Effects of Physical Training in Military Populations: A Meta-Analytic Summary

Ross R. Vickers, Jr.
Amanda C. Barnard

Naval Health Research Center

Report No. 11-17

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, nor the U.S. Government. Approved for public release; distribution is unlimited.

Naval Health Research Center
140 Sylvester Rd.
San Diego, California 92106-3521
Effects of Physical Training in Military Populations:

A Meta-Analytic Summary

Ross R. Vickers, Jr.
Amanda C. Barnard

Naval Health Research Center
140 Sylvester Road
San Diego, CA 92106-3521

e-mail: ross.vickers@med.navy.mil

This represents report 11-17 and was supported by the U.S. Department of Defense under work unit 60704. The views expressed in this article are those of the authors and do not 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. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects in research.
Abstract

Military physical training should increase combat readiness. Ability-performance modeling studies have shown that muscular strength and cardiovascular endurance are the keys to effective task performance, so training that improves these abilities improves readiness. This meta-analysis, which synthesized military physical training studies, showed that standard training practices produced large gains in muscular endurance and cardiovascular endurance. Standard training practices produced only minor improvements in muscular strength. Situational constraints limit the training options for military units, but some experimental programs have shown that training can be redesigned to promote muscular strength within those constraints. Modifications to training have the potential to better align training practices with readiness requirements.
Military units conduct physical training to improve combat readiness. Ability-performance modeling has shown that muscle strength and aerobic capacity are the keys to readiness (Vickers, 2009; Vickers, Hodgdon, & Beckett, 2009). Resistance training and endurance training develop these abilities (Jameson & Vickers, 2010; Vickers, Barnard, & Hervig, 2010), so appropriately designed physical training clearly could promote effective military task performance. However, military physical training is subject to time and equipment constraints. Military physical training must also promote nonphysical outcomes such as mental toughness and self-confidence. Given constraints and diffuse goals, the question arises: How much impact does traditional military physical training have on the physical abilities that are keys to readiness?

This paper reviews the research on the physical fitness outcomes from military physical training programs. The review constructs a broad training outcome profile and compares the strength and aerobic endurance outcomes with benchmarks from the physical training literature. This information provides the context for examining the alignment of military physical training with physical readiness requirements.

Methods

Literature Review

A literature search began by identifying relevant studies in the Defense Technical Information Center database and the Naval Health Research Center library. Subsequently, a PubMed search used the following keyword combinations: “physical” and “training,” “basic training” and “fitness,” and “military training” and “fitness” to identify additional articles. Article titles were reviewed as the first step in identifying relevant articles in the PubMed search.
Article abstracts were reviewed if the titles implied that the study dealt with military physical training outcomes. The full article was reviewed if the abstract indicated that the paper reported results from an experimental or quasi-experimental assessment of physical training outcomes. The article was retained for the meta-analysis if participants were drawn from a military population and the pre- and post-training means and SDs for one or more physical fitness outcomes were reported. The resulting set of 48 articles reported data from 96 samples (see Table 1).

Training Period

The analyses divided training effects into those produced by initial training and those produced by advanced training. Initial training studies took place during basic training, the first 8 to 12 weeks after entry into the service. Advanced training studies took the pre- and post-training measures between 12 weeks and 30 weeks after entry into the service. This temporal split reflected the fact that many recruits leave basic training and move on to some form of advanced entry-level training (e.g., Advanced Infantry Training, School of Infantry) during this time period. Too few data were available from studies conducted outside this entry-level time window to obtain reliable estimates of training effects for later phases of military service.

Training Outcomes

Training outcomes were grouped into four broad categories: muscle strength, muscle endurance, muscle power, and aerobic capacity (Hogan, 1985). Physical fitness tests were assigned to categories based on their common interpretations in the physical training literature and in work physiology research.

Any fitness test that had been investigated in only one or two samples was dropped from consideration. A single effect size (ES) might be a chance finding. Two ESs could differ widely,
thereby raising questions about the representativeness of the average value. Three ESs provided a minimum basis for establishing the consistency of training effects across samples and, if necessary, to identify a given sample as an outlier relative to two other ES estimates.

*Muscular strength.* Strength measures included dynamic lifts, isoinertial lifts, and isometric force tests. Dynamic lift tests were multijoint lifts similar to Olympic lifts (e.g., lifting a weight from the floor to shoulder height; see Stevenson, Bryant, Greenhorn, Deakin, & Smith, 1995). Isokinetic strength was excluded because this method had been used in just one study.

*Muscular endurance.* Muscle endurance measures assess the ability to generate continuous or repetitive submaximal forces. Push-ups, sit-ups, and pull-ups were the only measures in this category that met the minimal data requirement for analysis.

*Muscular power.* Muscular power, which is sometimes known as explosive strength, is measured by tests that require short bursts of maximal muscular exertion to propel an object. Initially, this category included sprints, vertical jump, and horizontal jump. A sprint was any continuous run covering up to 400 m, including shuttle runs. However, only the vertical jump had been studied often enough to be included in the analyses.

*Aerobic capacity.* Aerobic capacity was measured by laboratory tests and field tests. The laboratory tests were graded exercise tests that measured maximal oxygen uptake (VO$_{2\text{max}}$). The field tests were distance runs covering between 1.5 and 3 miles.

*Analysis Procedures*

A random effects (RE) meta-analysis was performed (Borenstein, Hedges, Higgins, & Rothstein, 2009). An RE model was chosen for two reasons. First, the results could be legitimately generalized beyond the studies in the analysis. Second, it was unlikely a priori that
all samples would share a common ES for any outcome, because training practices change over
time and training methods and emphases differ between services.

The RE analysis proceeded as follows. The pooled SD for the pre and posttraining scores
was computed. The difference between the posttraining average and the pretraining average was
divided by the pooled SD to obtain Cohen’s $d$. Hedge’s (1981) bias correction converted $d$ to $g$,
the ES index used in the analyses. An initial fixed effects analysis provided an estimate of $\tau^2$, the
RE variance. The RE weight for each ES was based on its total variance, which was the sum of
$\tau^2$ and sampling variance.

The assessment of strength outcomes required preliminary work before the RE analysis
could be carried out. Some studies administered multiple strength tests. Treating each strength
test as an independent estimate of training effects would have given too much weight to samples
that completed multiple tests. To avoid this, the strength test ESs were based on the overall
average scores for the set of strength tests used in the study. The mean values for individual tests
were averaged to obtain overall pre- and post-training strength test scores. The variances of the
overall average scores were computed by combining the reported SDs for the individual tests
scores with estimates of the inter-test correlations (see Nunnally & Bernstein, 1994, p. 162). As
no study reported the inter-test correlations, estimates had to be derived from results reported in
other studies. The inter-test correlation estimates were: dynamic strength, $r = .89$, $k = 5$, $N =
1,353$ participants; isoinertial strength, $r = .59$, $k = 8$, $N = 1,204$; isometric strength, $r = .41$, $k =
3$, $N = 301$. The pre and posttraining cumulative average scores and associated cumulative SDs
defined the strength ES.

The general RE method was applied individually to each outcome that was represented in
the data by three or more independent ES estimates. This decision balanced Borenstein et al.’s
(2009) recommendation of at least five ES estimates for RE modeling against the desire to have
the broadest possible coverage of training outcomes. However, the potential inaccuracy in $\tau^2$
estimates should be kept in mind when examining the results for infrequently studied outcomes.

Cohen’s (1988) ES criteria provided the framework for interpreting the pooled ES values. Each pooled effect was classified as trivial (i.e., $ES < 0.20$), small (i.e., $0.20 \leq ES < 0.50$), medium (i.e., $0.50 \leq ES < 0.80$), or large (i.e., $ES \geq 0.80$). These ES criteria have simple practical interpretations. A small ES meant that training shifted the score distribution enough so that 60% to 70% of program participants’ posttraining scores were higher than the pretraining average. A moderate ES meant that 70% to 80% of post-training scores exceeded the pre-training average. A large ES meant that more than 80% of posttraining scores exceeded the pretraining average. The training effect was trivial if fewer than 60% of posttraining scores exceeded the pretraining average.

An SPSS syntax program was constructed to implement Borenstein et al.’s (2009) procedures. Program accuracy was verified by repeating the analysis examples in the text. All analyses were performed using SPSS-PC, version 17, computer software.

Results

Standard Training

Initial training. Nine initial training outcomes met the $k \geq 3$ criterion (Table 2). Initial training produced moderate to large effects on muscular endurance (sit-ups, push-ups, pull-ups) and cardiovascular endurance (distance runs, VO$_{2\text{max}}$). The initial training effects for muscular strength and power were trivial except for a small effect on isometric strength.

Advanced training. The advanced training studies reported 31 ESs for eight outcomes (Table 3). The advanced training effects were uniformly weak. The dynamic strength outcome
(ES = .374) appears to be at odds with this generalization, but this appearance should be judged in the context of the wide confidence interval for the ES estimate. The confidence interval includes ES = .00, so there is reason to downplay the apparent deviation from the general trend.

**Benchmark Comparison**

The benchmark comparisons focused on muscle strength and cardiovascular endurance. The benchmarks were ESs associated with typical resistance and aerobic training programs implemented with untrained participants. A typical strength program involved three training sessions per week with three sets of each exercise and 8 to 10 repetitions per set. A typical aerobic training program consisted of training 3 days per week for 31 to 45 min at ~75% to 80% of VO\(_{2\text{max}}\). The typical resistance training program produced ES = 1.045 for isoinertial strength measures (\(k = 28\)). The typical aerobic training program produced ES = .756 for VO\(_{2\text{max}}\) (\(k = 31\)).

Military training gains were significantly smaller than the benchmarks. Strength gains were 14% of the benchmark (\(z = -10.09, p < .001\)). Aerobic gains were 79% of the benchmark (\(z = -4.04, p < .001\)).

**Experimental Training**

Can training be more productive within military training constraints? Past experiments provide an empirical basis for answering this question. The literature describing those experiments is diffuse. Training goals have varied from study to study and the experimental programs have been matched to the goals. The program differences make it unreasonable to speak of the typical or average experimental program training effect. Therefore, experimental programs are considered here on a study-by-study basis.

The experimental training programs review focused on two basic questions. First, did the experimental program yield training outcomes that were significantly better than the standard
program outcomes? If so, was the difference large enough to be practically important? The first
question had to be asked to rule out the possibility that an observed difference was a chance
finding (Abelson, 1997). The second question allowed for the possibility that a large sample size
had inflated the apparent importance of a small difference (Rosenthal & Rosnow, 1984).

Harman, et al. (2008). This experimental program, which was conducted during the
initial training period, added resistance training to the standard program. Compared with the
standard program, the experimental program produced larger gains in isoinertial strength,
VO$_{2\text{max}}$, and vertical jump. The experimental program produced smaller gains in sit-ups, push-
ups, and long-distance run performance, but these gains were only small relative to the large
improvement typical of standard training. The simple experimental ES was moderate to large
($0.74 \leq \text{ES} \leq 1.26$) for each of these “impaired” outcomes.

Knapik et al. (2005). This study, which was conducted during the initial training period,
introduced a complex variation on standard training. The experiment introduced ability group
runs, stretching, movement drills, and calisthenics. The calisthenics included push-ups, pull-ups,
and sit-ups. The experimental program produced dramatically larger gains in sit-up performance
(men, ES = 1.45, $t = 14.50$, $p < .001$; women, ES = 1.95, $t = 16.25$, $p < .001$). The experimental
push-up gains equaled the standard program gains (men, ES = .08, $t = 1.13$, $p = .130$; women, ES
= .03, $t = .32$, $p = .374$). Overall, the program improved sit-up outcomes and maintained push-up
outcomes.

Santtila, Keijo, Laura, and Heikki (2008). This study included two experimental
modifications to initial training. One experimental group participated in a program that
emphasized endurance training. The other experimental group participated in a program that
emphasized strength training. Both modifications were superimposed on standard training. The
programs produced small, statistically nonsignificant improvements in isoinertial strength and $VO_{2\text{max}}$. The experimental gains were not significantly different from the standard training gains.

*Williams, Rayson, and Jones (2002).* The experimental program in this study, which was conducted during the initial training period, included resistance training and augmented endurance training. The experimental program produced larger $VO_{2\text{max}}$ and dynamic strength gains than seen with standard training (Table 5). The experimental program produced small, statistically nonsignificant decrements in isometric strength that contrasted with small positive effects in standard training (see Table 2). The resulting difference was statistically significant for men.

*Rasch, Otott, Wilson, Brown, and Norton (1965).* This study implemented two experimental programs during initial training. Each experimental program combined interval training with circuit training. One group performed the two types of training on different days, while the other group performed both each training day. Isometric strength and pull-up performance were the only measured training outcomes that met the inclusion criteria for the analysis. The experimental outcomes were no different from the standard outcomes. This negative finding may not be surprising given that the nominal experimental programs were standard training for other countries.

*Knapik, Bahrke, Staab, Reynolds, and Vogel (1990).* This study, which was conducted during advanced training, examined the effects of systematically training for loaded road marches. The experiment added loaded training marches to the standard training program. Three experimental programs were defined by varying the number of marches per week. A group that performed no loaded marches provided a control comparison. The training outcomes were sit-ups, push-ups, and long-distance run performance. Experimental outcomes generally
approximated standard outcomes; differences were small and not statistically significant (Table 6).

*Kraemer et al. (2004)*. This study, which was conducted during the advanced training period, examined the separate and combined effects of augmented endurance training and resistance training. One experimental program focused solely on endurance training, a second experimental program focused entirely on resistance training. The two remaining experimental programs combined resistance training and endurance training. One combined program consisted of whole-body resistance training and endurance training; the other combined program consisted of upper body resistance training and endurance training. The experimental programs produced better outcomes than standard training in 11 of 12 comparisons with sit-ups, push-ups, and long-distance run performance as outcomes (Table 7). The differences were moderate to large, and 6 of the 11 were statistically significant. The exception to the general trend was the null effect of resistance training on long-distance run performance. However, even this result did not indicate a significantly poorer outcome because long-distance run performance improves only slightly in standard advanced training.

*Males, Sekulic, and Katic (2004)*. Males et al. studied an experimental program during advanced training. The program increased the frequency and intensity of standard training activities. Specifically, the experimental program (when compared with the control program) implemented more intense aerobic training that pushed the subjects into an anaerobic endurance zone. In addition, the experimental program included twice the amount of flexibility training as the control group. Relative to standard training, the experiment produced moderate to large, statistically significant improvements in isoinertial strength, sit-ups, push-ups, pull-ups, and long-distance runs (Table 8).
Knapik et al. (2004). This study evaluated a new Army Physical Readiness Training program implemented in advanced training. The new program combined progressive calisthenics with movement exercises, interval running, and ability-group endurance runs. The new program produced trivial improvements relative to standard training (Table 9). Those small differences were statistically significant because the sample was very large. However, this program may have two advantages over traditional training: the injury rate was lower (see Vickers, 2007, for the combined results of several studies), and the cardiovascular endurance gains were achieved with less running distance.

Summary of findings for experimental training programs. On the whole, the experimental programs produced better outcomes than standard training. When the specific effects of the various programs are considered, the programs implemented by Harman and colleagues (2008) and Williams (2005) are the most promising. Both of those programs improved strength and cardiovascular endurance, so both should better align training with performance requirements effects that should translate into improvements in task performance (Vickers, 2009; Vickers, Hodgdon, & Beckett, 2009).

Discussion

Initial training dramatically improved muscular endurance and cardiovascular endurance. Those outcomes contrasted markedly with trivial changes in muscular strength and muscular power.

Advanced training gains were much smaller than initial training gains. This result was predictable because training routinely produces smaller gains in subjects with higher levels of initial fitness than in less fit subjects (Jameson & Vickers, 2010; Vickers et al., 2010).
The alignment of training with military task demands could be improved. Performance modeling implicates muscular strength and cardiovascular endurance as the most important physical abilities for military tasks (see Vickers, 2009; Vickers et al., 2009). Standard training produces substantial cardiovascular endurance gains, but does little to develop muscular strength. Modifying programs to provide greater strength gains would better align training with performance requirements.

Program modifications are feasible. Experimental training programs have had much stronger effects on the critical abilities with little or no loss in the benefits derived from standard training. The experimental training programs that have been introduced fit the constraints found in military settings. The experimental modifications vary too widely to be encapsulated by a general summary, but there is evidence that properly designed programs can produce better muscular strength and cardiovascular endurance outcomes than current training practices (Harman et al., 2008; Williams et al., 2002). Other programs have shown that current training outcomes can be maintained while reducing the injury rate (see Vickers, 2007, summary of work by Knapik and colleagues). Greater use of either type of program will produce a higher net readiness level for military units.

The fitness elements that are most affected by standard training suggest a mechanism for aligning training with readiness requirements. Standard training has produced major gains in sit-up, push-up, pull-up, and long-distance run performances. Each of these outcomes is a common element of military physical fitness tests. The substantial improvements in these measures suggest that training programs “train to the test.” Introducing suitable physical readiness tests could bring this same tendency to bear on muscular strength and muscular power. For example, the Combat Fitness Test (CFT) recently implemented by the U.S. Marine Corps includes
elements that appear to depend on muscular power. CFT implementation can be expected to increase the emphasis on developing muscular power.

One major gap in the evidence could be critical for combat readiness. Studies of military physical training have been limited almost entirely to the first 6 months of service. The paucity of data on the effects of later training means that little or nothing is known about the effectiveness of training that takes place in the crucial period(s) leading up to combat deployment. Extrapolating from the modest gains seen in advanced training, subsequent training may do little more than maintain the fitness achieved during entry-level training. Simple maintenance is acceptable if earlier training has produced the physical capabilities needed to meet operational demands. In fact, the opportunity cost of investing time and other resources in physical training might not justify increased physical training. The time and other resources might be better spent on other performance improvement methods (e.g., small-unit tactical training).

The available evidence limits the precision and scope of the inferences that can be drawn from this review. Few outcomes have been studied enough to develop precise training effect estimates. The emphasis on recruit training, while admittedly focused on a critical period of each recruit’s service, makes it impossible to determine the effects of training later in the individual’s military service. Experimental programs have shown promise, but only a few possibilities have been investigated. The experiments that have been tried have not been replicated, so caution is appropriate when making assertions about their effects. Finally, the distinction between standard training and experiments may be inconsistent in a review that draws on data from a time span of nearly 50 years. Today’s experiment, if successful, can become tomorrow’s standard method.
Each of these limitations represents an opportunity for additional constructive research to align physical training with readiness.

In conclusion, initial military physical training develops muscular endurance and cardiovascular endurance. Advanced physical training adds little to the initial gains. Consideration should be given to modifying traditional training practices to improve muscular strength and muscular power outcomes. Recent U.S. Marine Corps instructions provide the policy context for novel program designs (Commandant of the Marine Corps, 2002, 2008). The road ahead should include systematic evaluation of training options designed to align physical training practices with anticipated combat requirements. The recent implementation of the CFT provides organizational incentive for changing training practices within the U.S. Marine Corps. Steps could be taken to capitalize on the resulting opportunity to implement and formally evaluate alternative training regimens.
References

*References marked with an asterisk indicates studies that provided data for the analyses.

Abelson, R. P. (1997). A retrospective on the significance test ban of 1999 (If there were no significance tests, they would be invented). In L. L. Harlow, S. A. Mulaik & J. H. Steiger (Eds.), What if there were no significance tests? (pp. 117–141). Mahwah, NJ: Lawrence Erlbaum.


Table 1

Demographic Characteristics

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>ΣN</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>68</td>
<td>304</td>
<td>21.23</td>
<td>2.76</td>
<td>18.0</td>
<td>29.0</td>
</tr>
<tr>
<td>Height</td>
<td>67</td>
<td>325</td>
<td>172.98</td>
<td>6.38</td>
<td>159.9</td>
<td>184.07</td>
</tr>
<tr>
<td>Weight</td>
<td>75</td>
<td>362</td>
<td>70.57</td>
<td>7.99</td>
<td>56.1</td>
<td>108.30</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>44</td>
<td>227</td>
<td>19.58</td>
<td>6.53</td>
<td>8.1</td>
<td>34.3</td>
</tr>
</tbody>
</table>
Table 2

Initial Training Effects

<table>
<thead>
<tr>
<th>Outcome</th>
<th>k</th>
<th>Avg</th>
<th>SE</th>
<th>Lower</th>
<th>Upper</th>
<th>Q</th>
<th>Sig</th>
<th>z</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sit-ups</td>
<td>17</td>
<td>1.690</td>
<td>.195</td>
<td>1.351</td>
<td>2.030</td>
<td>13.10</td>
<td>.666</td>
<td>8.69</td>
<td>.000</td>
</tr>
<tr>
<td>Push-ups</td>
<td>15</td>
<td>1.561</td>
<td>.073</td>
<td>1.433</td>
<td>1.689</td>
<td>21.04</td>
<td>.101</td>
<td>21.50</td>
<td>.000</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>25</td>
<td>1.094</td>
<td>.055</td>
<td>1.000</td>
<td>1.188</td>
<td>75.62</td>
<td>.000</td>
<td>19.89</td>
<td>.000</td>
</tr>
<tr>
<td>Moderate effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pull-ups</td>
<td>4</td>
<td>.631</td>
<td>.111</td>
<td>.370</td>
<td>.893</td>
<td>3.02</td>
<td>.389</td>
<td>5.68</td>
<td>.000</td>
</tr>
<tr>
<td>Laboratory VO_{2max}</td>
<td>26</td>
<td>.441</td>
<td>.078</td>
<td>.307</td>
<td>.574</td>
<td>28.29</td>
<td>.295</td>
<td>5.64</td>
<td>.000</td>
</tr>
<tr>
<td>Small effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isometric strength</td>
<td>20</td>
<td>.234</td>
<td>.076</td>
<td>.104</td>
<td>.365</td>
<td>23.55</td>
<td>.214</td>
<td>3.10</td>
<td>.001</td>
</tr>
<tr>
<td>Unimportant effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic strength</td>
<td>12</td>
<td>.147</td>
<td>.089</td>
<td>-.014</td>
<td>.307</td>
<td>10.23</td>
<td>.510</td>
<td>1.64</td>
<td>.050</td>
</tr>
<tr>
<td>Isoinertial strength</td>
<td>3</td>
<td>-.004</td>
<td>.280</td>
<td>-.822</td>
<td>.814</td>
<td>2.64</td>
<td>.267</td>
<td>-.01</td>
<td>.506</td>
</tr>
<tr>
<td>Isokinetic strength</td>
<td>4</td>
<td>-.063</td>
<td>.110</td>
<td>-.323</td>
<td>.196</td>
<td>2.17</td>
<td>.538</td>
<td>-.58</td>
<td>.718</td>
</tr>
<tr>
<td>Vertical jump</td>
<td>4</td>
<td>-.210</td>
<td>.117</td>
<td>-.485</td>
<td>.065</td>
<td>3.07</td>
<td>.380</td>
<td>-1.80</td>
<td>.964</td>
</tr>
</tbody>
</table>

Note. Table gives random-effects analysis results for all training outcomes with ES estimates based on three or more samples. Outcomes listed from largest ES to smallest ES. ES classification based on Cohen (1988).
Table 3

*Advanced Training Effects*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>k</th>
<th>Avg</th>
<th>SE</th>
<th>Lower</th>
<th>Upper</th>
<th>Q</th>
<th>Sig</th>
<th>z</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Small effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic strength</td>
<td>4</td>
<td>.374</td>
<td>.335</td>
<td>-.414</td>
<td>1.162</td>
<td>4.00</td>
<td>.262</td>
<td>1.12</td>
<td>.132</td>
</tr>
<tr>
<td>Isometric strength</td>
<td>5</td>
<td>.264</td>
<td>.144</td>
<td>-.042</td>
<td>.571</td>
<td>4.88</td>
<td>.300</td>
<td>1.84</td>
<td>.033</td>
</tr>
<tr>
<td>Sit-ups</td>
<td>5</td>
<td>.242</td>
<td>.068</td>
<td>.097</td>
<td>.387</td>
<td>4.72</td>
<td>.317</td>
<td>.355</td>
<td>.000</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>6</td>
<td>.218</td>
<td>.128</td>
<td>-.039</td>
<td>.475</td>
<td>12.49</td>
<td>.019</td>
<td>1.71</td>
<td>.044</td>
</tr>
<tr>
<td>Push-ups</td>
<td>3</td>
<td>.204</td>
<td>.039</td>
<td>.091</td>
<td>.318</td>
<td>2.15</td>
<td>.342</td>
<td>5.25</td>
<td>.000</td>
</tr>
<tr>
<td><strong>Unimportant effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pull-ups</td>
<td>4</td>
<td>-.057</td>
<td>.104</td>
<td>-.302</td>
<td>.187</td>
<td>3.00</td>
<td>.392</td>
<td>-.55</td>
<td>.710</td>
</tr>
<tr>
<td>Isoinertial strength</td>
<td>3</td>
<td>-.144</td>
<td>.096</td>
<td>-.425</td>
<td>.138</td>
<td>1.41</td>
<td>.495</td>
<td>-.149</td>
<td>.932</td>
</tr>
</tbody>
</table>

*Note.* Table gives random-effects analysis results for all training outcomes with ES estimates based on three or more samples. Outcomes listed from largest ES to smallest ES. ES classification based on Cohen (1988).
Table 4

Results of Harman et al. (2008) Resistance Training Program in Initial Training

<table>
<thead>
<tr>
<th>Outcome</th>
<th>g</th>
<th>SE&lt;sub&gt;g&lt;/sub&gt;</th>
<th>ES&lt;sup&gt;a&lt;/sup&gt;</th>
<th>z&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoinertial strength</td>
<td>1.26</td>
<td>.39</td>
<td>1.26</td>
<td>3.24</td>
<td>.001</td>
</tr>
<tr>
<td>Sit-ups</td>
<td>.77</td>
<td>.37</td>
<td>-.92</td>
<td>-2.49</td>
<td>.006</td>
</tr>
<tr>
<td>Push-ups</td>
<td>1.14</td>
<td>.38</td>
<td>-.42</td>
<td>-1.11</td>
<td>.134</td>
</tr>
<tr>
<td>Vertical jump</td>
<td>.34</td>
<td>.36</td>
<td>.55</td>
<td>1.53</td>
<td>.063</td>
</tr>
<tr>
<td>VO&lt;sub&gt;2&lt;/sub&gt;max</td>
<td>1.10</td>
<td>.38</td>
<td>.66</td>
<td>1.73</td>
<td>.041</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>.74</td>
<td>.37</td>
<td>-.35</td>
<td>-.96</td>
<td>.169</td>
</tr>
</tbody>
</table>

<sup>a</sup>ES = (Program g – Standard Training g) where the standard training g is the appropriate value from Table 3.

<sup>b</sup>z = ES/SE<sub>g</sub>. 
Table 5

Results of Williams et al. (2002) Resistance Training Experiment in Initial Training

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Gender</th>
<th>g</th>
<th>SE</th>
<th>ES(^a)</th>
<th>z(^b)</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic strength</td>
<td>Men</td>
<td>.71</td>
<td>.23</td>
<td>.56</td>
<td>2.45</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>.99</td>
<td>.50</td>
<td>.84</td>
<td>1.69</td>
<td>.046</td>
</tr>
<tr>
<td>Isometric strength</td>
<td>Men</td>
<td>-.32</td>
<td>.23</td>
<td>-.55</td>
<td>-2.41</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>-.05</td>
<td>.47</td>
<td>-.28</td>
<td>-.60</td>
<td>.273</td>
</tr>
<tr>
<td>VO(_2)max</td>
<td>Men</td>
<td>.99</td>
<td>.26</td>
<td>.55</td>
<td>2.11</td>
<td>.017</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>3.27</td>
<td>.80</td>
<td>2.83</td>
<td>3.54</td>
<td>.000</td>
</tr>
</tbody>
</table>

\(^a\)ES = (Program \(g\) – Standard Training \(g\)) where the standard training \(g\) is the appropriate value from Table 3.

\(^b\)z = ES/SE\(_g\).
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Condition</th>
<th>g</th>
<th>SE</th>
<th>ES</th>
<th>z</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sit-ups</td>
<td>0 marches/week</td>
<td>.61</td>
<td>.31</td>
<td>.37</td>
<td>1.19</td>
<td>.118</td>
</tr>
<tr>
<td></td>
<td>1 march/week</td>
<td>.00</td>
<td>.38</td>
<td>-.24</td>
<td>-.64</td>
<td>.262</td>
</tr>
<tr>
<td></td>
<td>2 marches/week</td>
<td>.22</td>
<td>.25</td>
<td>-.02</td>
<td>-.09</td>
<td>.465</td>
</tr>
<tr>
<td></td>
<td>4 marches/week</td>
<td>.12</td>
<td>.30</td>
<td>-.12</td>
<td>-.41</td>
<td>.342</td>
</tr>
<tr>
<td>Push-ups</td>
<td>0 marches/week</td>
<td>.81</td>
<td>.32</td>
<td>.61</td>
<td>1.89</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>1 march/week</td>
<td>.48</td>
<td>.39</td>
<td>.28</td>
<td>.71</td>
<td>.240</td>
</tr>
<tr>
<td></td>
<td>2 marches/week</td>
<td>.53</td>
<td>.26</td>
<td>.33</td>
<td>1.25</td>
<td>.105</td>
</tr>
<tr>
<td></td>
<td>4 marches/week</td>
<td>.53</td>
<td>.31</td>
<td>.33</td>
<td>1.05</td>
<td>.146</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>0 marches/week</td>
<td>-.16</td>
<td>.30</td>
<td>-.38</td>
<td>-1.26</td>
<td>.104</td>
</tr>
<tr>
<td></td>
<td>1 march/week</td>
<td>-.32</td>
<td>.38</td>
<td>-.54</td>
<td>-1.42</td>
<td>.078</td>
</tr>
<tr>
<td></td>
<td>2 marches/week</td>
<td>-.00</td>
<td>.25</td>
<td>-.22</td>
<td>-.87</td>
<td>.192</td>
</tr>
<tr>
<td></td>
<td>4 marches/week</td>
<td>-.23</td>
<td>.30</td>
<td>-.45</td>
<td>-1.49</td>
<td>.068</td>
</tr>
</tbody>
</table>

\(^a\)ES = (Program g – Standard Training g) where the standard training g is the appropriate value from Table 4.

\(^b\)z = ES/SE_g.
Table 7

Results of Kraemer et al. (2004) Experiment in Advanced Training

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Treatment</th>
<th>g</th>
<th>SE</th>
<th>ES&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Z&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sit-ups</td>
<td>Resistance</td>
<td>3.03</td>
<td>.68</td>
<td>2.79</td>
<td>4.10</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Endurance</td>
<td>.90</td>
<td>.50</td>
<td>.66</td>
<td>1.32</td>
<td>.094</td>
</tr>
<tr>
<td></td>
<td>Whole body + endurance</td>
<td>1.51</td>
<td>.51</td>
<td>1.27</td>
<td>2.49</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>Upper body + endurance</td>
<td>.91</td>
<td>.47</td>
<td>.67</td>
<td>1.42</td>
<td>.078</td>
</tr>
<tr>
<td>Push-ups</td>
<td>Resistance</td>
<td>1.96</td>
<td>.56</td>
<td>1.76</td>
<td>3.14</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>Endurance</td>
<td>.77</td>
<td>.49</td>
<td>.57</td>
<td>1.16</td>
<td>.124</td>
</tr>
<tr>
<td></td>
<td>Whole body + endurance</td>
<td>1.87</td>
<td>.55</td>
<td>1.67</td>
<td>3.03</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>Upper body + endurance</td>
<td>1.08</td>
<td>.48</td>
<td>.88</td>
<td>1.83</td>
<td>.034</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>Resistance</td>
<td>.00</td>
<td>.45</td>
<td>-.22</td>
<td>-.48</td>
<td>.314</td>
</tr>
<tr>
<td></td>
<td>Endurance</td>
<td>.79</td>
<td>.49</td>
<td>.57</td>
<td>1.14</td>
<td>.122</td>
</tr>
<tr>
<td></td>
<td>Whole body + endurance</td>
<td>.83</td>
<td>.47</td>
<td>.61</td>
<td>1.30</td>
<td>.096</td>
</tr>
<tr>
<td></td>
<td>Upper body + endurance</td>
<td>1.99</td>
<td>.56</td>
<td>1.77</td>
<td>3.16</td>
<td>.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>ES = (Program g – Standard Training g) where the standard training g is the appropriate value from Table 4.

<sup>b</sup>z = ES/SE<sub>g</sub>.
Table 8

Results of Males et al. (2004) Frequency and Intensity Experiment During Advanced Training

<table>
<thead>
<tr>
<th>Outcome</th>
<th>$g$</th>
<th>$SE$</th>
<th>$ES^a$</th>
<th>$Z^b$</th>
<th>$Sig$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoinertial</td>
<td>.60</td>
<td>.10</td>
<td>.74</td>
<td>7.44</td>
<td>.000</td>
</tr>
<tr>
<td>Sit-ups</td>
<td>1.13</td>
<td>.11</td>
<td>.89</td>
<td>8.07</td>
<td>.000</td>
</tr>
<tr>
<td>Push-ups</td>
<td>1.04</td>
<td>.11</td>
<td>.84</td>
<td>7.60</td>
<td>.000</td>
</tr>
<tr>
<td>Pull-ups</td>
<td>.54</td>
<td>.10</td>
<td>.60</td>
<td>5.97</td>
<td>.000</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>.74</td>
<td>.10</td>
<td>.52</td>
<td>5.22</td>
<td>.000</td>
</tr>
</tbody>
</table>

$^a ES = (Program\ g – Standard Training\ g)$ where the standard training $g$ is the appropriate value from Table 4.

$^b z = ES/SE_g$. 
Table 9

Results of Knapik et al. (2004) Modified Calisthenics Program in Advanced Training

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Gender</th>
<th>g</th>
<th>SE</th>
<th>ES(^a)</th>
<th>z(^b)</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sit-ups</td>
<td>Men</td>
<td>.38</td>
<td>.04</td>
<td>.14</td>
<td>3.45</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>.43</td>
<td>.11</td>
<td>.19</td>
<td>1.71</td>
<td>.044</td>
</tr>
<tr>
<td>Push-ups</td>
<td>Men</td>
<td>.27</td>
<td>.04</td>
<td>.07</td>
<td>1.65</td>
<td>.049</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>.35</td>
<td>.11</td>
<td>.15</td>
<td>1.33</td>
<td>.092</td>
</tr>
<tr>
<td>Long-distance run</td>
<td>Men</td>
<td>.15</td>
<td>.04</td>
<td>-.07</td>
<td>-1.70</td>
<td>.045</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>.23</td>
<td>.11</td>
<td>.01</td>
<td>.11</td>
<td>.457</td>
</tr>
</tbody>
</table>

\(^a\)ES = (Program g – Standard Training g) where the standard training g is the appropriate value from Table 4.

\(^b\)z = ES/SE\(_g\).
### ABSTRACT

Military physical training should increase combat readiness. Ability-performance modeling studies have shown that muscular strength and cardiovascular endurance are the keys to effective task performance, so training that improves these abilities improves readiness. This meta-analysis, which synthesized military physical training studies, showed that standard training practices produced large gains in muscular endurance and cardiovascular endurance. Standard training practices produced only minor improvements in muscular strength. Situational constraints limit the training options for military units, but some experimental programs have shown that training can be redesigned to promote muscular strength within those constraints. Modifications to past training practices have the potential to better align training practices with readiness requirements.