Analysis in Support of the Women in Service Restriction Review Study

Jason Jameson and Karen Kelly

Naval Health Research Center

San Diego, CA

Brian McGuire and Leon M. Pappa

U.S. Marine Corps Training and Education Command

Quantico, VA

SUMMARY

Problem

Military career advancement depends crucially on serving in combat roles, which is a restriction that has placed inherent limits on a woman's ability to advance to leadership roles. The Pentagon is currently removing these restrictions across all branches of the military. To ensure that this integration effort proceeds fairly while still maintaining the military's operational readiness, several questions related to *physical readiness* need to be answered: (1) What tests are available for estimating whether someone is physically prepared for combat? (2) How effective are these tests? (3) Are there substantial differences in performance between men and women on these tests? (4) Are these differences large enough that they indicate particularly challenging obstacles for women to overcome before they can serve in combat? What are these areas? Do they relate to strength, to endurance, or to particular types of strength? Answers to these questions naturally imply another: If these obstacles exist, can women overcome them?

Objective

In support of the U.S. Marine Corps, the objective of this report was to provide preliminary answers to these questions.[1] The focus of this early effort was to identify physical ability tests that might be used, and how effective they might be, for estimating physical preparedness for combat. Specifically, the Physical Fitness Test (PFT) and Combat Fitness Test (CFT) were identified as candidate predictors of physical combat readiness. The information gathered for this effort was also used to identify performance discrepancies between men and women, thus helping to identify specific problem areas for

---

[1] The last question is beyond the scope of this report, but it is a natural extension of this study, and we discuss it in the concluding section of this paper.

women. This information was then used to recommend minimum combat readiness standards for each

selection test.

Approach

To determine whether the candidate physical tests of combat readiness (i.e., the PFT and CFT)

were effective as predictors, a benchmark measure of combat performance was required. To meet this

need, the U.S. Marine Corps Training and Education Command developed several proxy tasks that

captured essential physical capabilities required for combat, including tasks that were taken from the

existing Training and Readiness requirements. These tasks were developed to serve as accurate

reflections of the occupational demands Marines face in theater. To gather a comprehensive and

representative sample of Marine Corps personnel, information was gathered from several hundred

active-duty Marines at Marine Corps Recruit Depot Parris Island, North Carolina, the School of

Infantry-East at Marine Corps Base Camp Lejeune, North Carolina, and The Basic School, at Marine

Corps Base Quantico, Virginia. This information consisted of performance outcomes on the combat

proxy tasks (CPTs; on either the individual proxy tasks, which were pass/fail, or a derived composite

score, calculated as the percentage of proxy tasks successfully completed), and archived data composed

of performance outcomes on the PFT and CFT component tasks (e.g., 3-mile run, flexed-arm hang;

ammunition lift, movement to contact, respectively). Of particular importance, the data set comprised

roughly equal numbers of men and women. Scores on the PFT and CFT component tasks were

compared to performance on the CPTs, and formal measures of predictive value (i.e., predictive validity,

as measured by statistical correlation) were derived.[2] Finally, minimum screening standards for combat

readiness were suggested for each of the component PFT and CFT tests after examining the group of

Marines deemed most prepared for combat. From this group—defined as those who had completed more

---

[2] The data set was analyzed by researchers at the Naval Health Research Center.

than 75% of the CPTs, and included both men and women—the worst performing individuals were isolated (defined by the worst performing decile). Mean performance of the weakest performers in the group deemed most prepared for combat was then calculated and adopted as the recommended minimum score for combat readiness.

Results and Discussion

Overall, the PFT and CFT are valid predictors of performance on the CPTs. Moreover, the CFT does a better job than the PFT of predicting performance on the CPTs. This is to be expected, since the CFT was intended to be a physical test battery that better reflected operational readiness. While the current data suggest that the selection tests are strong predictors of combat readiness, it is important to clarify the scope of the current investigation. First, the ceiling placed on the current scoring system for the PFT and CFT (e.g., once a Marine gets a maximum score, the test is terminated) is a common characteristic of large-scale field tests of physical fitness, but it artificially restricts the range of possible values individuals can obtain. This restriction in the range of possible values leads to underestimates of predictive validity.[3] To a certain extent, this issue does not affect the final conclusions of this study: since range restriction leads to *underestimated* predictive validities, the biased, calculated values can simply be viewed as very conservative estimates of the true predictive validity. Employing conservative estimates provides an even more severe and arguably better test of a predictor's value: if a given selection test can pass a very conservative criterion for what counts as a good predictor, then surely it would pass a more liberal one, and thus is likely an even better predictor than the results suggest.

Second, for the strength-based proxy tasks, weights were chosen to be reflective of the occupational demands Marines face during actual field operations. Further, to meet the logistical

---

[3] Sackett, P. R., & Yang, H. (2000). Correcting for range restriction: An expanded typology. *Journal of Applied Psychology, 85,* 112-118.

demands involved in testing hundreds of Marines, while still maintaining fidelity to operational

demands, maximum testing weights were implemented. As a result, in many cases, most Marines

(including both men and women) could complete the proxy tasks successfully. The table below provides

the percentages of successful completion for both men and women. These results can be interpreted as

confirmation that training is working: the majority of Marines were able to meet the strength and

endurance demands of combat, as measured by the proxy tasks.

Third, the relatively high rate of success across tasks does highlight, however, a noteworthy

discrepancy in performance between men and women for two of the tasks: pull-ups and the Clean and

Press. The difference in performance likely reflects sizeable differences in upper-body strength. An

important question to answer, for future studies, is just how much improvement could be expected for

women if they were specifically trained on upper-body strength. Currently, 12% of women were able to

complete 10 or more pull-ups, and a very small percentage, about 2%, could do 15 or more pull-ups. The

results suggest that lower-body strength and cardiorespiratory fitness are not the problem. Instead, the

biggest hurdle facing women is upper-body strength, and perhaps a lack of familiarity with the

appropriate technique to complete the tasks successfully (such as for the Clean and Press).

Pull-Ups		Deadlift (135 lb)		Clean & Press (115 lb)		120 mm Tank Loading		155 mm Round Carry		Obstacle Course Wall	
Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
409	378	409	378	409	378	409	378	409	378	409	378
15.69	3.57	100.0%	97.1%	80.0%	8.7%	99.8%	81.5%	99.8%	71.4%	98.8%	78.6%

INTRODUCTION

For decades women have been restricted from certain Military Occupational Specialties (MOSs), such as infantry, though there has been a gradual push, since World War II, to open up closed occupations to women.[4] However, military career advancement depends crucially on serving in combat roles, which is a restriction that has placed inherent limits on a woman's ability to advance to leadership roles. Based on Section 535 of the National Defense Authorization Act for Fiscal Year 2011, the Secretary of Defense has been directed to review the status of "laws, policies, and regulations, including collocation policy, that may restrict the services of female members of the Armed Forces to determine whether changes in such laws, policies, and regulations are needed to ensure that female members have an equitable opportunity to compete and excel in the Armed Forces"[1]. The Pentagon is currently removing these restrictions across all branches of the military. To ensure that this integration effort proceeds fairly while still maintaining the military's operational readiness, several questions related to *physical readiness* need to be answered: (1) What tests are available for estimating whether someone is physically prepared for combat? (2) How effective are these tests? (3) Are there substantial differences in performance between men and women on these tests? (4) Are these differences large enough that they indicate particularly challenging obstacles for women to overcome before they can serve in combat? What are these areas? Do they relate to strength, to endurance, or to particular types of strength? Answers to these questions naturally imply another: If these obstacles exist, can women overcome them?

---

[4] For a review of this history see, "Restrictions on Assignments of Military Women: A Brief History," published by the National Women's Law Center, http://www.nwlc.org/resource/restrictions-assignments-military-women-brief-history

In support of the U.S. Marine Corps, the objective of this report was to contribute preliminary answers to these questions.[5] The focus of this early effort was in identifying physical ability tests that might be used, and how effective they might be, for estimating physical preparedness for combat. Specifically, the Physical Fitness Test (PFT) and Combat Fitness Test (CFT) were identified as candidate predictors of physical combat readiness. The information gathered for this effort was also used to identify performance discrepancies between men and women, thus helping to identify specific problem areas for women. This information was then used to recommend minimum combat readiness standards for each selection test.

To maintain physical readiness, all Marines, regardless of gender, are required to pass 2 the PFT and CFT. Each test is composed of 3 component tasks. For the PFT, these tasks are pull-ups (men only), flexed-arm hang (FAH; women only), crunches, and a 3-mile run. For the CFT, these tasks are an ammunition lift (AL), a timed 880-yard sprint (movement to contact [MTC]), and a timed shuttle run (maneuver under fire [MANUF]).

For the pull-ups portion of the PFT, men must complete as many well-executed pull-ups (PU) as possible, with 20 or more yielding a maximum score of 100 points (no time limit). In contrast, women are tested with the FAH, which measures a candidate's ability to maintain elbow flexion by hanging on the pull-up bar for as long as possible. The FAH requires a static arm hold, which involves a static contraction of the muscles, rather than the dynamic contraction involved in pull-ups. Thus, performance on the two tasks cannot be compared because different muscle groups are used. For the second task of the PFT, both men and women are tested on the number of crunches completed within a 2-minute time limit. Each sit-up is given1 point, up to a maximum score of 100. And finally, in the third PFT task, men

---

[5] The last question is beyond the scope of this report, but it is a natural extension of this study, and we discuss it in the concluding section of this paper.

and women are timed on a 3-mile run, with a completion time of 18 minutes, yielding a maximum score of 100. One point is deducted for every 10 seconds above 18 minutes (e.g., a time of 18:10 yields a score of 99).

The CFT was developed with the intention of creating a physical fitness test with tasks that directly resemble demanding combat operations. Three tasks are used to estimate readiness for the physical challenges of combat: a timed 880-yard sprint (MTC); an AL, which involves lifting a 30-pound ammo can overhead from shoulder height for as many repetitions as possible within 2 minutes; and a MANUF, a timed 300-yard shuttle run incorporating several combat-related tasks.

The primary aim of this study was to determine whether the benchmark physical fitness tests—the PFT and the CFT—can serve as valid predictors of successful completion of combat-related tasks. An additional aim of this study was to determine which benchmark physical fitness tests are most predictive of performance on these tasks. This information will then be used to guide inference about the appropriate minimum test standards to set for each selection test.

METHODS

Role and Responsibilities

The Training and Education Command (TECOM) was responsible for research design and methods, data collection, and study approval (Marine Corps Combat Development Command Institutional Review Board protocol #DoD N-40078., "Assessment of United States Marine Corps Closed Military Occupational Specialty Physical Performance Standards" [CMOS]). Data were collected and transferred to the Naval Health Research Center (NHRC), where Drs. Kelly and Jameson provided statistical analysis and data interpretation as well as a written report.

Data Collection

Data were collected by TECOM researchers (McGuire and Pappa) from three different sites: Marine Corps Recruit Depot Parris Island, South Carolina; School of Infantry-East at Marine Corps Base Camp Lejeune, North Carolina; and The Basic School, at Marine Corps Base Quantico, Virginia. To be included in the study, participants were required to be active-duty Marines and fit for full duty. Participants included 410 men and 380 women.[6] A briefing on the events was conducted, and all volunteers provided informed consent. The combat-related tasks were designated as the physical training (PT) for the day, which is a mandatory requirement for active-duty Marines. The volunteer rate was high (98%). A total of 380 female Marines and 410 male Marines were tested from the three locations. The bulk of Marines tested were students from School of Infantry-East, and therefore were closer in age and training status to new recruits.

Warm-Up and Testing

All participants participated in a dynamic warm-up prior to testing. The dynamic warm-up consisted of the following:

-50 m jog

-50 m backward run

-25 m walking toe touch

-25 m butt kickers

-25 m lateral shuffle

-10X body weight squats

---

[6] The analysis excluded from consideration 1 man and 2 women due to missing data.

-10X push-ups

-10X burpees

Following the warm-up period, participants engaged in the combat proxy tasks (CPTs). The order of the tests is as follows:

1. Maximum set of pull-ups

2. Deadlift

    – 1 rep max progressing ladder (60, 70, 80, 95, 115, 135 lb)

3. Clean and Press

    – 1 rep max progressing ladder (70, 80, 95, 115 lb, and 6 lifts of 65 lb)

4. 120 mm (replica) Tank Round Lift and Load

    – Lift and load five (5) 120 mm projectiles (replica rounds weighing 55 lb)

      in 35 seconds

5. 155 mm (replica) Artillery Round Lift and Carry

    – Pick up and carry a 155 mm projectile (replica round weighing 95 lb) a distance of 50 m

      in 2 minutes

6. Negotiate 7-foot Obstacle Course Wall (execute lower-level entry)

    – with a 20-inch assist box (to standardize the 1- or 2-person lift in a Military Operations

      on Urban Terrain environment), while wearing a fighting load (Kevlar helmet, flak jacket,

      plate carrier, and Small Arms Protective Inserts [SAPI])

Instructions and a demonstration of correct movement and posture of each task were provided to each participant prior to testing. Before each event, participants were given the opportunity to practice the events with lighter weights. For the Clean and Press and Deadlift tasks, incremental lifts were used

to gradually warm up the participants and advance according to their own comfort level as measured by their own self-report and observations by the investigators responsible for data collection.

Analysis of Closed MOS Physical Performance Standards Testing

The majority of test events performed in the "CMOS were either a pass/fail event for a maximum weight repetition or for a timed event. A total of 6 events were used: pull-ups, deadlifts, clean and presses, 120 mm tank loading drill, a 155 mm artillery round carry, and negotiating an obstacle course wall while wearing a fighting load.

RESULTS

The intent of the analysis was to determine the value of the PFT and CFT components as predictors of performance on the CPTs. The overall demographic characteristics (age, height, and weight) are provided in Table 1a. The demographic breakdown by gender is provided in Table 1b. Frequency distributions for scores on the PFT and CFT components tasks are provided in Appendix B.

Table 1a. Descriptive statistics for basic demographics

	N	Minimum	Maximum	Mean	SD
Age	788	17	42	22.39	4.442
Height	788	58	77	66.95	3.717
Weight	788	94	239	150.92	27.400

Table 1b. Descriptive statistics for basic demographics, by gender

	Gender	N	Mean	SD	SE
Age	Male	409	22.50	4.342	.215
	Female	379	22.27	4.551	.234
Height	Male	409	69.42	2.665	.132
	Female	379	64.29	2.715	.139
Weight	Male	409	167.57	25.442	1.258
	Female	379	132.95	15.517	.797

The descriptive statistics for the breakdown of the PFT, by overall score and by component task, are provided in Table 2. The CFT results are presented in Table 3.

Table 2. Descriptive statistics for the PFT overall score and component tasks

	N	Range	Minimum	Maximum	Mean	SD
PFT Pull-Ups	435	24	3	27	15.92	4.351
PFT FAH	353	45	25	70	65.56	8.614
PFT Crunches	788	55	45	100	96.30	8.846
PFT 3-Mile Run	788	863	994	1857	1372.20	154.754
PFT Overall Score	788	180	120	300	261.53	28.017

A pull-ups task that included both men and women was conducted. Overall, across both genders, the average number of pull-ups completed was 9.87 (SD = 8.00). By gender, the results are very different, illustrated in Figure 1 below.

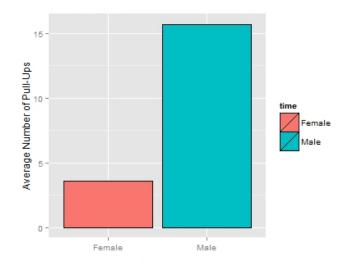**Figure 1. Average number of pull-ups, by gender**

Table 3. Descriptive statistics for the CFT overall score and CPTs

	N	Range	Minimum	Maximum	Mean	SD
CFT MTC	787	166	126	292	191.62	26.526
CFT AL	787	115	10	125	77.42	23.402
CFT MANUF	787	200	108	308	171.67	35.621
CFT Overall Score	787	123	177	300	286.32	14.230

The breakdown, by gender, for performance on the PFT is presented in Table 4, and for the CFT, in Table 5.

Table 4. Performance on the PFT, by gender

	Gender	N	Mean	SD	SE
PFT Crunches	Male	409	98.89	4.862	.240
	Female	379	93.50	11.058	.568
PFT 3-Mile Run	Male	409	1281.54	115.118	5.692
	Female	379	1470.04	130.632	6.710
PFT Overall Score	Male	409	259.63	27.710	1.370
	Female	379	263.58	28.238	1.450

Table 5. Performance on the CFT, by gender

	Gender	N	Mean	SD	SE
CFT MTC	Male	409	173.39	15.242	.754
	Female	378	211.34	21.563	1.109
CFT AL	Male	409	96.52	9.591	.474
	Female	378	56.76	14.784	.760
CFT MANUF	Male	409	145.19	18.386	.909
	Female	378	200.34	26.356	1.356
CFT Overall Score	Male	409	289.05	12.098	.598
	Female	378	283.37	15.715	.808

Results for the Combat Proxy Tasks

Overall, there was a high rate of successful completion on the CPTs, with the notable exception being the Clean and Press. The results are presented in Table 6.

Table 6. Overall percentage of successful completion on the CPTs

Mission Readiness Task	N	%. Successful Completion
Deadlift:135 lb	788	99%
155mm Artillery Round Lift & Carry	788	86%
120 mm Tank Round Lift & Load	788	91%
Obstacle Course Wall	788	89%
Clean & Press: 115 lb	788	46%

The relatively large drop for the Clean and Press can be attributed to a much smaller completion rate for women. A gender breakdown for performance on the CPTs is presented in Table 7 and Figure 2.

Table 7. Percentage of successful completion on the

mission readiness tasks, by gender

	Gender	N	% Successful Completion
Deadlift:115 lb	Male	409	100%
	Female	379	99%
155 mm Artillery Round Lift & Carry	Male	409	100%
	Female	379	72%
120 mm Tank Round Lift & Load	Male	409	100%
	Female	379	82%
Obstacle Course Wall	Male	409	99%
	Female	379	79%
Clean & Press: 115 lb	Male	409	80%
	Female	379	9%

Figure 2. Percentage of successful completion on the mission readiness tasks, by gender

Given the unusually low rate of completion, let us consider where the breakdown occurred. More detailed results regarding the Clean and Press are presented in Table 8 and Figure 3.

Table 8. Percentage of successful completion on the MRTs, by gender

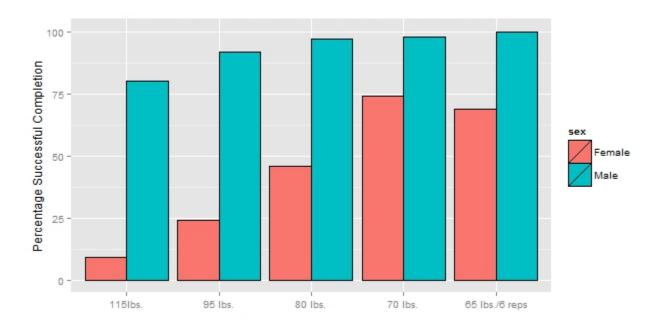|	Gender	N	% Successful Completion
Clean & Press: 115 lb	Male	409	80%
	Female	379	9%
Clean & Press: 95 lb	Male	409	92%
	Female	379	24%
Clean & Press: 80 lb	Male	409	97%
	Female	379	46%
Clean & Press: 70 lb	Male	409	98%
	Female	379	74%
Clean & Press: 6 reps 65 lb	Male	409	100%
	Female	379	69%

Figure 3. Percentage of successful completion on the Clean and Press, by gender

As these results show, men were more successful than women on the Clean and Press, with the substantial breakdown for women occurring between 70 and 80 lb.

Estimating the Value of the PFT and CFT Components Tasks

To estimate the value of the PFT and CFT component tasks as predictors of combat readiness, statistical correlations were computed between the component tasks and (1) an overall mission readiness score (the percentage of tasks successfully completed, computed across all CPTs), and (2) the individual tasks themselves (see Tables 9a and 9b). Both Pearson product-moment correlation coefficients and Spearman's ρ correlations were computed; placing maximum scores on several of the tests led to restricted range and non-normal distributions. Spearman's ρ is the nonparametric measure of

association that does not require normality and partially corrects for range restriction.[7] Green indicates excellent predictors; light green, very good predictors; and yellow, good (meaning valid) predictors. As an indication of validity, a standard interpretive benchmark will be adopted: correlations in the range of .30 to .40 are considered valid predictors.[8]

Table 9a. Correlations between the PFT and CFT component tasks and overall mission readiness (Pearson's *r*)

Pearson's *r*	Crunches	3-Mile Run	MTC	AL	MANUF	Overall Mission Readiness
Pull- ups	0.37	-0.64	-0.73	0.74	-0.73	0.67
Crunches		-0.34	-0.37	0.39	-0.38	0.37
3-mile run			0.79	-0.63	0.73	-0.58
MTC				-0.75	0.82	-0.69
AL					-0.8	0.74
MANUF						-0.74

Moreover, the pull-ups task was the currently added one, conducted in conjunction with the CPTs, and includes both men and women.

---

[7] McGuire, B., Vickers, R. R., Reynolds, J. H., Curry, A., Bockelman, T., & Massimo, R. (2011). Examination of pull-ups and push-ups as possible alternatives to the flexed-arm hang on the Marine Corps Physical Fitness Test.

[8] Hardison, C. M., Sims, C. S., & Wong, E., (2010). The Air Force Officer Qualifying Test: Validity, fairness, bias, Santa Monica, CA: RAND Corporation, TR-744-AF.

Table 9b. Correlations between the PFT and CFT component tasks and overall mission readiness (Spearman's ρ)

Spearman's ρ	Crunches	3-Mile Run	MTC	AL	MANUF	Overall Mission Readiness
Pull-ups	0.46	-0.66	-0.76	0.74	-0.76	0.75
Crunches		-0.37	-0.43	0.43	-0.41	0.38
3-mile run			0.8	-0.63	0.74	-0.61
MTC				-0.75	0.83	-0.71
AL					-0.78	0.76
MANUF						-0.77

Table 10a contains the point-biserial correlations between the individual proxy tests and the PFT and CFT components tasks. The point-biserial correlation[9] is the special case of Pearson's *r* when one of the variables is dichotomous.

---

[9] Following Pett (1997), the following conventions have been used to interpret the size of these point-biserial correlations:

Very strong: ≥.81

Strong: .49-.80 (green)

Moderate: .25-.48 (yellow)

Weak to moderate: .00-.24 (red)

Pett, M.A. (1997). *Nonparametric Statistics for Health Care Research.* Thousand Oaks, CA: Sage Publications, Inc.

Table 10a. Correlations between the PFT and CFT component tasks and the individual proxy tasks (point-biserial correlation)[10]

Point Biserial Correlation, r_{pb}	Deadlift	Clean & Press	Tank Loading Drill	Artillery Round Carry	Wall Negotiation
Pull-ups	*	0.70	0.31	0.44	0.35
Crunches	*	0.32	0.13	0.34	0.20
3-mile run	*	-0.53	-0.3	-0.41	-0.32
MTC	*	-0.64	-0.35	-0.47[11]	-0.34
AL	*	0.69	0.33	0.48	0.32
MANUF	*	-0.68	-0.36	-0.47	-0.36

Nearly all Marines completed the deadlift, so no meaningful correlations can be extracted. Where there was greater variability in testing (meaning that a larger range of values were observed, such as the Clean and Press), the correlations were less susceptible to range restriction, and thus are likely more meaningful in providing a more accurate point estimate of the population correlation coefficient. Table 10b contains results for the FAH. These results are solely from women and thus should not be over-interpreted (i.e., compared directly in terms of magnitude to the results above), since the absolute magnitude of the correlations for FAH are likely underestimates due to the restriction of range).

---

[10] A biserial correlational analysis was also conducted, since one could argue that the test variables were artificially dichotomized. For the sake of brevity, these results were excluded, since they did not change the overall conclusions of the study.

[11] Despite not meeting the conventions suggested by Pett, the MTC and MANUF are classified here as strong, because they have consistently risen to the top as predictors.

Table 10b. Correlations between the FAH and the individual proxy tasks (point-biserial correlation)

Point Biserial Correlation, r_{pb}	Deadlift	Clean & Press	Tank Loading Drill	Artillery Round Carry	Wall Negotiation
FAH	*	0.09	0.11	0.28	0.08

DISCUSSION

The CFT does a better overall job than the PFT in predicting performance on the combat-related tasks. All of the component CFT tasks (AL, MTC, MANUF) do an excellent job of predicting performance on the combat-related tasks. For the component CFT tasks, the AL does the best individual job of predicting performance on the proxy tasks, but the strength of the relationship between the tasks is essentially the same across all the tasks (correlations, $r \approx .7$). For the PFT tasks, pull-ups predict performance better than the run task, which does a better job than crunches. While crunches themselves are the weakest predictor, *relatively speaking*, they nevertheless serve as a valid predictor ($r = .33$) of performance on the overall CPTs score (the percentage of tasks successfully completed), according to criteria used to evaluate the predictive validity of screening procedures (i.e., for a valid predictor, $r = .30–.40$). However, because performance on these tasks was limited to the weights, repetitions, times, and so forth required to do the task rather than measuring maximum physical capacity in executing them, the resulting restricted range may *underestimate* the true predictive validity of these tasks (that is, the true population correlation coefficient, ρ). It is understood that these restrictions were made in order to reflect the occupational demands as reflected by the combat-related tasks; thus it can be inferred that completion of the tasks is reflective of occupational/operational capability not physical ability. Finally, women who were able to perform 1 pull-up performed better than women who could not do a pull-up,

but who achieved a 70-second FAH. That is, for women who could perform 1 pull-up, there was a trend[12] for better performance (compared with women who could do 70 seconds, but who could not do any pull-ups) on the CFT component tasks and the combat-related tasks. Overall, these results suggest that the PFT and CFT serve as a sound basis for making valid inferences about a Marine's physical capability to perform well in combat.

CONCLUSION

NHRC Recommendations to TECOM

Based on the findings, it is recommended that the Marine Corps continue to use the Initial Strength Test (IST) as a prescreen for qualification to become a Marine. However, if a new recruit aims to go into a combat arm MOS, it is recommended that he or she successfully complete an "IST+," which is based on the evidence showing correlation between PFT and CFT components and combat-related tasks. Recommendations for this additional test are as follows.

IST+

Minimum Combat Readiness Standards	% of CPTs Completed	
	Good performers (>75%)	Best performers (100%)
AL	45	70
MTC (m:s)	3:48	3:24
Pull-Ups	3[13]	5

---

[12] Though in nearly all cases (with AL being the exception), the effect was not statistically significant. This is likely a result of insufficient power because of the restricted selection criteria (N ~ 50 women for each group).

[13] The statistical decision procedure that we employed to determine minimum standards generated 1 pull-up as the recommended minimum. This result was due to the composition of the lowest performing decile, which consisted almost entirely of women who could not do at least 1 pull-up. Given the importance of upper-body strength for combat tasks, to recommend 1 pull-up seems to us unreasonable, and we defer to the existing minimum standard of 3 pull-ups.

These recommended numbers to "pass" the IST+ are based on recent data collected on active-duty U.S. Marines students (men and women) who have little combat training and likely no operational experience. For those designated or who want to go into a combat role, these values are the lowest recommended because they are strongly influenced by the scores of female students who are early in their Marine Corps training. These values were calculated by examining the Marines who were most successful on the CPTs. "Success" was defined as passing at least 75% of the proxy tasks, or passing 100% of the tasks. Of the good performers (N = 620), 66% were men and 34% were women. Of the best performers (N = 351), 92% were men and 8% were women.
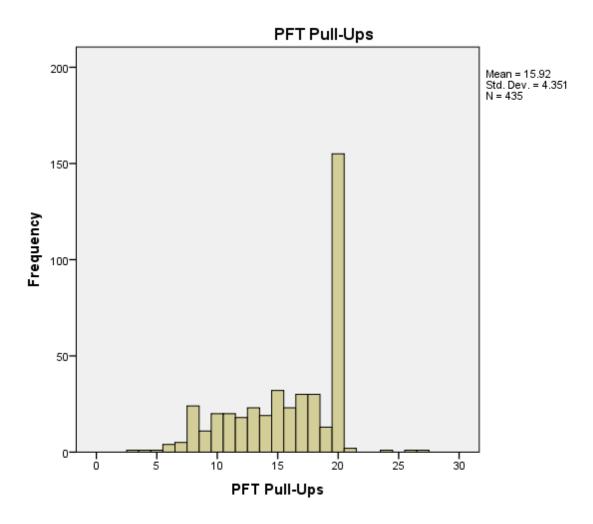
We then examined the worst performers within both groups, as defined by the worst performing decile values on pull-ups, AL, and MTC. We calculated the mean for all Marines whose performance values were worse than the lowest decile cutoff score. Currently, these numbers are higher than those that need to be achieved on the CFT for the AL and MTC. Thus, it can alternatively be recommended that the current minimum standards for the CFT be the passing score for the IST+. However, based on this recommendation, the current minimum passing standards for the CFT may be too low.
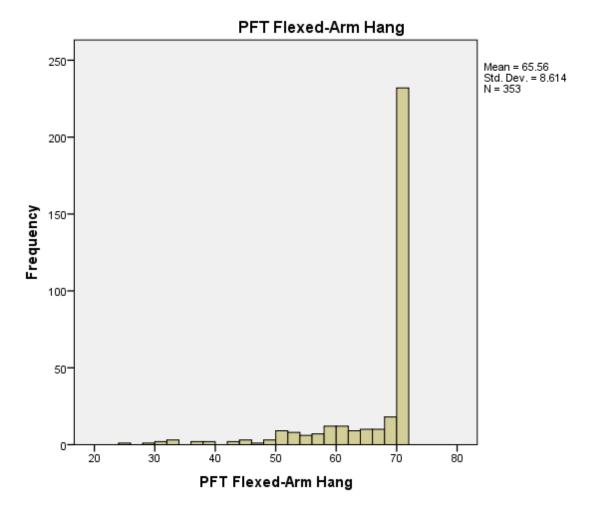
Further, it is recommended that there be a "training to task" period. For example, if a recruit would like to go into a combat arm MOS and fails the IST+, he or she should be given an opportunity to train and re-take the IST+ within a predetermined time period (perhaps no less than 8 weeks) to allow for time to train appropriately and for measurable strength changes to be made. Moreover, for women, greater emphasis should be given on developing upper-body strength, which appears to be the primary factor for women in terms of reduced success on proxy tasks. Previous collaborative NHRC and TECOM research concluded that pull-up training was effective at increasing the number of pull-ups completed by women as well as improving overall upper-body muscle endurance, which translated to other upper-body exercises such as push-ups. A separate Marine Corps study (1993) also showed that a
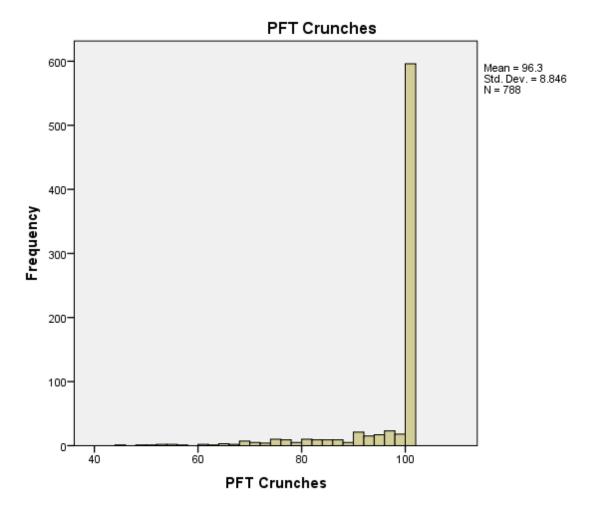
12-week training program increased the number of women who could do more than 3 pull-ups by 30%. In addition, the researchers suggested that pull-ups are a good indicator of upper-body strength and that pull-up training may be the best way to increase female Marine upper-body strength. Collectively these data suggest that while the women in this study were not successful in upper-body proxy tasks, with proper training it is likely that they would be able to achieve the minimum or greater than the minimum standard.

APPENDIX A

The appendices that follow included additional information about the distributions of scores overall, broken down by gender, as well as more detailed information regarding the data collection procedure.

PFT Pull-Ups

Mean = 15.92
Std. Dev. = 4.351
N = 435

PFT Flexed-Arm Hang

Mean = 65.56
Std. Dev. = 8.614
N = 353

PFT Crunches

Mean = 96.3
Std. Dev. = 8.846
N = 788

PFT Crunches

PFT 3-mile Run

Mean = 1372.2
Std. Dev. = 154.754
N = 788

CFT Movement to Contact

Mean = 191.62
Std. Dev. = 26.526
N = 787

CFT Ammunition Lift

Mean = 77.42
Std. Dev. = 23.402
N = 787

CFT Maneuver Under Fire

Mean = 171.67
Std. Dev. = 35.621
N = 787

Distribution of PFT 3-mile run times

Distribution of CFT MTC Times

Distribution of CFT MANUF times

APPENDIX B

(The analysis below was conducted by TECOM). Tables B1–3 depict how each testing site performed, and Table B4 depicts how all Marines performed on the "Analysis of Closed MOS Physical Performance Standards Testing" (CMOS).

Pull-Ups

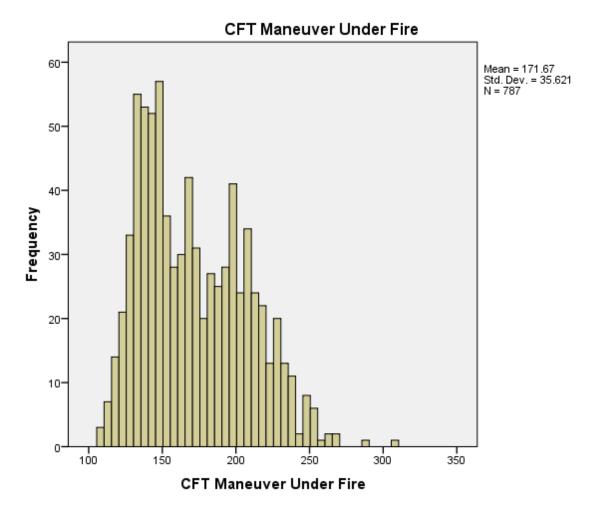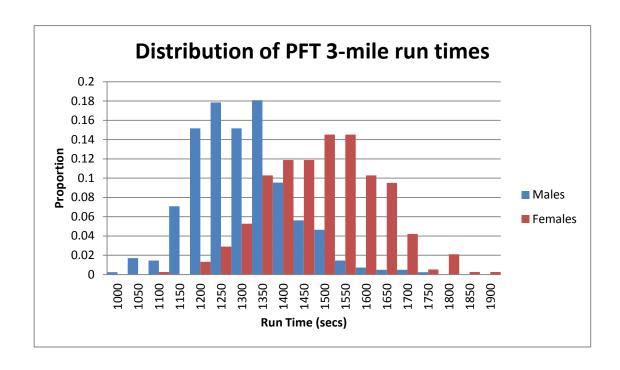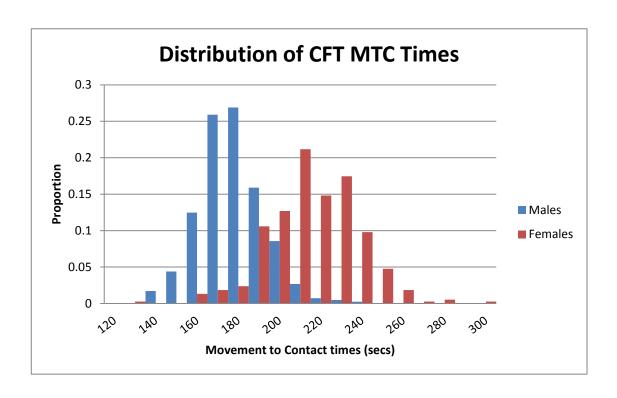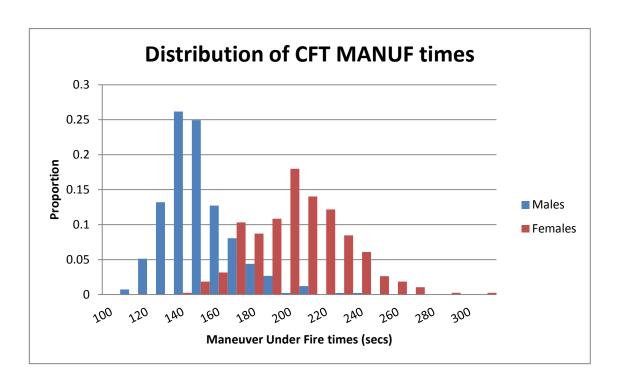Dead-hang pull-ups as an event in the PFT will not become mandatory for the Marine Corps until 2014. Even then PFTs are not required to be performed until the last day of June 2014. Fiscal year 2012 GCE Testing conducted by the TECOM Marine Air Ground Task Force Training and Education Standards Division showed no correlation between the FAH and upper-body strength. To ensure accurate and meaningful correlations between CMOS testing events, PFT and CFT scores, it was necessary to collect pull-up data as a CMOS testing event. For fairness and equality, all participants were required to complete 1 maximum set of pull-ups. Of the participants in the research, men performed an average of 16 pull-ups and women performed and average of 3.6 pull-ups. Pull-ups were performed in PT gear and were the first event. The female average for pull-ups performed was 3.6 and the male average for pull-ups was 15.7.

Deadlift

Deadlifts were performed at incremental weights of 60, 70, 80, 95, 115 and 135 lb. Of the listed weights, Marines progressed to a maximum weight that they could successfully deadlift 1 time. One hundred percent of the male population was able to deadlift 135 lb.

Of the female population, <1% could not deadlift 95 and 115 lb, and 2.6% could not deadlift 135 lb. Initially this is not surprising as the deadlift is primarily a lower-body weight movement utilizing the larger and stronger muscles of the posterior chain.

Clean and Press

The Clean and Press is functional movement and was used to replicate the task of picking an item from the ground and lifting to an overhead position. The Clean and Press was performed at incremental weights of 70, 80, 95, and 115 lb. Additionally, Marines were given 1 minute to clean and press 65 lb 6 times.

Of the listed weights, Marines progressed to a maximum weight that they could successfully clean and press 1 time. All Marines attempted the 6 reps 65 lb, with a time limit of 1 minute. Below is the breakdown of performance by men and women and the associated weights:

26.6% of female population and 1.9% of men could not clean and press 70 lb.

53.95% of female population and 3.17% of men could not clean and press 80 lb.

76.32% of female population and 7.56% of men could not clean and press 95 lb.

91.32% of female population and 20% of men could not clean and press 115 lb.

31.32% of female population and <1% of men could not clean and press 65 lb 6 times with a time limit of 1 minute.

With all weights there is a large disparity between male and female failure weights. What this suggests is that women have less upper-body strength than men. Indeed 91.32% of women failed to lift 115 lb, whereas 20% of men failed at that weight. At the lesser weight of 65 lb, we had our test

population clean and press that weight six times in 1 minute or less. While only 26.6% of women could

not lift 70 lb, when attempting to perform the lesser weight of 65 lb for 6 repetitions 31.32% of them fail

that task.

120 mm Tank Loading Simulation (Gunnery Skills Test)

Less than 1% of men could not complete the tank loading drill in the allotted time. 18.68% of

women could not complete the tank loading drill in the allotted time. Tank loading simulation strictly

measured a Marines ability to lift and transfer a simulated round weighing 55 lb. There were 5 simulated

rounds and the task was to be completed in 35 seconds or less. It would be very likely that failure rates

would increase in a more confined space and actually taking a round out of a horizontal tube and placing

into a horizontal breech.

155 mm Artillery Lift and Carry (Stow Ordnance)

Less than 1% of men could not complete the 155 mm artillery round lift and carry in the allotted

time; 28.42% of women could not complete the 155 mm artillery round lift and carry in the allotted

time.

The 155 mm artillery lift and carry consisted of picking up a replica 155 mm artillery round

weighing 95 lb, and carrying it a distance of 50 m in under 2 minutes. Marines were not required to

place the round on their shoulder and were allowed to cradle the round. While 28.42% of women failed

to complete this task, it is extremely likely that if required to "shoulder" the round and/or carry multiple

rounds, that failure rate would increase.

Obstacle Course Wall with assist box (execute lower-level entry with 1- or 2-person assist)

Less than 1.2% of men could not get over the obstacle course wall using an assist box, while wearing Kevlar helmet, flak jacket, plate carrier, and SAPI; 21.32% of women could not get over the obstacle course wall using an assist box, while wearing Kevlar helmet, flak jacket, plate carrier, and SAPI. This proxy event simulates the infantry task of a lower-level entry utilizing a 1- or 2-person on assist. To standardize the event, a box was used instead of a Marine's hand or leg.

Table B1. Test sample composed of 4 MCT companies.

MCT Performance for CMOS PROXY				
	Women		**Men**	
Total	**238**		**228**	
Event	**# of failures**	**% fail**	**# of failures**	**% fail**
wall	48	20.17%	4	1.75%
artillery carry	80	33.61%	1	0.44%
tank loading	44	18.49%	0	0.00%
Clean & Press				
70 lb	78	32.77%	8	3.51%
80 lb	143	60.08%	13	5.70%
95 lb	191	80.25%	25	10.96%
115 lb	224	94.12%	66	28.95%
6x 65 lb	90	37.82%	2	0.88%
Deadlift				
60 lb	0	0.00%	0	0.00%
70 lb	0	0.00%	0	0.00%
80 lb	0	0.00%	0	0.00%
95 lb	0	0.00%	0	0.00%
115 lb	2	0.84%	0	0.00%
135 lb	2	0.84%	0	0.00%
avg pull-	2.625		14	

ups					

Table B2. Test sample was of enlisted and officer

permanent personnel assigned to PISC[*]

Parris Island South Carolina performance				
	Women		**Men**	
Total	**103**		**127**	
Event	**# of failures**	**% fail**	**# of failures**	**% fail**
wall	33	32.04%	1	0.79%
artillery carry	28	27.18%	0	0.00%
tank loading	26	25.24%	1	0.79%
Clean & Press				
70 lb	22	21.36%	0	0.00%
80 lb	57	55.34%	0	0.00%
95 lb	81	78.64%	6	4.72%
115 lb	95	92.23%	16	12.60%
6x 65 lb	29	28.16%	0	0.00%
Deadlift				
60 lb	0	0.00%	0	0.00%
70 lb	0	0.00%	0	0.00%
80 lb	0	0.00%	0	0.00%
95 lb	1	0.97%	0	0.00%
115 lb	1	0.97%	0	0.00%
135 lb	8	7.77%	0	0.00%
avg pull-ups	5		17.35	

Table B3. Test sample composed of 2 TBS companies*

The Basic School, Quantico VA

	Women		Men	
Total	**39**		**55**	
Event	**# of failures**	**% fail**	**# of failures**	**% fail**
wall	0	0.00%	0	0.00%
artillery carry	0	0.00%	0	0.00%
tank loading	1	2.56%	0	0.00%
Clean & Press				
70 lb	1	2.56%	0	0.00%
80 lb	5	12.82%	0	0.00%
95 lb	18	46.15%	0	0.00%
115 lb	28	71.79%	0	0.00%
6x 65 lb	0	0.00%	0	0.00%
Deadlift				
60 lb	0	0.00%	0	0.00%
70 lb	0	0.00%	0	0.00%
80 lb	0	0.00%	0	0.00%
95 lb	0	0.00%	0	0.00%
115 lb	0	0.00%	0	0.00%
135 lb	0	0.00%	0	0.00%
avg pull-ups	7.6		21.2	

Table B4

TOTAL CMOS PROXY				
	Women		**Men**	
Total	**380**		**410**	
Event	**# of failures**	**% fail**	**# of failures**	**% fail**
wall	81	21.32%	5	1.22%
artillery carry	108	28.42%	1	0.24%
tank	71	18.68%	1	0.24%

loading				
Clean & Press				
70 lb	101	26.58%	8	1.95%
80 lb	205	53.95%	13	3.17%
95 lb	290	76.32%	31	7.56%
115 lb	347	91.32%	82	20.00%
6x 65 lb	119	31.32%	2	0.49%
Deadlift				
60 lb	0	0.00%	0	0.00%
70 lb	0	0.00%	0	0.00%
80 lb	0	0.00%	0	0.00%
95 lb	1	0.26%	0	0.00%
115 lb	3	0.79%	0	0.00%
135 lb	10	2.63%	0	0.00%
avg pull-ups	3.6		15.7	