Examination of Pull-Ups and Push-Ups as Possible Alternatives to the Flexed Arm Hang on the Marine Corps Physical Fitness Test

Brian McGuire
Ross R. Vickers, Jr.
John H. Reynolds
Anne Curry
Timothy Bockelman
Ryan Massimo

Naval Health Research Center

Report No. 11-21

Naval Health Research Center
140 Sylvester Rd.
San Diego, California 92106-3521

Examination of Pull-Ups and Push-Ups as Possible Alternatives to the Flexed Arm Hang on the Marine Corps Physical Fitness Test

Brian McGuire[a], Ross R. Vickers, Jr.[b], John H. Reynolds[c], Anne Curry[d], Timothy Bockelman[d], and Ryan Massimo[e]

[a]U.S. Marine Corps Training and Education Command (G-3), Quantico, VA; [b]Naval Health Research Center, San Diego, CA; [c]Marine Corps Warfighting Lab, Quantico, VA; [d]Marine Corps Recruit Depot Parris Island, SC; [e]Semper Fit, Quantico, VA

Report No. 11-21 was supported by the U.S. Marine Corps Training and Education Command and by the Office of Naval Research, Arlington, VA, under Work Unit No. 60704. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. Approved for public release; distribution is unlimited. Human subjects participated in this study after giving their free and informed consent. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects in research.

Abstract

The Flexed-Arm Hang (FAH) has been an event on the Marine Corps Physical Fitness Test since 1975. This study evaluated alternative tests that would avoid deficiencies in the FAH as a test of dynamic upper body strength and determine the best test of dynamic upper body strength for female Marines within certain parameters (e.g., minimal equipment, training for the test enhances physical performance in dynamic military tasks). The sample consisted of 318 female Marine volunteers from Marine Corps Training and Education Command units. Participants performed four tests: dead-hang pull-ups, movement pull-ups, self-paced push-ups, and cadence push-ups. Participants' most recent Physical Fitness Test (PFT) score and their FAH time were collected from existing records. Scores were similar to those of females Marine Corps-wide. The sample pull-ups scores were positively skewed, FAH times were negatively skewed, and push-up scores were normally distributed. Intertest correlations were consistent with meta-analytic results establishing all three tests as measures of a common muscle endurance factor. Age group differences were minor. Pull-up training improved performance, but push-up training did not. Study participants test performances were comparable to those in other military samples. The FAH was sensitive to differences at the low end of the muscle endurance continuum; pull-ups were sensitive to muscle strength and endurance differences at the high end of the continuum. Pull-ups are a suitable and feasible test of upper body strength and endurance for female Marines. Considering the numbers of Marines who may not be able to complete pull-ups initially upon implementation of such a test, combining the FAH and pull-ups would be a suitable implementation measure. A push-up test is another option, though it is less of a test of strength than the pull-ups. Encouraging pull-ups training would do more to improve muscle strength and endurance than encouraging push-up training.

The Department of Defense requires physical fitness testing for all branches of military service (Assistant Secretary of Defense, 2002). The testing must assess cardiorespiratory fitness and muscle strength and endurance. The flexed-arm hang (FAH) has been part of the U.S. Marine Corps Physical Fitness Test (PFT) for women since 1975. This test has served as a measure of upper body strength and endurance. Concerns have been raised regarding the FAH's effectiveness. Most recently, a recommendation from the 2010 Sergeants Major Symposium was to replace the FAH because it is perceived as an ineffective test of upper body strength. Following this symposium, the Training and Education Command (TECOM) was tasked with determining the most effective and feasible upper body strength test for female Marines.

Available research supports the perception that the FAH is a poor test of upper body strength. The FAH is a somewhat better index of isometric upper body muscle endurance than strength. Muscle strength and muscle endurance are distinct but correlated physical abilities. The FAH is a poor strength indicator but an acceptable isometric muscular endurance measure. When considered as a muscle endurance measure, the FAH is somewhat comparable to the pull-up and push-up when measuring muscle endurance (Table 1). The average factor loading for the FAH is lower than the loadings for the other two tests, but the confidence intervals for those loadings overlap. Allowing for the uncertainty associated with each average loading, all three tests could be considered equivalent measures for men and for women. The essential point is that, in contrast to muscular strength, the FAH, pull-ups, and push-ups are all effective measures of the same construct, upper body endurance.

The meta-analytic results are not directly applicable to the PFT FAH test. The meta-analytic data came from studies in which the FAH was performed to voluntary exhaustion. PFT FAH testing stops after 70 s. Most female Marines do not reach voluntary exhaustion in this time. Consequently, the same maximum score is assigned to women who almost certainly differ in upper body muscle endurance. The range of endurance differences is uncertain, but it may be quite wide.

The truncated FAH scores can be unfair. Consider two groups of women. Women in the first group have exceptional upper body muscle endurance strength, but only mediocre cardiorespiratory endurance. Women in the second group have exceptional cardiorespiratory endurance, but only mediocre upper body muscle strength. In the current PFT, women in the first group would receive 100 points for the FAH, but less than 100 points on the 3-mile run. Women in the second group would receive 100 points on both tests. A fair test arguably would give approximately equal scores to the women in both groups.

Faced with the problematic perceptions and measurement limitations of the FAH as currently administered, an expert panel was convened to identify alternative upper body endurance tests. Pull-up and push-up tests were recommended. These tests have been studied in female Marines in the past, but the previous examinations have been characterized by study design limitations (e.g., limited sample size) and have not examined different test administration methods. This report compares the performance of four FAH test alternatives: dead-hang pull-ups, movement pull-ups, self-paced push-ups, and cadence push-ups, in a moderately large sample.

Method

Subjects

Study participants were 318 female volunteers from the permanent party rosters at eight TECOM Formal Learning Centers and four TECOM Headquarters elements (Table 2). The average participant was 26 years of age (Table 3). Nearly all participants (97%) represented the two youngest age groups used specified in PFT standards (Table 4).

Study participants were more physically fit than an average female Marine (Table 3). The average PFT score was slightly higher than the average female Marine PFT score in 2010. FAH performance was markedly better than average.

Testing Schedule

Volunteers at each test site gave informed consent before participating in up to four test sessions. On the first study day, participants performed the dead-hang pull-up in the morning and the self-paced push-up in the afternoon. A rest day followed the first test day. On the third study day, participants performed the movement pull-up in the morning and the cadence push-up in the afternoon.

Work schedules prevented some study participants from attending the scheduled sessions. Individual or small group test sessions were scheduled to permit those individuals to participate. This decision was based on the judgment that broader sampling was more important than rigid scheduling. Some participants failed to complete all four tests. No injuries were reported during the testing.

Test Procedures

Dead-hang pull-ups. The participant grasped a pull-up bar with the palms of her hands facing away from her body. The test subject then hung from the bar with her arms fully extended and without swaying. Pull-ups were performed from this position by lifting the body until the chin was over the pull-up bar. The participant then lowered herself to the starting position and repeated the pull-up as many times as possible to voluntary exhaustion or until the test administrator instructed her to stop. Pull-ups were not counted if any torso or leg movement assisted in the completion of a pull-up. The participant was allowed to continue when a pull-up was performed improperly, but that pull-up was not counted. Test score was the number of pull-ups successfully completed.

Movement pull-ups. Participants grasped a pull-up bar. The palms of the hands could face away from the body or toward the body. Most participants chose to have their palms face toward their bodies. After coming to a dead hang, the participant lifted her body until her chin was above the pull-up bar. She then returned to the dead-hang position and started another pull-up. The sequence was repeated to voluntary exhaustion or until the test administrator stopped the test. Participants were allowed to move their bodies during the pull-ups, but the movement was limited to swinging slightly on the bar and moderate flexing at the waist and knees. The pull-up was not counted if the participant's movement brought her knees as high as her waist. Participants were allowed to continue when a pull-up was performed improperly, but that pull-up was not counted. Test score was the number of pull-ups successfully completed.

Self-paced push-ups. The test was initiated with the test subject starting in a front leaning rest position. The participant chose a hand position that was comfortable for her. A push-up was performed by lowering the body toward the floor and then pushing back to an up position with arms fully extended and body forming a straight line from head to heel (i.e., a plank position). The test subject maintained the straight body line while lowering her body until her upper arms were parallel to the floor. Still maintaining the straight line, the participant then pushed up until her arms were again fully extended. The down–up cycle was repeated as many times as possible in 2 min. During testing, the participant was allowed to rest in the up position and to make minor adjustments in placement of hands and feet. The subject was not allowed to lift her hands or feet off the ground once the test started. The test administrator counted each properly performed push-up. Push-ups that were not properly performed were not counted, but the test subject was allowed to continue, and any subsequent push-ups that were properly performed were counted. Test score was the number of push-ups successfully performed in the time allowed for the test.

Cadence push-ups. The cadence push-up test differed from the self-paced push-up in four respects. Hand placement was prescribed as hands under the shoulders. Test performance started in the down position. The down position was defined as elbows to 90 degrees. The performance cadence was fixed by a recorded verbal instruction indicating “Up” and “Down.” The alternating instructions were spaced at 1-s intervals for a rate of 30 push-ups per minute. The test continued for 2 min or until the participant was instructed to stop. The test subject was stopped if she failed to maintain the proper push-up form or failed to keep pace with the verbal instructions. The test score was the number of push-ups successfully completed. The test lasted at most 2 min, so the maximum possible score was 60 push-ups.

Appendix A provides the detailed test instructions given to study participants.

Data Analysis

Data analyses were performed with the SPSS-PC computer package, Version 17 (SPSS, Inc., Chicago, IL). Initial descriptive analyses included the Kolmogorov–Smirnov test to determine whether the test score distributions were approximately normal (Siegel, 1956). The distributions were approximately normal for age, PFT, and push-up scores. Analysis of variance (ANOVA) and *t* tests were used subsequently to test for differences in these variables. Marked skew was evident for the FAH and pull-up scores. Nonparametric analyses were used to evaluate group differences for these variables. The Kruskal–Wallis test was the nonparametric ANOVA equivalent. The Mann–Whitney U test was the nonparametric *t* test equivalent. Cohen's (1988) criteria were applied to classify effects as trivial, small, medium, or large.

Results

Descriptive Statistics

Pull-Up Performance

Low scores were typical for pull-ups (Table 5). Roughly 2 in 5 participants (43.2%) performed at least one dead-hang pull-up. About 1 in 5 (21.5%) performed ≥ 3 dead-hang pull-ups, the current minimum standard for male Marines to receive a passing score on this test in the PFT. Allowing movement in the pull-up improved the average pull-up score ~1 pull-up. With this allowance, 55.3% of study participants performed at least one pull-up and 37.4% performed ≥ 3 pull-ups.

Push-Up Performance

Push-up performance differed markedly between the two tests. The average study participant performed roughly twice as many self-paced push-ups as cadence push-ups (Table 5).

Skewness

The Kolmogorov–Smirnov test indicated that FAH and pull-up scores were markedly skewed (Table 5). Figure 1 shows the FAH score distribution and Figure 2 shows pull-up score distributions. The percentiles derived from the figures and from the push-up scores are given in Table 6.

The skewed distributions affected subsequent data analysis decisions. The deviation from a normal distribution means that standard parametric statistical analyses may be misleading for the FAH and pull-ups. Nonparametric statistics were used to analyze these scores.

Fitness Test Associations

Test Score Correlations

Tests that measure the same general construct should be positively correlated. Given previous evidence that all three basic tests measure the same muscle endurance construct, scores on all of the tests should be positively correlated and they were (Table 7). Pull-up–push-up correlations were moderate, ranging from r = .401 to r = .514. Small to moderate FAH correlations ranged from r = .249 to r = .361.

Alternative forms of the same test combined muscle endurance variance general with test-specific variance. The combination should yield stronger correlations than those derived from scores on two different tests. The large correlations obtained with different variants of the same test were consistent with this expectation: pull-ups, r = .892; push-ups, r = .672.

These correlations were consistent with meta-analytic results demonstrating that the FAH, push-ups, and pull-ups measured the same construct. When the factor loadings in Table 1 are combined with sampling variability, the expected correlations for the different types of tests ranged from r = .271 to r = .529. Excluding the correlations between two variants of the same test, the observed correlations ranged from r = .249 to r = .514. The observed range would have narrowed if the FAH had been continued to voluntary exhaustion as it has been in the studies that contributed to the meta-analysis.

Skewness Effects

The skew in the pull-up and FAH test scores effectively represented range restrictions that will attenuate correlations (Sackett & Yang, 2000**).** Spearman's rank-order correlation (ρ), which has been given in parentheses in Table 7, provided a partial correction for this effect. The differences between the rank-order correlations and the corresponding product-moment correlations illustrate the tendency toward lower correlations when data were skewed. The tendency was particularly noteworthy when the negative FAH skew was combined with the positive pull-up skew. The rank-order correlations were between .080 and .104 larger than the product-moment correlations.

The rank-order correlations between the two pull-up tests and between the two push-up tests were slightly smaller than the product moment correlations. These changes suggest that outlier data points affected both correlations. These figures illustrate the importance of allowing for skew when analyzing the FAH and pull-up data.

Age Effects

Age allowances are a standard PFT element. The participants in the present sample were drawn primarily from the two youngest age groups used to define the U.S. Marine Corps PFT (Table 4), so age group comparisons were limited to these two groups (Table 8). Based on those comparisons, PFT standards should require the Marines in the older group to perform 1 more pull-up, 5 more self-paced push-ups, or 2 more cadence push-ups. However, only the push-up requirement would be based on statistically significant group differences.

Training Programs

Training Program Definitions

Potential participants had been provided an optional training plan 6 weeks prior to testing (Appendix B). Before testing, participants were asked four questions about their training patterns. The first question asked whether they had trained for pull-ups. If the answer was “Yes,” the second question asked if they had trained for pull-ups at least twice each week. The third question asked whether they had trained for push-ups. If the answer was “Yes,” the fourth question asked whether they had trained for push-ups at least twice each week.

The initial training program definitions considered pull-ups and push-ups separately. Individuals were characterized as having undergone consistent pull-up training if they answered “Yes” to questions 1 and 2, as having undergone sporadic pull-up training if they answered “Yes” to question 1 and “No” to question 2, or as having undergone no pull-up training if they answered “No” to question 1. By these definitions, 130 (41.4%) Marines participated in consistent pull-up training, 38 (12.1%) Marines participated in sporadic pull-up training, and 146 (46.5%) Marines did not train for pull-ups.

The same training program definitions were applied to push-up training. Based on the answers to the third and fourth training questions, 148 (47.1%) Marines participated in consistent push-up training programs, 29 (9.2%) Marines participated in sporadic push-up training, and 148 (47.1%) Marines did not train for push-ups.

An overall training classification was derived by combining the pull-up and push-up training definitions. Most participants who trained for either test trained for both tests (Table 9). A strong association, $\kappa = .676$, between the two types of training justified the construction of an overall training classification. The two largest groups consisted of 123 (39.2%) Marines who did no training for either test and 111 (35.4%) Marines who trained consistently for both tests. The overall classification characterized the training programs as “No Training,” and “Complete Training,” respectively. The remaining 80 (25.5%) Marines were classified as having undergone “Partial Training.” The incomplete training group included the 18 (5.7%) Marines trained consistently for push-ups but not pull-ups and the 15 Marines (4.8%) who trained consistently for pull-ups but not push-ups (4.8%).

Training Status and Participant Characteristics

Training group status was not related to age or general physical fitness (Table 10).

Overall Training Effects

Omnibus tests of the association of overall training status with performance indicated that training was not related to test performance (Table 11). However, a planned comparison of the No Training and Complete Training groups showed significant differences favoring training for self-paced push-ups, $t = 2.11$, $p < .019$, one-tailed, and movement pull-ups, $z = 1.68$, $p < .048$, one-tailed.

Test-Specific Training Effects

Additional analyses tested the hypothesis that a specific type of training might affect only the targeted tests (Table 12). The omnibus tests for differences as a function of pull-up training status indicated no association of training with Dead-hang pull-up performance, $p = .211$, but a weak association with movement pull-up performance, $p = .059$. Planned comparisons contrasting No Training and Consistent Training indicated that training significantly improved both movement pull-up performance, $z = 2.41$, $p < .008$, one–tailed, and dead-hang pull-up performance, $z = 1.72$, $p < .043$, one-tailed.

Push-up training did not improve push-up performance (Table 12). Although trained participants performed better than untrained participants, the differences did not approach statistical significance, $p > .625$. The planned contrast of Consistent Training with No Training approached statistical significance for the Self-paced push-up, $t = 1.61$, $p < .052$, one-tailed. That contrast was clearly not significant for the Cadence push-up, $t = .74$, $p < .232$, one-tailed.

Generality of Pull-Up Training Effects

Improving scores on a single test of upper body muscle endurance might not mean that a training program improved overall muscle endurance. If the objective is to develop general upper body muscle endurance, training should improve scores on all valid tests of upper body muscle endurance. The relationship of pull-up training with push-up performance was examined to determine whether pull-up training produced the desired general effects.

Pull-up training approached the ideal. Self-paced push-up performance was significantly better, $t = 2.59$, $p < .005$, one-tailed, for the trained participants, $M = 31.12$, $SD = 12.73$, than for the untrained participants, $M = 27.39$, $SD = 10.95$. A similar trend was observed for cadence push-ups: trained, $M = 16.33$, $SD = 7.81$; untrained, $M = 15.40$, $SD = 8.00$. However, that trend was not statistically significant, $t = 0.97$, $p < .156$, one-tailed.

This analysis was not extended to consider the effect of push-up training on pull-up performance. Push-up training could not produce general muscle endurance effects because it did not improve push-up performance (Table 12).

Training Consistency

Pull-up training may be effective even if it is only sporadic. Pull-up performance was slightly better in the Inconsistent Training group than in the Consistent training group (Table 12). Post hoc comparisons indicated that the two groups did either pull-up test, $p > .294$, one-tailed for each test.

Comparison to Reference Populations

Comparisons to results obtained in other test populations places the current results in a broader context for interpretations. The following comparisons allow for the skew and training effects documented in the prior analyses.

General Comparisons

The average study participant's PFT score was slightly better than those of the average female Marine (Table 13). Overall, the study participants had a higher average FAH time and were more likely to reach the maximum FAH score. The general trend toward better performance was evident within age groups as well. However, the differences represented small effect sizes, ES < .17 in all cases. By Cohen's (1988) criteria, none of the differences were large enough to be of theoretical or practical significance.

The study participants' average FAH time was significantly better than that of the typical female Marine in 2002. By Cohen's (1988) criteria, the typical difference was small, but potentially important. The same trend toward better performance was evident for the proportion of tests receiving the maximum FAH score. However, those differences were neither statistically significant nor large enough to be of practical importance.

Different populations had to be considered to evaluate the female Marines' performance on the other tests. With respect to pull-ups, the study participants' dead-hang performance was comparable to that of female West Point cadets, and their movement performance was comparable to that of a 2002 sample of Marines who participated in an experimental pull-up training program.

The comparisons for push-up performance produced mixed results. Self-paced push-up performance of younger Marines was poor relative to their Army counterparts, but older female Marines performance better that their Army counterparts. Marines performed significantly better when the two age groups were combined, but the difference was significant because the two groups were relatively large.

The study participants' relatively poor cadence push-up performance was the only substantial difference between the present sample and a reference group. When expressed as an effect size, the difference was ES = 2.80, a figure more than three times Cohen's (1988) minimum criterion for a large effect, ES = .80. This substantial difference may have little to do with population differences in upper body muscle endurance. Observations during the test sessions suggested that study participants scores were substantially affected by difficulty in matching the required push-up rhythm. Participants were likely to be stopped because they did not maintain the cadence rather than because they were unable to perform another push-up. It seems likely that practice would reduce the difference.

Training Effect Comparison

The generality of pull-up training effects could be evaluated because the participants in a 1993 Marine Corps study (Anonymous, 1993) completed a 12-week supervised physical training program. The pull-up test was administered before training began (Inventory), after 6 weeks of training (Intermediate), and at the end of training (Final). The percentage of women who performed three or more pull-ups was reported for each test administration, so this index could be examined to assess training effects. The 12-week training program increased the number of women who were able to perform ≥ 3 pull-ups by 30% (Table 14). Despite the substantial training improvement, appropriate comparisons consistently favored the Marines in this study.

The untrained individuals in this study provided the proper comparison for the Inventory test. The current study participants' 30.6% pass rate was significantly (p = .031), better than the 15.8% pass rate at the beginning of the 2002 study.

The trained individuals in this study were the proper comparison group for the 1993 Intermediate test results. The participants in this study performed better than those in the earlier study whether the comparison was based on Consistent training, Inconsistent training, or a combination of the two (Table 14). However, the difference was statistically significant only in the two comparisons that included the Consistent training group from this study.

Discussion

All of the tests considered in this study are viable candidates use as muscle endurance measures in the PFT. Vickers (in review) meta-analysis of the structure of physical abilities indicated that all three tests measure the same general muscle endurance construct. The positive intertest correlations were consistent with the evidence supporting that conclusion, particularly after allowing for skew in the test scores. Thus, the current data were consistent with a large body of evidence establishing that the FAH, pull-up, and push-up tests are valid measures of the same muscle endurance construct.

Although the tests were valid upper body muscle endurance indicators, they were not equivalent indicators. The test score distributions indicated that different tests had different sensitivity ranges. The FAH score distribution was skewed toward lower scores with a fixed upper boundary of 70 s. Given this distribution, FAH was sensitive to differences near the lower end of the endurance distribution, but not to differences near the upper end. The pull-up score distributions were skewed toward higher scores, with many women unable to perform a single pull-up. Those distributions made pull-ups sensitive to differences in the upper part of the muscle endurance distribution, but not to differences near the lower end. In contrast with FAH and pull-ups, push-up scores were normally distributed. This suggests that push-ups were sensitive over the full muscle endurance continuum.

The findings present two options for measuring upper body muscle endurance. Adopting a push-up test would cover the full endurance range with one test. At this time, the test score distributions for the PFT FAH and for pull-ups indicate that neither type of test will cover the full muscle endurance range. However, a composite test that combined the FAH with pull-ups could cover the full range. The FAH would be sensitive to individual differences in the lower part of the endurance continuum, and the pull-up would be sensitive to differences in the upper part of the continuum.

Pull-ups or push-ups could be incorporated into the PFT with modest age allowances. Based on this study, adding one pull-up or five push-ups to the standards for those aged 17–26 years would give appropriate standards for 27- to 39-year-olds. Allowances for female Marines in the 40+ age range could not be estimated from the data.

Pull-up training improved test scores, but push-up training did not. The reason for this difference is not known at this time, but it is reasonable to assume that the challenge of lifting one's entire body weight when performing a pull-up is greater than the challenge provided by performing a push-up. If upper body muscle endurance gains are proportional to the training challenge, greater gains with pull-up training would correspond to the greater challenge provided by that training. One implication is that a pull-up test should be preferred to a push-up test if the PFT is modified. PFT modification will be followed by "training to the test," and pull-up training will yield greater benefits than push-up training.

The pull-up test score distributions provide another argument favoring a pull-up test over a push-up test. The pull-up test is the best option for meeting the stated DoD objective of measuring both strength and endurance. A woman must be able to lift her body weight through a full range of motion to complete a pull-up. Strength is the maximum force that a muscle or muscle group can generate. By this definition, a person's strength in the muscle groups involved in the pull-up must be at least equal to his or her body weight or he or she will be unable to perform the exercise. Recent evidence indicates a strong correlation of one-repetition maximum (1-RM) pull-down strength with the number of pull-ups multiplied by body weight (Halet, Mayhew, Murphy, & Fanthorpe, 2009). This interpretation is supported by evidence that 1-RM strength can be estimated from submaximal lifts, particularly when the lifted weight is near the 1-RM so that only a few repetitions are performed (Brechue & Mayhew, 2009; Desgorces, Berthelot, Dietrich, & Testa, 2010; Mayhew, Johnson, Lamonte, Lauber, & Kemmler, 2008; Mayhew, Ware, Cannon, Corbett, Chapman, Bemben…, Slovak, 2002; Reynolds, Gordon, & Robergs, 2006; Whisenant, Panton, East, & Broeder, 2003). If tests that require many repetitions to voluntary exhaustion are less accurate strength indicators than tests that require only a few repetitions to voluntary exhaustion, it follows that pull-ups are preferable to push-ups when assessing strength.

The current training program produced better results than the experimental program carried out in 1993. This finding is important when considering the feasibility of introducing pull-up training. The earlier program generated a strong sense of group cohesion and benefitted from exceptional leadership. These characteristics raised doubts about how well the results would generalize to other settings. Specifically, the report asserted that “Due to these factors, the results are likely to be considerably better than could be achieved by 50 average female Marines under normal working conditions” (Anonymous, 1993). The current training was conducted at diverse sites following with the usual variations in leadership and modifications to the recommended training program. The beneficial training effects seen in this study should allay doubts that the training will improve pull-up performance in the general female Marine population. This finding makes it unlikely that some units would have a marked advantage over others because of exceptional training.

For the most part, study participants’ performance was approximately equal to the performance seen in reference groups. This point obviously is particularly important when generalizing from the present data to the overarching female Marine population. The participants’ average PFT and FAH scores were greater than the corresponding averages for the general female Marine population, but the absolute differences were too small to be important (Cohen, 1988). The general trend also was evident within age groups. Overall, the most important result of these comparisons to other military populations is that the observed differences between the current sample and the reference populations translated into effect sizes that Cohen (1988) would classify as trivial or small.

Cadence push-up performance was the only large difference between the current sample and a reference group. Based on observations during data collection, practice on this test might eliminate most or all of the difference. Participants had difficulty maintaining the required push-up rhythm. This difficulty appeared to be more important than fatigue in determining test scores.

In summary, pull-ups, push-ups, and the FAH are valid measures of upper body muscle endurance for female Marines, with the pull-up being a better measure of muscular strength. However, the pull-up is insensitive to individual differences in the lower part of the endurance distribution, and the FAH, as administered in the PFT, is insensitive to individual differences in the upper part of the endurance distribution. Push-ups appear to be sensitive to differences over the full range of upper body endurance. With these points in mind, two alternatives to the FAH can be suggested to ensure valid measurement over the full range of upper body muscle endurance. One option is to use push-ups as a stand-alone test. The other option is to combine pull-ups and the FAH for a composite test that takes advantage of the tests' differential sensitivity for endurance differences. Both options could be implemented with minimal age allowances. Regardless of which option might be chosen, pull-up training may be the best way to increase general upper body muscle endurance in female Marines. It is also arguable that the pull-up test comes closest to the stated DoD policy objective of measuring both muscle strength and muscle endurance.

References

Anonymous. (1993). Female Physical Fitness Test Study (Decision Paper).

Assistant Secretary of Defense. (2002). DoD Physical Fitness and Body Fat Programs Procedures (DoD Inst. 1308.3 of 5 Nov 02). Washington, DC: Author.

Brechue, W. F., & Mayhew, J. L. (2009). Upper-body work capacity and 1RM prediction are unaltered by increasing muscular strength in college football players. *Journal of Strength and Conditioning Research*, *23*, 2477–2486.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine, 19*, 3127–3131.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Desgorces, F. D., Berthelot, G., Dietrich, G., & Testa, M. S. (2010). Local muscular endurance and prediction of 1 repetition maximum for bench in 4 athletic populations. *Journal of Strength and Conditioning Research*, *24*, 394–400.

Halet, K. A., Mayhew, J. L., Murphy, C., & Fanthorpe, J. (2009). Relationship of 1 repetition maximum lat-pull to pull-up and lat-pull repetitions in elite collegiate women swimmers. *Journal of Strength and Conditioning Research*, *23*, 1496–1502.

Knapik, J., Banderet, L., Bahrke, M., O'Connor, J., Jones, B., & Vogel, J. (1993). *Army Physical Fitness Test (APFT): Normative data on 6022 soldiers* (Tech. Rep. No. T-94-7). Natick, MA: U. S. Army Institute of Environmental Medicine.

Mayhew, J. L., Johnson, B. D., Lamonte, M. J., Lauber, D., & Kemmler, W. (2008). Accuracy of prediction equations for determining one repetition maximum bench press in women before and after resistance training. *Journal of Strength and Conditioning Research*, *22*, 1570–1577.

Mayhew, J. L., Ware, J. S., Cannon, K., Corbett, S., Chapman, P. P., Bemben M. G., Ward, T. E., Farris, B., Juraszek, J., & Slovak, J. P. (2002). Validation of the NFL-225 test for predicting 1-RM bench press performance in college football players. *Journal of Sports Medicine and Physical Fitness*, *42*, 304–308.

Reynolds, J. M., Gordon, T. J., & Robergs, R. A. (2006). Prediction of one repetition maximum strength from multiple repetition maximum testing and anthropometry. *Journal of Strength and Conditioning Research*, *20*, 584–592.

Sackett, P. R., & Yang, H. (2000). Correction for restriction of range: an expanded typology. *Journal of Applied Psychology, 85*, 112–118.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. NY: McGraw-Hill.

Vickers, R. R., Jr. (in review) Construct validity of physical fitness tests (Tech. Rep.). San Diego, CA: Naval Health Research Center.

Whisenant, M. J., Panton, L. B., East, W. B., & Broeder, C. E. (2003). Validation of submaximal prediction equations for the 1 repetition maximum bench press test on a group of collegiate football players. *Journal of Strength and Conditioning Research*, *17*, 221–227.

Appendix A. Detailed Test Instructions

SELF-PACED PUSH-UP INSTRUCTIONS

THE PUSH-UP EVENT MEASURES THE ENDURANCE OF THE CHEST, SHOULDER, AND TRICEPS MUSCLES. ON THE COMMAND “GET SET,” ASSUME THE FRONT-LEANING REST POSITION BY PLACING YOUR HANDS WHERE THEY ARE COMFORTABLE FOR YOU. YOUR FEET MAY BE TOGETHER OR UP TO 12 INCHES APART (MEASURED BETWEEN THE FEET). WHEN VIEWED FROM THE SIDE, YOUR BODY SHOULD FORM A GENERALLY STRAIGHT LINE FROM YOUR SHOULDERS TO YOUR ANKLES. ON THE COMMAND “GO,” BEGIN THE PUSH-UP BY BENDING YOUR ELBOWS AND LOWERING YOUR ENTIRE BODY AS A SINGLE UNIT UNTIL YOUR UPPER ARMS ARE AT LEAST PARALLEL TO THE GROUND. THEN, RETURN TO THE STARTING POSITION BY RAISING YOUR ENTIRE BODY UNTIL YOUR ARMS ARE FULLY EXTENDED. YOUR BODY MUST REMAIN RIGID IN A GENERALLY STRAIGHT LINE AND MOVE AS A UNIT WHILE PERFORMING EACH REPETITION. AT THE END OF EACH REPETITION, THE SCORER WILL STATE THE NUMBER OF REPETITIONS YOU HAVE COMPLETED CORRECTLY. IF YOU FAIL TO KEEP YOUR BODY GENERALLY STRAIGHT, TO LOWER YOUR WHOLE BODY UNTIL YOUR UPPER ARMS ARE AT LEAST PARALLEL TO THE GROUND, OR TO EXTEND YOUR ARMS COMPLETELY, THAT REPETITION WILL NOT COUNT, AND THE SCORER WILL REPEAT THE NUMBER OF THE LAST CORRECTLY PERFORMED REPETITION.”

IF YOU FAIL TO PERFORM THE FIRST 10 PUSH-UPS CORRECTLY, THE SCORER WILL TELL YOU TO GO TO YOUR KNEES AND WILL EXPLAIN YOUR

DEFICIENCIES. YOU WILL THEN BE SENT TO THE END OF THE LINE TO BE RETESTED. AFTER THE FIRST 10 PUSH-UPS HAVE BEEN PERFORMED AND COUNTED, NO RESTARTS ARE ALLOWED. THE TEST WILL CONTINUE, AND ANY INCORRECTLY PERFORMED PUSH-UPS WILL NOT BE COUNTED. AN ALTERED, FRONT-LEANING REST POSITION IS THE ONLY AUTHORIZED REST POSITION. THAT IS, YOU MAY SAG IN THE MIDDLE OR FLEX YOUR BACK. WHEN FLEXING YOUR BACK, YOU MAY BEND YOUR KNEES, BUT NOT TO SUCH AN EXTENT THAT YOU ARE SUPPORTING MOST OF YOUR BODY WEIGHT WITH YOUR LEGS. IF THIS OCCURS, YOUR PERFORMANCE WILL BE TERMINATED. YOU MUST RETURN TO, AND PAUSE IN, THE CORRECT STARTING POSITION BEFORE CONTINUING. IF YOU REST ON THE GROUND OR RAISE EITHER HAND OR FOOT FROM THE GROUND, YOUR PERFORMANCE WILL BE TERMINATED. YOU MAY REPOSITION YOUR HANDS AND/OR FEET DURING THE EVENT AS LONG AS THEY REMAIN IN CONTACT WITH THE GROUND AT ALL TIMES. CORRECT PERFORMANCE IS IMPORTANT. YOU WILL HAVE 2 MINUTES IN WHICH TO DO AS MANY PUSH-UPS AS YOU CAN. WATCH THIS DEMONSTRATION.

CANDIDATE FITNESS ASSESSMENT PULL-UP INSTRUCTIONS

PULL-UPS MEASURE MUSCULAR STRENGTH AND ENDURANCE OF THE SHOULDER AND BACK. WHEN INSTRUCTED, MOUNT THE BAR WITH THE PALMS OF YOUR HANDS FACING AWAY FROM YOU WITH YOUR ARMS FULLY EXTENDED IN A DEAD-HANG POSITION. AT THE COMMAND ‘BEGIN,’ RAISE YOUR BODY UNTIL YOUR CHIN IS RAISED ABOVE THE BAR. YOU MAY NOT SWING, KICK, OR

BICYCLE YOUR LEGS DURING UPWARD MOVEMENT. RETURN TO THE DEAD-HANG POSITION. EXECUTE AS MANY REPETITIONS AS YOU CAN.

MOVEMENT PULL-UP INSTRUCTIONS

PULL-UPS MEASURE MUSCULAR STRENGTH AND ENDURANCE OF THE SHOULDER AND BACK. WHEN INSTRUCTED, MOUNT THE BAR WITH ANY GRIP YOU CHOOSE, PALMS FACING AWAY FROM YOU OR PALMS TOWARD YOU WITH YOUR ARMS EXTENDED IN A DEAD-HANG POSITION. AT THE COMMAND 'BEGIN,' RAISE YOUR BODY UNTIL YOUR CHIN IS RAISED ABOVE THE BAR. YOU MAY SWING OR KICK YOUR LEGS AS LONG AS YOUR KNEES ARE NOT RAISED ABOVE YOUR WAIST. YOUR FEET MAY NOT TOUCH THE PULL-UP BAR SUPPORTS. RETURN TO THE DEAD-HANG POSITION. EXECUTE AS MANY REPETITIONS AS YOU CAN.

CADENCE PUSH-UP INSTRUCTIONS

THE CADENCE PUSH-UP EVENT MEASURES THE STRENGTH AND ENDURANCE OF THE CHEST, SHOULDER, AND TRICEPS MUSCLES. LIE PRONE, READY TO PERFORM A FULL PUSH-UP. HANDS SHOULD BE SHOULDER-WIDTH APART AND JUST UNDER YOUR SHOULDERS. FINGERS SHOULD BE FACING FORWARD. ELBOWS ARE BENT. THE TAPE WILL COUNT DOWN FROM 5 TO 1. THE NEXT COMMAND WILL BE "UP." THE UP POSITION IS ELBOWS LOCKED, BODY STRAIGHT (BUTTOCKS IN LINE WITH BACK AND LEGS). HIPS MAY NOT BE FLEXED. THE HEAD AND NECK SHOULD BE IN LINE WITH THE BACK. THIS WILL

BE FOLLOWED BY "DOWN." THE DOWN POSITION IS BACK STRAIGHT WITH ELBOWS BENT TO AT LEAST 90 DEGREES. CONTINUE FOR AS LONG AS POSSIBLE, STAYING WITH THE CADENCE. NO RESTING IS PERMITTED, AND HAND POSITION CANNOT BE CHANGED. THE TEST IS FINISHED WHEN PUSH-UPS ARE NOT PROPERLY EXECUTED OR DO NOT STAY ON THE CADENCE. THE MAXIMUM NUMBER OF PUSH-UPS IS 60 OVER 2 MINUTES.

Appendix B. Training Program

General Instructions

Initial load for weighted exercises should be that which enables completion of the set with momentary muscular fatigue in the last repetition. Load should be progressively increased with time. Rest 1-2 min between sets.

The program is to be monitored by a Command Physical Training Representative, Certified Athletic Trainer, or Semper Fit Personal Trainer.

Exercise Schedule for Week 1

Day 1			
Exercise/Movement	Load	Reps	Sets
Pull-Ups (Baseline Test)	BW	Max	1
Bent-Over Row (Pronated)		12-15	3-4
3-Way Shoulder		12-15	3-4
Dip	BW	12-15	3-4
Biceps Curl (Bar)		12-15	3-4

Day 2			
Exercise/Movement	Load	Reps	Sets
Push-Ups (Baseline Test)	BW	Max	1
Low Row		12-15	3-4
Military Press (Bar)		12-15	3-4
Pull-Over		12-15	3-4
Reverse Biceps Curl (Bar)		12-15	3-4

Day 3			
Exercise/Movement	Load	Reps	Sets
Jumping Pull-Ups	BW	12-15	3-4
Lat Pull-Down (Supinated)		12-15	3-4
Upright Row (DB)		12-15	3-4
Close-Grip Bench Press		12-15	3-4
Biceps Curl (DB)	BW	12-15	3-4

Exercise Schedule for Week 2

Day 1			
Exercise/Movement	Load	Reps	Sets
Weighted Push-Ups		12-15	3-4
Single-Arm Row		12-15	3-4
Front Plate Raise w/ Hold		12-15	3-4
Triceps Push-Down (Pronated)		12-15	3-4
Hammer Curl (DB)		12-15	3-4

Day 2			
Exercise/Movement	Load	Reps	Sets
Supinated Pull-Ups	BW	12-15	3-4
Bent-Over Row (Supinated)		12-15	3-4
3-Way Shoulder		12-15	3-4
Reverse Dip	BW	12-15	3-4
Biceps Curl (Bar)		12-15	3-4

Day 3			
Exercise/Movement	Load	Reps	Sets
Close-Grip Push-Ups	BW	12-15	3-4
Low Row		12-15	3-4
Military Press (DB)		12-15	3-4
Pull-Over	BW	12-15	3-4
Reverse Biceps Curl (Bar)		12-15	3-4

Exercise Schedule for Week 3

Day 1			
Exercise/Movement	Load	Reps	Sets
Pronated Pull-Ups	BW	12-15	3-4
Lat Pull-Down (Pronated)		12-15	3-4
Upright Row (Bar)		12-15	3-4
Close-Grip Bench Press		12-15	3-4
Biceps Curl (DB)	BW	12-15	3-4

Day 2			
Exercise/Movement	Load	Reps	Sets
Push-Ups	BW	12-15	3-4
Single-Arm Row		12-15	3-4
Front Plate Raise w/ Hold		12-15	3-4
Triceps Push-Down (Supinated)		12-15	3-4
Hammer Curl (DB)		12-15	3-4

Day 3			
Exercise/Movement	Load	Reps	Sets
Jumping Pull-Ups	BW	12-15	3-4
Bent-Over Row (Pronated)		12-15	3-4
3-Way Shoulder		12-15	3-4
Dip	BW	12-15	3-4
Biceps Curl (Bar)		12-15	3-4

Exercise Schedule for Week 4

Day 1			
Exercise/Movement	Load	Reps	Sets
Weighted Push-Ups		12-15	3-4
Low Row		12-15	3-4
Military Press (Bar)		12-15	3-4
Pull-Over		12-15	3-4
Reverse Biceps Curl (Bar)		12-15	3-4

Day 2			
Exercise/Movement	Load	Reps	Sets
Supinated Pull-Ups	BW	12-15	3-4
Lat Pull-Down (Supinated)		12-15	3-4
Upright Row (DB)		12-15	3-4
Close-Grip Bench Press		12-15	3-4
Biceps Curl (DB)	BW	12-15	3-4

Day 3			
Exercise/Movement	Load	Reps	Sets
Close-Grip Push-Ups	BW	12-15	3-4
Single-Arm Row		12-15	3-4
Front Plate Raise w/ Hold		12-15	3-4
Triceps Push-Down (Pronated)		12-15	3-4
Hammer Curl (DB)		12-15	3-4

Exercise Schedule for Week 5

Day 1			
Exercise/Movement	Load	Reps	Sets
Pronated Pull-Ups	BW	10-12	4-5
Bent-Over Row (Supinated)		10-12	4-5
3-Way Shoulder		10-12	4-5
Reverse Dip	BW	10-12	4-5
Biceps Curl (Bar)		10-12	4-5

Day 2			
Exercise/Movement	Load	Reps	Sets
Push-Ups	BW	10-12	4-5
Low Row		10-12	4-5
Military Press (DB)		10-12	4-5
Pull-Over		10-12	4-5
Reverse Biceps Curl (Bar)		10-12	4-5

Day 3			
Exercise/Movement	Load	Reps	Sets
Jumping Pull-Ups	BW	10-12	4-5
Lat Pull-Down (Pronated)		10-12	4-5
Upright Row (Bar)		10-12	4-5
Close-Grip Bench Press		10-12	4-5
Biceps Curl (DB)	BW	10-12	4-5

Exercise Schedule for Week 6

Day 1			
Exercise/Movement	Load	Reps	Sets
Weighted Push-Ups		10-12	4-5
Single-Arm Row		10-12	4-5
Front Plate Raise w/ Hold		10-12	4-5
Triceps Push-Down(Supinated)		10-12	4-5
Hammer Curl (DB)		10-12	4-5

Day 2			
Exercise/Movement	Load	Reps	Sets
Supinated Pull-Ups	BW	10-12	4-5
Bent-Over Row (Pronated)		10-12	4-5
3-Way Shoulder		10-12	4-5
Dip	BW	10-12	4-5
Biceps Curl (Bar)		10-12	4-5

Day 3			
Exercise/Movement	Load	Reps	Sets
Close-Grip Push-Ups	BW	10-12	4-5
Low Row		10-12	4-5
Military Press (Bar)		10-12	4-5
Pull-Over		10-12	4-5
Reverse Biceps Curl (Bar)		10-12	4-5

Table 1

Gender Differences in Test Validity

	Men		Women				
Indicator	*M*	*SE*	*M*	*SE*	Difference	*z*	Sig
Pull-up	.760	.032	.578	.040	.182***	3.51	.000
Push-up	.690	.032	.691	.063	-.001	-.01	.495
FAH	.778	.054	.641	.071	.137	1.53	.062

*$p < .05$, two-tailed; **$p < .01$; ***$p < .001$.
FAH, flexed-arm hang.

Table 2

Female Marine Upper Body Physical Fitness Testing Sites

Location	FLC/HQ Element	Subjects (N = 318)
NAS Pensacola	Marine Air Training Support Group 21 (MATSG21)	26
MCRD Parris Island	Marine Corps Recruit Depot (MCRDPI)	101
MCRD San Diego	Marine Corps Recruit Depot (MCRDSD)	49
MCB Twentynine Palms	Communications Electronic School	42
	Marine Air Ground Task Force (MAGTF) Training Command	23
MCB Quantico	Training Command Headquarters	7
	Training and Education Command (TECOM) Headquarters	2
	Officer Candidate School (OCS)	10
	The Basic School (TBS)	14
MCB Camp Lejeune	School of Infantry East (SOIE)	20
	Marine Corps Combat Service Support Schools (MCCSSS)	21
	Engineer School	3

Table 3

Descriptive Statistics for Sample Characteristics

	Valid	Missing	*M*	*SD*	Min	Max	K-S Z[a]	Sig
Age, y	302	22	26.37	5.78	18	45	2.02	.001
FAH	298	26	64.38	9.63	23	70	6.06	.000
PFT	303	21	253.82	33.38	140	300	1.85	.002

[a]K-S Z = z score from the Kolmogorov–Smirnov test for a normal distribution.

Table 4

Age Group Distribution

Age group, y	*n*	%	Valid %
17–26	173	53.4	57.3
27–39	121	37.3	40.1
40–45	8	2.5	2.6
Total	302	93.2	
Missing	22	6.8	

Table 5

Descriptive Statistics for Pull-Ups and Push-Ups

	Valid	Missing	*M*	*SD*	Min	Max	K-S Z[a]	Sig
Pull-ups								
Dead-hang	317	7	1.63	2.92	0	18	5.14	.000
Movement	313	11	2.59	3.50	0	20	4.06	.000
Push-ups								
Self-paced	313	11	29.10	12.42	2	63	.88	.428
Cadence	310	14	15.70	8.06	0	40	.64	.802

[a]K-S Z = *z* score for the Kolmogorov–Smirnov test for a normal distribution.

Table 6

Distribution Percentiles for Test Scores

	Percentile											
	10th	20th	25th	30th	40th	50th	60th	70th	75th	80th	90th	95th
Sample												
FAH	49	57	61	65	70	70	70	70	70	70	70	70
PFT	204	228	234	241	250	262	270	277	280	284	290	296
Pull-ups												
Dead-hang	0	0	0	0	0	0	1	2	2	3	5	8
Movement	0	0	0	0	0	1	2	3	4	5	8	10
Push-ups												
Self-paced	13	18	20	21	25	30	32	35	38	40	46	51
Cadence	5	9	10	12	14	16	18	20	21	22	26	29

Table 7

Correlations Among Upper Body Muscular Endurance Tests

	FAH	Dead-hang	Movement	Self-paced	Cadence
FAH	1.000				
Pull-ups					
Dead-hang	.249 (.341)	1.000			
Movement	.334 (.438)	.892 (.845)	1.000		
Push-ups					
Self-paced	.361 (.383)	.434 (.496)	.514 (.572)	1.000	
Cadence	.345 (.375)	.401 (.489)	.445 (.494)	.672 (.652)	1.000

Note. All correlations were significant at the 0.01 level (one-tailed).

Table 8

Age Group Comparisons

	17–26 Years			27–39 Years				
	N	*M*	*SD*	*N*	*M*	*SD*	Δ Test	Sig
Dead-hang	172	1.29	2.32	121	1.99	3.45	1.50[a]	.135
Movement	172	2.24	3.04	117	3.04	4.07	1.42[a]	.155
Self-paced	171	26.97	12.39	119	31.88	11.86	3.38[b]	.001
Cadence	171	14.82	7.71	115	16.64	8.66	1.86[b]	.064
FAH	162	62.99	10.49	118	66.15	7.84	2.75[a]	.006

[a]Mann–Whitney U z score. [b]Student's t-test.

Table 9

Training Patterns

	Training status	Pull-up specific training program: None	Inconsistent	Consistent	Total
Push-up specific training program	None	123[a]	10	15	148
	Inconsistent	5	20	4	29
	Consistent	18	8	111	137
	Total	146	38	130[b]	314

Note. $\chi^2 = 258.43$, 4 *df*, $p = .000$; $\kappa = .676$, $t = 15.08$, $p = .000$.
[a]No Training group. [b]Complete Training group.

Table 10

Association of Participant Characteristics With Overall Training

	No Training[a]			Partial Training[a]			Complete Training[a]				
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*F*	Sig
Age	117	26.65	5.69	75	25.71	5.70	106	26.47	6.01	.64	.528
FAH	119	63.69	10.99	77	65.38	7.88	98	64.20	9.27	.72	.489[b]
PFT	119	253.06	36.49	78	256.17	30.44	102	251.84	31.79	.38	.682

[a]See text for training classification definitions. [b]Kruskal–Wallis test for group differences, $\chi^2 = 0.18$, 2 *df*, $p = .910$.

Table 11

Comparison of Overall Training Programs

	No Training[a]			Partial Training[a]			Complete Training[a]				
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	Test	Sig
Pull-ups											
Dead-hang	116	1.51	2.93	75	1.67	2.63	109	1.48	2.69	.80[b]	.670
Movement	116	2.30	3.48	73	2.67	3.59	109	2.72	3.46	2.36[b]	.308
Push-ups											
Self-paced	115	27.43	11.48	75	29.13	12.90	109	31.02	13.35	.47[c]	.625
Cadence	115	15.17	8.25	73	16.63	7.91	109	15.97	8.06	.75[c]	.473

[a]See text for training classification definitions. [b]Test statistic is the Kruskal–Wallis χ^2 test with 2 *df*. [c]Test statistic is a one-way ANOVA *F* test with 2 *df*.

Table 12

Training Effects Estimated for Test-Specific Training Programs

	No Training[a]			Inconsistent Training[a]			Consistent Training[a]				
	N	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	Test	Sig
Pull-up Training											
Dead-hang	146	1.47	2.89	38	1.84	2.99	130	1.64	2.73	3.12[b]	.211
Movement	144	2.17	3.34	36	2.94	3.90	130	2.89	3.48	5.65[b]	.059
Push-up Training											
Self-paced	146	28.01	11.49	28	28.96	15.31	136	30.38	12.82	.47[c]	.625
Cadence	144	15.39	7.76	27	16.37	10.55	136	16.07	7.83	.33[c]	.781

[a]See text for training classification definitions. [b]Test statistic is the Kruskal–Wallis χ^2 test with 2 *df*. [c]Test statistic is a one-way analysis of variance *F* test with 2 *df*.

Table 13

Comparison to Reference Populations

	Current Sample M	Current Sample SD	Reference Population M	Reference Population SD	Test	Sig	ES[a]
PFT							
USMC	253.8	33.4	252.6	33.6	t = .65	.516	.04
17–26	248.8	35.7	251.4	33.7	t = -.92	.361	-.07
27–39	261.7	28.1	257.7	32.4	t = 1.57	.119	.14
40–45	246.8	42.2	244.9	35.5	t = .12	.907	.05
FAH Time							
2010							
USMC	64.4	9.6	63.5	10.9	t = 1.67	.097	.11
17–26	63.0	10.5	63.0	11.2	t = -.03	.975	.00
27–39	66.2	7.8	64.9	9.5	t = 1.71	.090	.16
40–45	61.8	14.7	63.9	10.4	t = -.41	.692	-.14
2002							
USMC	64.4	9.6	61.5	13.1	t = 5.11	.000	.30
17–26	63.0	10.5	61.2	13.2	t = 2.18	.031	.17
27–39	66.2	7.8	62.7	12.5	t = 4.78	.000	.45
40–45	61.8	14.7	59.2	14.5	t = .49	.640	.18
FAH % Maximum[b]							
2010							
USMC	60.4%		60.0%		χ^2 = 0.00	.994	.01
17–26	53.5%		58.3%		χ^2 = 1.55	.215	-.11
27–39	70.0%		65.7%		χ^2 = 0.95	.331	.11
40–45	62.5%		62.2%		χ^2 = 0.00	.994	.01
2002							
USMC	60.4%		55.3%		χ^2 = 2.72	.099	.12
17–26	53.5%		53.5%		χ^2 = 0.00	.993	.00
27–39	70.0%		60.8%		χ^2 = 3.50	.061	.22
40–45	62.5%		47.8%		χ^2 = 0.62	.463	.33
Pull-ups							
Dead-hang[c]	1.24[g]	1.78[g]	1.08	1.60	z = .26	.799	.09
Movement[d]	2.59	3.50	3.66	3.91	z = -1.10	.268	-.31
Push-ups							
Self-paced[e]							
17–26	27.0	12.4	30.7	11.0	t = -3.94	.000	-.30
27–41	31.9	11.9	23.7	10.7	t = 7.53	.000	.69
Combined	29.1	12.4	27.3	11.4	t = 2.57	.011	.15
Cadence[f]	15.7	8.1	37.8	7.6	t = 48.31	.000	-2.48

Note. The t test was used to compare groups for push-ups because those scores were approximately normally distributed. The Mann–Whitney U test was used to compare groups on pull-ups because those scores were highly skewed.

[a]Effect size (ES) computed as (Sample Mean – Reference Mean)/Sample Standard Deviation or as ES = ln(Odds Ratio)/1.81 for proportions (Chinn, 2000). [b]Percentage of females receiving the maximum possible FAH score. [c]Reference group is female West Point Cadets for 2011. [d]Reference group is 1993 female Marine pull-up study sample. Comparison includes only calendar year 2011 Marines who trained consistently. [e]Reference group is

Knapik, Banderet, Bahrke, O'Connor, Jones, and Vogel, 1993. [f]Reference group is female Coast Guardsmen for 2011. [h]Marine Corps data rescored from 0 to 5 to match West Point data.
FAH, flexed-arm hang; PFT, Physical Fitness Test; USMC, U.S. Marine Corps.

Table 14

Comparison to 1993 Marine Corps Study Sample

		P[a]	F	%	P[a]	F	%	χ^2	Sig	OR	ES
Inventory	N	9	48	15.8	44	100	30.6	4.59	.031	2.35	.47
Intermediate	I	12	34	26.1	15	21	41.7	2.22	.148	2.02	.39
Intermediate	C	12	34	26.1	56	74	43.1	4.14	.042	2.14	.42
Intermediate	I+C	12	34	26.1	71	95	42.8	4.21	.040	2.12	.42

Note. Inventory and Intermediate refer to the measurement times for the calendar year 1993 sample. The N, I, and C designations identify the CY11 comparison groups: N = no training; I = Inconsistent training; C = Consistent training. P = Pass. F = Fail. OR = Odds ratio for passing. ES = Effect size = ln(OR)/1.81 (Chinn, 2000).

Figure 1

(a) Dead-Hang Pull-Up Score Distributions

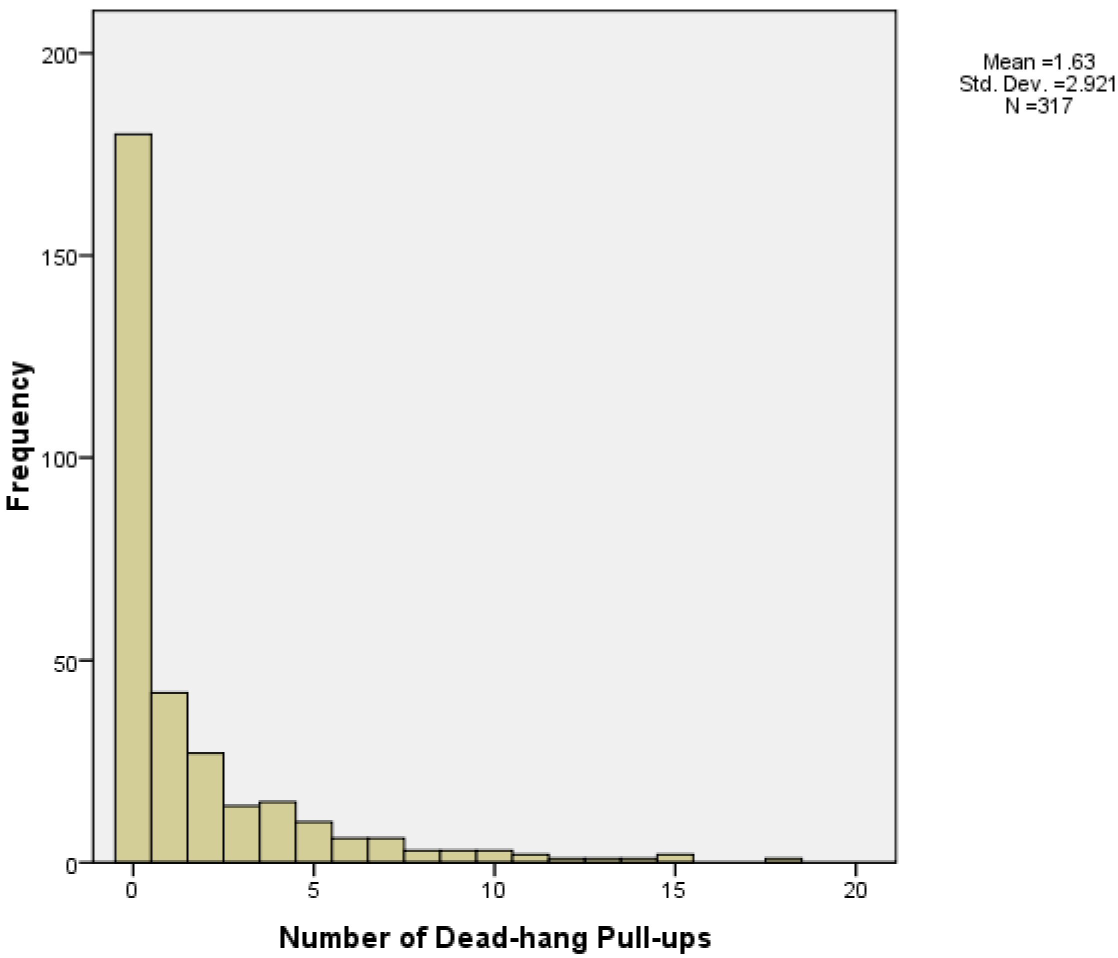


Figure 1 (continued)

(b) Movement Pull-Up Score Distributions

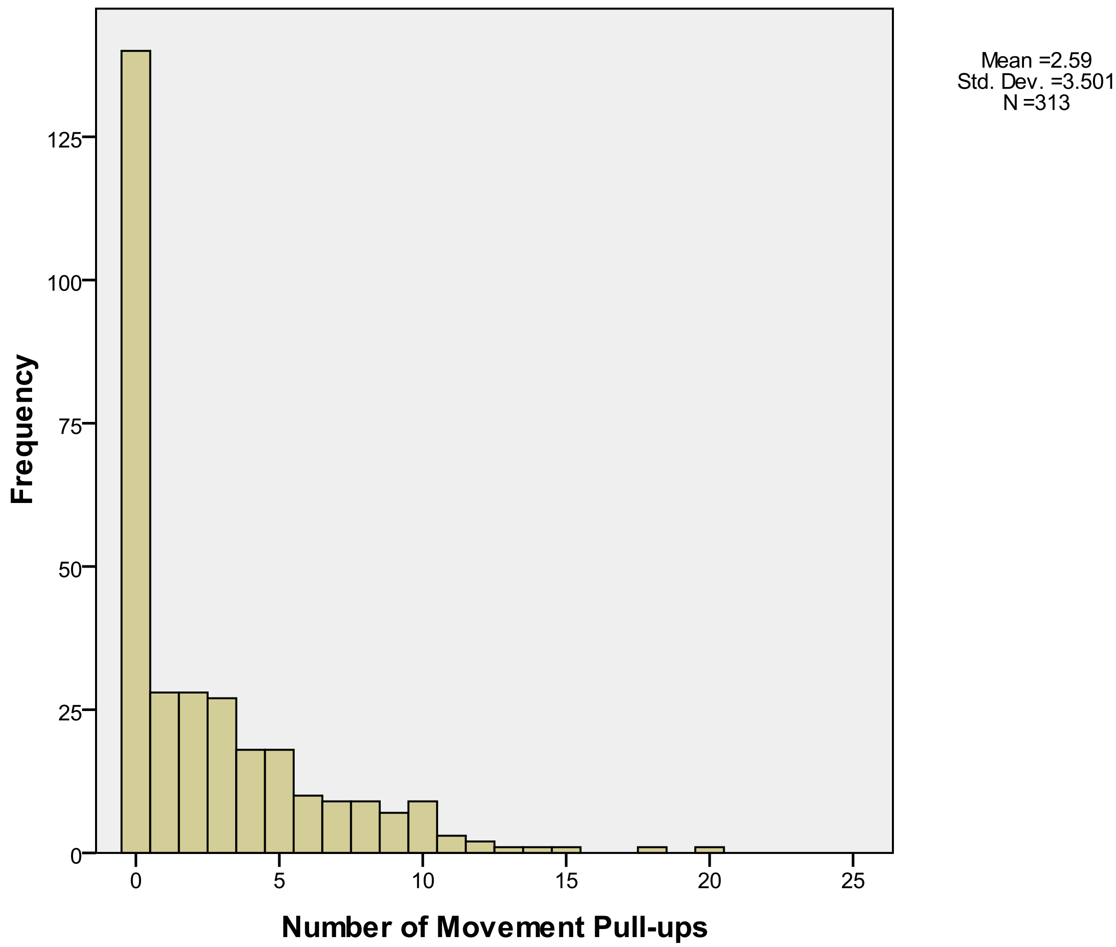
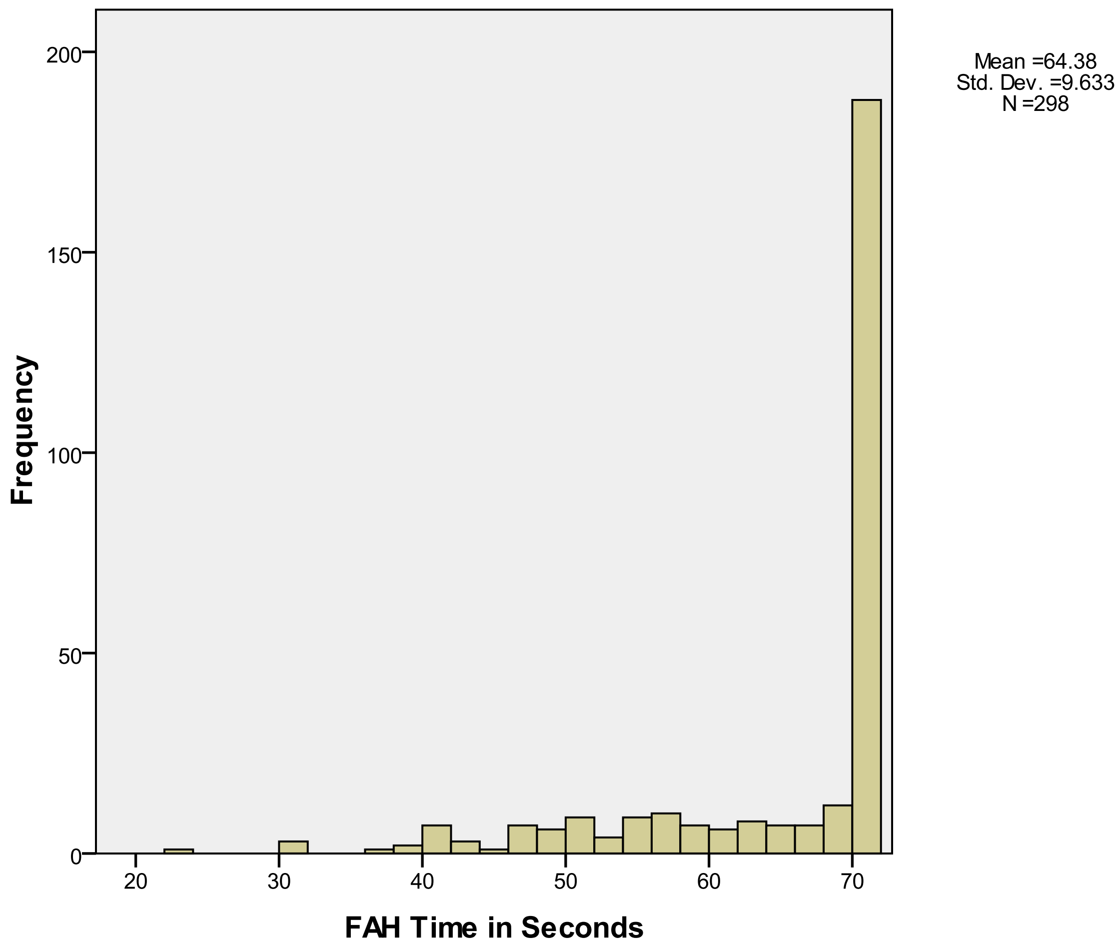


Figure 2

FAH Score Distribution



REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD MM YY)
15 06 11

2. REPORT TYPE
Technical

3. DATES COVERED (from – to)
Jan 2011 – Jun 2011

4. TITLE
Examination of Pull-ups and Push-ups as Possible Alternatives to the Flexed Arm Hang on the Marine Corps Physical Fitness Test

5a. Contract Number:
5b. Grant Number:
5c. Program Element Number:
5d. Project Number:
5e. Task Number:
5f. Work Unit Number: 60704

6. AUTHORS
McGuire, Brian; Vickers, Ross; Reynolds, John; Curry, Anne; Bockelman, Timothy; Massimo, Ryan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commanding Officer
Naval Health Research Center
140 Sylvester Rd
San Diego, CA 92106-3521

8. PERFORMING ORGANIZATION REPORT NUMBER
Report No. 11-21

8. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
Commanding Officer
Naval Medical Research Center
503 Robert Grant Ave
Silver Spring, MD 20910-7500

Chief, Bureau of Medicine and Surgery
(MED 00), Navy Dept
2300 E Street NW
Washington, DC 20372-5300

10. SPONSOR/MONITOR'S ACRONYM(S)
NMRC/BUMED

11. SPONSOR/MONITOR'S REPORT NUMBER(s)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Flexed-Arm Hang (FAH) has been an event on the Marine Corps Physical Fitness Test since 1975. This study evaluated alternative tests that would avoid deficiencies in the FAH as a test of dynamic upper body strength and determine the best test of dynamic upper body strength for female Marines within certain parameters (e.g., minimal equipment, training for the test enhances physical performance in dynamic military tasks). The sample consisted of 318 female Marine volunteers from Marine Corps Training and Education Command units. Participants performed four tests: dead-hang pull-ups, movement pull-ups, self-paced push-ups, and cadence push-ups. Participants' most recent Physical Fitness Test (PFT) score and their FAH time were collected from existing records. Scores were similar to those of females Marine Corps-wide. The sample pull-ups scores were positively skewed, FAH times were negatively skewed, and push-up scores were normally distributed. Intertest correlations were consistent with meta-analytic results establishing all three tests as measures of a common muscle endurance factor. Age group differences were minor. Pull-up training improved performance, but push-up training did not. Study participants test performances were comparable to those in other military samples. The FAH was sensitive to differences at the low end of the muscle endurance continuum; pull-ups were sensitive to muscle strength and endurance differences at the high end of the continuum. Pull-ups are a suitable and feasible test of upper body strength and endurance for female Marines. Considering the numbers of Marines who may not be able to complete pull-ups initially upon implementation of such a test, combining the FAH and pull-ups would be a suitable implementation measure. A push-up test is another option, though it is less of a test of strength than the pull-ups. Encouraging pull-ups training would do more to improve muscle strength and endurance than encouraging push-up training.

15. SUBJECT TERMS
X Insert key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	18a. NAME OF RESPONSIBLE PERSON
UNCL	UNCL	UNCL	UNCL	50	Commanding Officer

18b. TELEPHONE NUMBER (INCLUDING AREA CODE)
COMM/DSN: (619) 553-8429

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18