (age = 23.96 ± 3.74 years, height = 178.82 ± 7.51 cm, mass = 79.66 ± 12.66 kg) and 25 female (age = 22.00 ± 2.02 years, height = 165.40 ± 10.24 cm, mass = 63.98 ± 11.07 kg) recreationally active adults.

Intervention(s): Participants completed a clinical movement screen under a control condition and while wearing an

18.10-kg weighted vest as well as tests of balance, range of motion, and strength.

Main Outcome Measure(s): Item score differences were assessed using Wilcoxon signed rank tests for matched pairs. Interactions between (1) balance, range of motion, and strength and (2) load condition were modeled using penalized varying-coefficients regression with item scores as the dependent measure.

Results: Except for the hurdle step, item scores were lower in the weighted-vest than in the control condition for all tests ($P < .05$). Except for rotary stability, F statistics were significant for all models (P values $< .05$, F^2 values = 0.22 – 0.77). Main effects of balance, range of motion, and strength on Functional Movement Screen scores were observed ($P < .05$); however, little evidence was found to suggest that these attributes mediated load-related decreases in Functional Movement Screen item scores.

Conclusions: Balance, range of motion, and strength affected movement quality but did not mediate the effect of the load treatment.

Key Words: weight bearing, postural balance, range of motion, muscle strength

Key Points

- Functional Movement Screen item scores decreased when testing was administered with a standardized external load.
- Balance, range of motion, and strength appeared to be important contributors to item score variability, with or without an external load.
- Balance, range of motion, and strength did not mediate the effect of load on item scores.

Predicting and promoting physical performance are important goals in the tactical professional community. Within the military specifically, broad human-performance-optimization initiatives¹ have been launched that take a multifaceted approach to addressing performance deficits. Movement screens have become increasingly integrated with these initiatives as a correlate of injury risk and performance, and their use in other tactical professional settings has followed suit.^{2,3}

The popularity of movement screening has grown dramatically in recent years as evidenced by the number of screens that have been developed. Examples include the Functional Movement Screen (FMS),^{4,5} the Return to Duty screen,⁶ and the Athletic Ability Assessment⁷ to name a few. Of these, the FMS is likely the most popular and well

researched with extensive application in tactical professional populations.² However, despite its popularity, the FMS is a poor predictor of physical performance outcomes.⁸

It has recently been shown that FMS scores decreased when the screen was administered using a standardized external load.⁹ Further, item scores from an FMS administered when wearing an external load were better predictors of tactical performance than were conventional FMS item scores.⁹ This latter finding might suggest that a load-enhanced screening can sensitize tools such as the FMS to movement deficits that are relevant to physical performance. Yet it remains unclear what mechanisms might account for this improved predictability and to what extent they can be evaluated using load-enhanced screen-

ing. Importantly, understanding these mechanisms could potentially improve our ability to identify and develop high-performing tactical professionals.

Performance is a multidimensional construct with several underlying factors. Classical models¹⁰ of human performance identify components such as strength, speed, power, agility, balance, flexibility, coordination, and endurance, of which several are also suggested to influence performance on FMS movement tests.^{4,5,11} The factors that mutually affect physical performance and clinical movement screens—particularly when such screens are modified to incorporate external load carriage—may mediate the observed changes in movement quality and the resulting increase in its association with tactical performance. We propose here that balance, range of motion, and strength are potential factors underlying these observations. Of the traits that are theorized to be assessed by clinical movement quality scales, each of these 3 has a role in promoting athleticism^{12–14} and each has been shown to interact with external loading.^{15–17}

This article presents part 1 of a 2-part work on the relationship between clinical movement screens and physical performance. Part 1 focused on clinically observable (ie, noninstrumented) aspects of movement-screen performance. Specifically, the purpose of this investigation was to examine the roles played by balance, range of motion, and strength in mediating load-related decreases in clinically rated movement quality, specifically FMS item scores. Part 2 investigated the relationship between these same factors and the dynamic behavior of the movement system during similar clinical movement tests. The knowledge derived from these works could provide a theoretical basis for the development of tactical training programs and refined movement-quality assessments with increased predictive validity for a variety of outcomes. We hypothesized that decreases in FMS item scores in the weighted-vest condition would be smaller in individuals with high levels of balance, range of motion, and strength.

METHODS

This study used a randomized crossover trial to quantify the mediating effects of balance, range of motion, and strength on within-participants differences in movement quality related to external loading. Approval was obtained from the institutional review board of the University of North Carolina at Greensboro. Data were collected in a laboratory setting by a single investigator experienced in the required measurement techniques. Twenty-five male (age = 23.96 ± 3.74 years, height = 178.82 ± 7.51 cm, mass = 79.66 ± 12.66 kg) and 25 female (age = 22.00 ± 2.02 years, height = 165.40 ± 10.24 cm, mass = 63.98 ± 11.07 kg) recreationally active adults were recruited. Participation was limited to individuals between 18 and 34 years of age in order to reflect the recruitment pool for military and tactical occupations. Participation was additionally limited by the following exclusion criteria: (1) reporting less than 90 minutes of physical activity per week, (2) chronic instability of the lower extremity, (3) clinical hypermobility condition, (4) dizziness or balance problem, or (5) recent (<6 months) musculoskeletal injury. All volunteers provided written consent to participate and completed a physical activity readiness questionnaire to

screen for disqualifying injuries or medical conditions before data collection.

Procedures

Participants reported to the laboratory for a single data-collection session. After consent and completion of the physical activity readiness questionnaire, participants proceeded through the data collection in the following order: (1) balance, (2) range of motion, (3) FMS testing, (4) strength testing. This was determined to be the optimal order for minimizing the influence of fatigue on our outcomes. Testing procedures lasted approximately 2 hours.

Balance

Balance was tested in quiet, single-legged stance using a portable Accusway force plate and Balance Clinic software (Advanced Mechanical Technology, Inc, Watertown, MA). Participants were instructed to remain as motionless as possible for the duration of the test. Testing was conducted using the nondominant (ie, nonkicking) limb with hands on hips and eyes closed. Mediolateral and anteroposterior center-of-pressure (COP) coordinates were calculated from the raw force data sampled at 100 Hz. These data were used to create a resultant displacement time series, which was then differentiated with respect to time to yield a resultant COP velocity (CP_V) series. The mean of this velocity series was recorded for each participant. The first 10 seconds of CP_V data was used for analysis in this investigation.

Range of Motion

Range of motion was quantified using 3 validated clinical measures: (1) the Apley scratch test quantified range of motion in the shoulders and thoracic spine, (2) the sit-and-reach test measured hip and trunk flexibility, and (3) the weight-bearing-lunge test measured dorsiflexion range of motion.

The Apley scratch test closely mirrors the FMS shoulder-mobility test and has demonstrated high interrater (0.89–0.97) and intrarater (0.92–0.99) reliability.¹⁸ The test begins with participants standing with arms at their sides. When directed, the participant attempts to touch the hands together behind his or her back. With one hand, the participant reaches behind his or her head and down the back. The other hand reaches behind his or her lower back and up the spine. The distance between the participant's hands is measured with tape and recorded (in centimeters) as the score. In the present study, the average score of the left and right sides was used for analysis.

The sit-and-reach test was conducted using a 30.5-cm wooden box in accordance with the procedures outlined by Ayala et al,¹⁹ in which a test-retest reliability coefficient of 0.92 was reported. Participants sat on the floor with 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 22-cm 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.

Participants repeated the testing procedures until their scores stabilized to within 1 cm for 3 successive efforts.

The weight-bearing–lunge test was conducted according to the methods of Hoch et al.²⁰ This test began with the participant facing a wall and standing with the test foot aligned with a strip of measuring tape placed perpendicular to the wall. The nontest foot was stepped back 12 to 18 in (30.5–45.7 cm) for support. While keeping the heel of the test foot firmly on the ground, the participant was instructed to bend at the knee until his or her knee contacted the wall. After being familiarized with the task, participants 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. Interrater and intrarater correlation coefficients of 0.97 or greater have been reported for measuring the resulting angle or distance from the wall in this position.²¹ For this study, 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.

Strength

Strength-testing procedures were selected on the basis of reliability and safety for the population of interest. One-repetition–maximum testing is the criterion standard for strength assessment; however, a reliable estimate cannot be obtained during a single session in untrained populations²² and may additionally be unsafe for these individuals.²³ We therefore used alternative methods that are feasible in untrained populations and are strongly associated with their repetition-maximum analogs.

The modified YMCA bench-press test was used to assess upper body strength. The test was conducted using sex-specific standardized weights: 36.4 kg (80 lb) for men and 15.9 kg (35 lb) for women.²³ The test began with the participant positioned on a standard weight bench grasping the bar at a comfortable position. A metronome was then set to 60 beats/min and the participant was instructed to perform bench presses at 30 repetitions/min such that each beat of the metronome coincided with the bar reaching the up (fully extended) or down (bar-on-chest) position. When the participant was no longer able to maintain the 30-repetitions/min cadence or could no longer continue, that number of repetitions was recorded as the final score. This score is a strong predictor of 1-repetition–maximum bench-press loads (Pearson $r = 0.87$, average standard error of the estimate = 5.55).²³ A truncated familiarization trial was performed so as to allow participants to become accustomed to the weight and cadence of the test.

Countermovement-jump peak power has been shown to estimate 1-repetition–maximum back squat with high fidelity (test-retest reliability = 0.98, Pearson $r = 0.92$)²⁴ and was therefore selected to assess lower body strength. Each jump test requires the participant to exert maximal effort. Participants were allotted 1 practice trial and 3 test trials with approximately 1 minute of rest between efforts. They began by standing on a force plate (model 4060-NC; Bertec Corp, Columbus, OH) with hands on hips. When instructed, they crouched to a preferred depth (countermovement), immediately jumped as high as possible, and then landed on the force plate. Vertical ground reaction force was sampled at 1000 Hz and low-pass filtered at 40

Hz using The MotionMonitor software (Innovative Sports Training Inc, Chicago, IL). Data were recorded from the sampling buffer starting 1 second before the activation of a threshold trigger that marked the initiation of the countermovement. Because the first 1 second of data corresponded to quiet standing, we assumed that initial center-of-mass velocity was zero. Instantaneous velocity was then calculated using the forward-dynamics approach.²⁵ 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 used for analysis.

Functional Movement Screen

After a familiarization round, the FMS^{4,5} was administered both under conventional conditions (FMS_C) and while wearing an adjustable vest weighing 18.10 kg (FMS_W; MiR Vest Inc, San Jose, CA). This is comparable with loads used in other studies^{15,26} involving military personnel but may be less than the average combat loads in recent conflicts. We elected to use a standardized 18.10-kg load as this amount is sufficient to challenge FMS performance⁹ and has a basis in previous research.²⁶ For men and women respectively, the vest added an average of 22.72% or 28.29% of the participant's body weight. The weighted-vest and control conditions were randomized.

The FMS was administered by an experienced, although not FMS certified, member of the investigative team with established reliability in the relevant measures (Cohen $\kappa = 0.67$ –1.00; S.M.G. et al, unpublished data, 2015). The tests administered are listed in the “Results” and were scored according to the following criteria: (1) participant was unable to complete the movement; (2) participant was able to complete the task with errors noted; (3) participant was able to complete the task without error.

Statistical Analysis

In order to compare our results with previous work, we tested decreases in FMS item scores related to the weighted-vest condition with directional Wilcoxon signed-rank tests for matched pairs. We then tested our hypotheses concerning effect modification 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 in the following paragraphs.) Other options for analyzing this type of data include ordinal logistic regression or multinomial logistic regression. Our choice of statistical approach was based on the need to account for within-subjects effects and interaction effects in the same model. We further required a suitable means of performing model selection in a high-dimensional predictor space.

We modeled differential covariate effects in the 2 testing conditions using a varying coefficients structure.²⁷ Although this type of model is traditionally used to analyze effects that vary over time, it can be applied similarly to analyze effects that vary by condition.²⁷ Regardless of the order in which the tests were administered, the design matrix was specified such that data from the control condition were modeled as the first of 2 coefficients for each variable. The second coefficient, corresponding to the

Table 1. Descriptive Summary of Analyzed Covariates

Variable	Mean ± SD
Age, y	22.98 ± 3.13
Height, cm	172.11 ± 11.18
Mass, kg	71.82 ± 14.18
Weight-bearing–lunge test, cm	9.83 ± 3.83
Apley scratch test, cm	13.13 ± 7.43
Sit-and-reach test, cm	27.22 ± 10.18
YMCA bench-press test, No. of repetitions	27.80 ± 16.43
Countermovement jump, W	3308.63 ± 1025.04
Resultant center-of-pressure velocity, cm/s	4.96 ± 1.46

Abbreviation: YMCA, Young Men's Christian Association.

weighted-vest condition, represented the change in the effect of the covariate relative to the control condition. Thus, this coefficient can be interpreted as a covariate × condition interaction term. Note that, unlike the examples offered by Hess et al,²⁷ the modifying factor in our model was not time but rather weighted-vest condition. Because all data were collected on the same day, our set of independent variables was the same for each model. This is reasonable because we did not expect intrinsic performance attributes to vary within the span of a few minutes and bias associated with condition order is accommodated through randomization.

We hypothesized that the decrease in weighted-vest FMS item scores relative to the control condition would be smaller for those participants showing greater levels of balance, range of motion, and strength. With 2 exceptions, this would be visible in our models as a positive relationship between the item score and our 3 mediators at time point 2. The 2 exceptions are the Apley scratch test and resultant CP_v, for which higher values were interpreted as worse and the predicted signs of their coefficients in the weighted-vest condition were therefore negative.

Several nuisance variables, as well as their interactions with the weighted-vest condition, were also accounted for in our models: age, sex, height, and mass. It is apparent from our list of independent variables that the models of interest in our study were likely substantially underpowered for conventional regression techniques. This was especially true for the detection of interaction effects. Model selection under these circumstances can be achieved through penalization.²⁸ Penalized regression methods minimize an error term just as more familiar forms of regression do but are subject to additional constraints on the magnitude of the coefficients. These constraints are incorporated using a

data-driven tuning parameter, here denoted as Λ , which is usually selected on the basis of some information criterion or cross-validation procedure. The effect of using 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. Though estimation and significance testing are performed at an optimized value of Λ , relative importance among variables can be approximated by observing the order in which predictors are retained in the model as the penalty is progressively relaxed from a point at which all coefficients are equal to zero.²⁷

In this investigation, the tuning parameter (Λ) associated with minimum cross-validation error 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. All analyses in the present study were conducted using R (The R Foundation, Vienna University of Economics and Business, Vienna, Austria) with add-on packages *grpreg*²⁸ and *boot*.²⁹ A significance level of $\alpha = .05$ was specified a priori.

RESULTS

The weighted-vest condition was associated with a decrease in item scores for each FMS test except the hurdle step. A descriptive summary of the predictor variables modeled in this study is presented in Table 1. A count summary and paired difference tests for FMS item scores in the 2 experimental conditions are shown in Table 2. Model summary statistics are provided in Table 3. With the exception of the rotary-stability test, each model was significant at the .05 level and accounted for a moderate to large proportion of the variance (adjusted $R^2 = 0.22$ – 0.77). Exponentiated coefficients for individual predictors are presented in Table 4. These coefficients may be interpreted as the factors by which the outcome score is expected to 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 than or less than 1 correspond to positive and negative effects, respectively. For a given model, the relative importance of the various predictors after accounting for the nuisance parameters can be seen in Table 5, which shows the order in which variables were

Table 2. Summary and Paired Differences in Item Scores for Each Functional Movement-Screen Test^a

Test	Control Rating			Weighted-Vest Rating			Wilcoxon Signed Rank Test V Statistic ^b	P Value
	1	2	3	1	2	3		
Deep squat	6	31	13	13	27	10	71.5	.002 ^c
Hurdle step	1	29	20	1	30	19	36.0	.40
In-line lunge	0	18	32	3	25	22	120.0	.002 ^c
Shoulder mobility	2	15	33	25	24	1	1081.0	<.001 ^c
Active straight-leg raise	2	16	32	4	19	27	40.0	.01 ^c
Trunk-stability push-up	11	2	37	23	15	12	378.0	<.001 ^c
Rotary stability	0	42	8	1	46	3	44.0	.03 ^c

^a Total counts are shown for each item score category (1–3) under both conditions (control and weight vest).

^b Wilcoxon signed rank test statistics are shown for differences between conditions. The alternative hypothesis was directional (ie, weighted-vest condition < control condition).

^c Different ($P < .05$).

Table 3. Penalty Parameters (λ) and Final Model Summary Statistics

Outcome	λ	Features	F Value	Degrees of Freedom	P Value	R^2 Value
Deep squat	0.014	9	9.125	9,90	<.001	0.42
Hurdle step	0.019	9	5.477	9,90	<.001	0.29
In-line lunge	0.011	11	3.490	11,88	<.001	0.22
Shoulder mobility	0.005	15	22.620	15,84	<.001	0.77
Active straight-leg raise	0.017	9	8.330	9,90	<.001	0.40
Trunk-stability push-up	0.028	8	13.579	8,91	<.001	0.50
Rotary stability	0.021	6	0.814	6,93	.56	0.00

selected as the penalty parameter was incrementally relaxed.

Nuisance Parameters

In general, body weight was the most influential nuisance covariate, reducing test performance in the deep squat, active straight-leg raise, and trunk-stability push-up versus increasing performance in the shoulder-mobility test (Table 4). Sex was also differential in its effect depending on the test. Male sex was associated with poorer performance in the deep squat (Table 4) and better performance in the trunk-stability push-up (Table 4), each of these being relatively strong effects. Height was predictive of better hurdle-step (Table 4) performance. No effects related to height, body weight, or sex indicated a mediating role on

the effect of the weight vest. Age was the only nuisance parameter in which a mediating effect was observed, with lower scores in the weighted-vest condition of the shoulder-mobility test (Table 4) being associated with greater age. No general effects of age were observed.

Balance, Range of Motion, and Strength

A significant effect was observed for the YMCA bench-press repetitions, for which a greater number of repetitions was predictive of higher item scores in the deep-squat (Table 4) and trunk-stability push-up (Table 4) tests. For the countermovement jump, higher peak power was predictive of lower hurdle-step scores (Table 4).

High scores on the weight-bearing-lunge test predicted better performance in the deep squat and in-line lunge (Table 4). In the weighted-vest condition specifically, the same variable was predictive of greater shoulder-mobility (Table 4) test performance. Higher scores on the sit-and-reach test were associated with better performance in the active straight-leg raise (Table 4). Finally, lower (ie, better) Apley scratch test scores predicted better performance in the shoulder-mobility (Table 4) test.

Lastly, greater mean CP_V was a significant predictor of better hurdle-step performance (Table 4).

DISCUSSION

Paired differences in FMS item scores closely mirrored previously reported changes.⁹ One noteworthy finding concerns the distribution of scores in the 2 conditions. Particularly in the control condition, very few scores of 1 were observed. This may support what has been suggested regarding an inability of clinical movement tests to differentiate performance at higher levels.⁷ The load-bearing treatment used in this study may therefore have served to normalize the score distribution.

Previous authors⁹ have shown that FMS item scores, except for the trunk-stability push-up, were not related to tactical criterion performance tasks unless the screen was performed with an external load. Considering the possibility that the additional load highlighted performance-relevant attributes during the screening process, we hypothesized that the external-load treatment preferentially taxed those participants with relatively low levels of balance, range of motion, and strength. We therefore hypothesized that high levels of balance, range of motion, and strength would be associated with smaller decreases in FMS item scores when comparing FMS_W with FMS_C . This hypothesis would be supported by coefficients corresponding to the weighted-vest condition (“*W”) being selected early and exhibiting a relatively large effect in the appropriate direction, which was generally not the case.

Table 4. Coefficients and t Statistics for All Models

Outcome	Variable	Coefficient	t Value	P Value
Deep squat	Sex	1.27	2.60	.01
	Mass	0.99	-2.13	.04
	Intercept \times weighted-vest condition ^a	0.89	-2.23	.03
	Weight-bearing-lunge test	1.06	6.98	<.001
	YMCA bench-press test	1.01	2.97	.003
Hurdle step	Height	1.01	3.10	.003
	Countermovement jump	1.00	-2.11	.04
	Resultant center-of-pressure velocity	1.04	2.61	.01
In-line lunge	Mass	0.99	-2.81	.006
	Weight-bearing-lunge test	1.03	3.16	.002
Shoulder mobility	(Intercept)	3.26	2.72	.007
	Age \times weighted-vest condition	0.98	-2.76	.007
	Mass	1.01	2.19	.03
	Weight-bearing-lunge test \times weighted-vest condition	1.02	2.31	.02
	Apley scratch test	0.97	-4.74	<.001
Active straight-leg raise	(Intercept)	3.17	2.38	.02
	Mass	0.99	-2.31	.02
	Sit-and-reach test	1.02	5.52	<.001
Trunk-stability push-up	Sex	2.03	7.33	<.001
	Mass	0.99	-2.38	.02
	YMCA bench-press test	1.01	4.28	<.001
Rotary stability	Model not significant			

Abbreviation: YMCA, Young Men’s Christian Association.

^a The intercept \times weighted-vest condition represents the global incremental or decremental effect of the weighted-vest condition compared with the control condition.

Table 5. Order of Variable Selection^a Extended on Next Page

Rank	Deep Squat	Hurdle Step	In-Line Lunge
1	Age ^b	Age ^b	Age ^b
1	Height ^b	Height ^b	Height ^b
1	Sex ^b	Sex ^b	Sex ^b
1	Mass ^b	Mass ^b	Mass ^b
5	Weight-bearing–lunge test ^b	Resultant center-of-pressure velocity ^b	Weight-bearing–lunge test ^b
6	Intercept × weighted-vest condition ^b	Apley scratch test ^b	Intercept × weighted-vest condition
7	YMCA bench-press test ^b	Countermovement jump ^{b,c}	Resultant center-of-pressure velocity ^{b,c}
8	Apley scratch test ^b	Sit-and-reach test ^{b,c}	YMCA bench-press test ^{b,c}
9	Resultant center-of-pressure velocity ^b	Sit-and-reach test × weighted-vest condition ^b	Age × weighted-vest condition
10	Countermovement jump ^d	Weight-bearing–lunge test	Weight-bearing–lunge test × weighted-vest condition ^b
11	Sit-and-reach test ^d	Sex × weighted-vest condition	Countermovement jump ^b
12	YMCA bench-press test × weighted-vest condition ^d	YMCA bench-press test × weighted-vest condition	Height × weighted-vest condition ^b
13	Apley scratch test × weighted-vest condition	Resultant center-of-pressure velocity × weighted-vest condition ^d	YMCA bench-press test × weighted-vest condition ^b
14	Countermovement jump × weighted-vest condition ^e	YMCA bench-press test ^d	Apley scratch test
15	Sex × weighted-vest condition ^e	Intercept × weighted-vest condition	Resultant center-of-pressure velocity × weighted-vest condition
16	Age × weighted-vest condition	Mass × weighted-vest condition	Mass × weighted-vest condition
17	Resultant center-of-pressure velocity × weighted-vest condition	Apley scratch test × weighted-vest condition ^e	Countermovement jump × weighted-vest condition
18	Weight-bearing–lunge test × weighted-vest condition	Height × weighted-vest condition ^e	Apley scratch test ^{c,d}
19	Height × weighted-vest condition	Weight-bearing–lunge test × weighted-vest condition	Sit-and-reach test × weighted-vest condition ^d
20	Sit-and-reach test × weighted-vest condition	Countermovement jump × weighted-vest condition	Sit-and-reach test
21	Mass × weighted-vest condition	Age × weighted-vest condition	Sex × weighted-vest condition

Abbreviation: YMCA, Young Men’s Christian Association.

^a Relative importance of covariates is shown as a function of increasing the penalty parameter (λ). The nuisance variables (age, sex, height, and mass) were not penalized; therefore, they were present in all models.

^b Variable was selected in the final model. Variables that were selected but not shown in Table 4 did not reach statistical significance in their coefficients.

^c Variables that share this superscript were selected in the same iteration.

^d Variables that share this superscript were selected in the same iteration.

^e Variables that share this superscript were selected in the same iteration.

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 selected was the weight-bearing–lunge test in the shoulder-mobility model, for which a positive relationship was observed. This is not an unreasonable finding, as both tests depend on range of motion; however, if any single range-of-motion predictor was to be selected for this outcome, we would expect it to be the Apley scratch test.

In contrast, several of our mediator variables showed noteworthy global effects. The weight-bearing–lunge test was the most important predictor of the deep squat, supporting what has been suggested by previous researchers,¹¹ as well as the in-line lunge. These movements are particularly susceptible to bottom-up deviation cascades in

which dorsiflexion range of motion is limited. 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.

Our balance variable, mean CP_V , was retained as the most important variable in the hurdle-step model. Traditionally, greater CP_V would be interpreted to reflect poorer balance control.³⁰ However, in this case, it predicted higher hurdle-step scores. Although balance is one of the attributes purported to be assessed during clinical movement screens, data relating static balance and FMS component tests are limited. One group⁸ found no relation between in-line–lunge scores and COP excursion, whereas another³¹ observed an inverse relationship between hurdle-step

Table 5. Extended From Previous Page

Shoulder Mobility	Active Straight-Leg Raise	Trunk-Stability Push-Up	Rotary Stability
Age ^b	Age ^b	Age ^b	Age ^b
Height ^b	Height ^b	Height ^b	Height ^b
Sex ^b	Sex ^b	Sex ^b	Sex ^b
Mass ^b	Mass ^b	Mass ^b	Mass ^b
Apley scratch test × weighted-vest condition ^b	Sit-and-reach test ^b	Intercept × weighted-vest condition ^b	Weight-bearing–lunge test × weighted-vest condition ^b
Intercept × weighted-vest condition	Age × weighted-vest condition ^b	Sit-and-reach test × weighted-vest condition ^b	Sit-and-reach test × weighted-vest condition ^b
Age × weighted-vest condition ^b	Weight-bearing–lunge test ^b	YMCA bench-press test ^b	YMCA bench-press test
Apley scratch test ^b	Apley scratch test × weighted-vest condition ^b	YMCA bench-press test × weighted-vest condition ^b	Sit-and-reach test
Sit-and-reach test × weighted-vest condition ^b	YMCA bench-press test ^b	Resultant center-of-pressure velocity	Apley scratch test × weighted-vest condition
YMCA bench-press test ^b	Resultant center-of-pressure velocity	Apley scratch test × weighted-vest condition	Weight-bearing–lunge test
Countermovement jump × weighted-vest condition	Weight-bearing–lunge test × weighted-vest condition	Countermovement jump × weighted-vest condition	YMCA bench-press test × weighted-vest condition
Resultant center-of-pressure velocity ^b	YMCA bench-press test × weighted-vest condition	Weight-bearing–lunge test	Apley scratch test
YMCA bench-press test × weighted-vest condition ^b	Apley scratch test	Weight-bearing–lunge test × weighted-vest condition	Countermovement jump
Sit-and-reach test ^b	Resultant center-of-pressure velocity × weighted-vest condition	Countermovement jump ^c	Countermovement jump × weighted-vest condition
Weight-bearing–lunge test ^{b,d}	Sit-and-reach test × weighted-vest condition	Sit-and-reach test ^c	Age × weighted-vest condition
Weight-bearing–lunge test × weighted-vest condition ^{b,d}	Sex × weighted-vest condition	Height × weighted-vest condition	Resultant center-of-pressure velocity
Sex × weighted-vest condition	Countermovement jump	Apley scratch test	Height × weighted-vest condition
Countermovement jump ^b	Mass × weighted-vest condition	Mass × weighted-vest condition	Intercept × weighted-vest condition
Mass × weighted-vest condition	Countermovement jump × weighted-vest condition	Sex × weighted-vest condition	Mass × weighted-vest condition
Resultant center-of-pressure velocity × weighted-vest condition	Height × weighted-vest condition	Resultant center-of-pressure velocity × weighted-vest condition	Resultant center-of-pressure velocity × weighted-vest condition
Height × weighted-vest condition	Intercept × weighted-vest condition	Age × weighted-vest condition	Sex × weighted-vest condition

performance and COP standard deviation. Previous data from our laboratory⁹ indicated that anteroposterior CP_V, albeit from double-legged standing, was positively associated with higher scores in the deep squat, weighted deep squat, and weighted in-line lunge. This was initially interpreted as a spurious result potentially related to the small sample size or the use of a double-legged–standing protocol, which may lack discriminatory ability in young, healthy populations. Although it is difficult to directly compare the present results with the previous findings, the pattern suggests at least 2 possibilities. First, it could be that the variance in CP_V that is predictive of lower FMS component scores is actually related to a confounder variable such as height or weight. This possibility seems unlikely based on our control and model selection procedures. Alternatively, higher CP_V in this data set may actually reflect better postural control. It is possible that balance limitations can result in compensatory decreases in postural motion when individuals are not confident in exploring their postural-control space. Although it is somewhat paradoxical to assert that the participants who followed instructions better actually demonstrated worse performance, research³² suggested that increases in “ex-

ploratory” sway behavior may play a functional role in maintaining balance when sensory information is lacking. The lower CP_V in our sample may also indicate a more constrained postural-control strategy that restricted the performance of dynamic tasks.³³

One limitation that should be considered in interpreting the present findings concerns the measures used to evaluate performance attributes. These measures were chosen on the basis of their reliability and feasibility among the population of interest. However, other performance metrics may exist that more precisely quantify attributes associated with functional movement or other aspects of performance may be more relevant for a given clinical movement test. Specifically, outcomes related to agility, power, and coordination may play a role. These can be evaluated through timed cone drills, medicine-ball throws, or Olympic-style weightlifting, among other means.

Lastly, we note that the FMS was proposed to assess efficiency or quality of movement and was presented as a potentially separate component of performance that might complement multidimensional models such as that of Fleishman.¹⁰ In contrast, the hypotheses in our investigation

considered clinical movement screens to be a convenient, feasible method of observing previously identified performance domains. Although some of the effects presented may be interpreted to support the notion of a latent movement-quality capacity, we interpreted the variation in predictors and directionality to suggest that the FMS items did not load on a single factor. This is consistent with the authors³⁴ of a large-scale analysis who concluded that the underlying structure of the FMS composite score was not unidimensional.

CONCLUSIONS

Our findings confirm that a moderate to large portion of the variance ($R^2 = 0.22\text{--}0.77$) in FMS scores was explained through models that included balance, range-of-motion, and strength 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 balance, range of motion, and strength prevented movement-quality decreases related to external loading. This suggests that the differential abilities of FMS_C and FMS_W in predicting tactical performance outcomes are perhaps attributable to other factors. In conclusion, although load-enhanced clinical scoring of movement behaviors may be a viable means of predicting physical performance, further research is needed to understand the complex relationships among movement-quality and performance attributes. In addition to the agility, power, and coordination outcomes mentioned earlier, future authors should also analyze other features of clinical movement tests. Such features may include more specifically coded behaviors similar to the criteria for loss of points in FMS component tests or semistructured participant feedback regarding perceived effort and degree of difficulty required to complete a given task.

REFERENCES

- Deuster PA, O'Connor FG, Henry KA, et al. Human performance optimization: an evolving charge to the Department of Defense. *Mil Med*. 2007;172(11):1133–1137.
- Lisman P, O'Connor FG, Deuster PA, Knapik JJ. Functional movement screen and aerobic fitness predict injuries in military training. *Med Sci Sports Exerc*. 2013;45(4):636–643.
- 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(2):146–153.
- 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(3):132–139.
- 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(2):62–72.
- Thelen MD, Koppenhaver SL, Hoppes CW, Shutt C, Musen JL, Williams MK. Reliability of a novel return to duty screening tool for military clinicians. *US Army Med Dep J*. October–December 2015:14–23.
- 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(7):862–873.
- 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(3):197–202.
- Glass SM, Ross SE. Modified functional movement screening as a predictor of tactical performance potential in recreationally active adults. *Int J Sports Phys Ther*. 2015;10(5):612–621.
- Fleishman EA. *The Structure and Measurement of Physical Fitness*. Englewood Cliffs, NJ: Prentice-Hall; 1964.
- Butler RJ, Plisky PJ, Southers C, Scoma C, Kiesel KB. Biomechanical analysis of the different classifications of the Functional Movement Screen deep squat test. *Sports Biomech*. 2010;9(4):270–279.
- 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(1):115–120.
- 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(4):867–873.
- Yaggie JA, Campbell BM. Effects of balance training on selected skills. *J Strength Cond Res*. 2006;20(2):422–428.
- Heller MF, Challis JH, Sharkey NA. Changes in postural sway as a consequence of wearing a military backpack. *Gait Posture*. 2009;30(1):115–117.
- De Maio M, Onate J, Swain D, Morrison S, Ringleb S, Naiak D. Physical performance decrements in military personnel wearing personal protective equipment (PPE). <http://www.dtic.mil/dtic/tr/fulltext/u2/a567881.pdf>. Accessed June 27, 2018.
- Park H, Noll G, Branson D, Peksoz S, Petrova A, Goad C. Impact of wearing body armor on lower body mobility. *Cloth Textiles Res J*. 2011;29(3):232–247.
- Ross M, Paterno G. Reliability of the Apley's range of motion test and dominant and nondominant upper extremity differences. *Eur J Phys Med Rehabil*. 1998;8(6):166–169.
- Ayala F, Sainz de 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(4):219–226.
- Hoch MC, Staton GS, McKeon PO. Dorsiflexion range of motion significantly influences dynamic balance. *J Sci Med Sport*. 2011;14(1):90–92.
- Bennell K, Talbot R, Wajswelner H, Techovanich W, Kelly D, Hall AJ. Intra-rater and inter-rater reliability of a weight-bearing lunge measure of ankle dorsiflexion. *Aust J Physiother*. 1998;44(3):175–180.
- 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(5):1418–1422.
- 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(3):440–445.
- 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(3):534–539.
- 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(3):668–674.
- Harman EA, Gutekunst DJ, Frykman PN, et al. Prediction of simulated battlefield physical performance from field-expedient tests. *Mil Med*. 2008;173(1):36–41.
- Hess W, Persson M, Rubenbauer S, Gertheiss J. Using lasso-type penalties to model time-varying covariate effects in panel data regressions: a novel approach illustrated by the “Death of Distance”

- in international trade. *Research Institute of Industrial Economics*. 2013; Working Paper No. 961.
28. Breheny P, Huang J. Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. *Stat Comput*. 2015;25(2):173–187.
 29. Davison AC, Hinkley DV. *Bootstrap Methods and Their Applications*. Cambridge University Press: Cambridge, UK; 1997.
 30. 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(1):15–23.
 31. Clifton DR, Harrison BC, Hertel J, Hart JM. Relationship between functional assessments and exercise-related changes during static balance. *J Strength Cond Res*. 2013;27(4):966–972.
 32. Carpenter MG, Murnaghan CD, Inglis JT. Shifting the balance: evidence of an exploratory role for postural sway. *Neuroscience*. 2010;171(1):196–204.
 33. Riccio GE, Stoffregen TA. Affordances as constraints on the control of stance. *Hum Mov Sci*. 1988;7(2–4):265–300.
 34. 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(3):672–678.

Address correspondence to Stephen M. Glass, PhD, Department of Otolaryngology, The Ohio State University, 915 Olentangy River Road, Columbus, OH 43212. Address e-mail to stephen.glass@osumc.edu.