



Resistance Training Increases the Variability of Strength Test Sores

Ross R. Vickers, Jr.
Amanda C. Barnard
Linda K. Hervig

Naval Health Research Center

Document No. 09-29

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

Resistance Training Increases the Variability of Strength Test Scores

Ross R. Vickers, Jr.
Amanda C. Barnard
Linda K. Hervig

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

This represents report 09-29 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

Informal observations made while conducting a meta-analysis of resistance training programs suggested that the between-person variation in strength test scores is greater after training than before. This study treated the informal observation as a hypothesis to be evaluated. The odds were 2.5:1 that the standard deviation after training would be larger than the standard deviation before training if a sample underwent resistance training, compared with 1:1 odds for control groups in training studies. This difference supported the study hypothesis. Extending the analysis to subcategories based on age, gender, training experience, and training program characteristics, the training effect was present in all subgroups, but it was significantly stronger in some than others (e.g., older or novice lifters). The training effect did not increase with program length. This fact and the weak effect in experienced lifters suggested that the training effect might be a product of neuromuscular adaptations occurring early in training. Whatever its source, the training effect seldom will change inferences about whether resistance training has increased the strength of program participants, but it could be important when predicting how training affects the ability to meet basic performance standards.

Resistance training improves strength (Falk & Tenenbaum, 1996; Payne, Morrow, Johnson, & Dalton, 1997; Peterson, Rhea, & Alvar, 2004; Rhea & Alderman, 2004; Rhea, Alvar, & Burkett, 2002; Rhea, Alvar, Burkette, & Ball, 2003; Wolfe, LeMura, & Cole, 2004). An informal observation made recently while conducting a meta-analysis of the literature suggested a second training effect. Training appeared to increase test score variability. Specifically, posttraining standard deviations appeared to be consistently larger than pretraining standard deviations.

If the hypothesized increase in variability exists, knowledge of this fact would be important in two respects. Increased variability implies that different people respond differently to training programs. Some trainees must benefit more than others if test score variation is to increase. Understanding who benefits most and why could lead to the design of more efficient training regimens. Increased variability would also be important for any attempt to model the effects of resistance training. Models would have to include changes in variation in addition to changes in the average strength test score to provide a complete and accurate representation of training effects.

This report provides evidence that test score variation generally does increase from pretraining to posttraining. The report also identifies the types of training subjects for whom this effect is most pronounced.

Methods

Literature Review

The data for the present analyses came from 196 studies identified in a previous meta-analysis of the resistance training literature (Vickers, Hervig, & Barnard, manuscript in preparation). The meta-analysis relied on the reference lists from prior meta-analyses and a computer search of the PubMed database to identify articles. The search was limited to English language reports published in various journals. Details of the search can be found in the earlier report.

Data Coding

The information extracted during the meta-analysis included the sample size, means, and standard deviations for pretest and posttest strength measurements. This information was recorded for every strength test used in a total of 377 samples described in the 196 studies. Gender, age, and training history were among the other variables recorded in the meta-analysis. For gender, samples were classified as consisting of men, women, or both men and women. For age, samples were classified as “younger” if the available information indicated that the average age was <50 years. Samples were classified as “older” if the available information indicated that the average age was ≥ 50 years. These determinations were based on quantitative information (i.e., the average age of the study participants or an age range for the participants) when it was available. When quantitative information was not available, the age category was based on qualitative information that strongly suggested membership in one of the two categories (e.g., college students, free-living elderly). Training history distinguished samples of experienced lifters from novice lifters. Experienced lifters were individuals with recent and probably ongoing training experience (e.g., weight lifters, athletes in weight training programs). Novice lifters included any individuals who had no recent history of resistance training. Some nominal novices may have had lifting experience in the past, but the study descriptions suggested that many had no such experience. While recognizing the imprecision of the term, the novice designation was adopted to emphasize the difference between these study participants and the recent, probably ongoing training of experienced lifters.

Analysis Procedures

The nonparametric sign test (Siegel, 1956) was the primary analysis procedure. This statistic was chosen because it did not require any assumptions about the distributions of the standard deviations or the correlation of pretraining standard deviations with posttraining standard deviations. The *z* score from the sign test was used to decide whether the hypothesized training effect was present.

Output from the sign test was used to compute the odds of a training effect. The sign test output provided a count of the number of positive events (i.e., final standard deviation [F] > initial standard deviation [I]) and negative events (i.e., F ≤ I). This information was used to compute odds, $Odds = p_s / p_n = p_s /(1 - p_s)$.

Subgroups were compared to determine the generality of the training effect. Odds ratios (ORs) were computed for this purpose. The subgroup represented by the largest number of test scores was identified for each demographic variable and each program characteristic. The OR for a given group was the ratio of the odds for that group relative to the chosen reference group. The χ^2 test associated with each OR indicated whether the difference in the odds for the two groups being compared was statistically significant. ORs were converted to effect size (ES) estimates by taking their natural logarithms and dividing by 1.81 (Chinn, 2000). Cohen’s (1988) ES criteria were used to classify the group differences as trivial (ES < .20), small (.20 ≤ ES < .50), moderate (.50 ≤ ES <.80), or large (ES ≥ .80).

The sign test would be significant if the posttest standard deviation consistently was larger than the pretest standard deviation. This condition could be satisfied even if the difference in the standard deviations was small. For example, the posttest standard deviation might be 1% larger than the pretest standard deviation in every case. The training effect would be present, but there would be good reason to doubt that it was important. The geometric mean of the ratio of posttest to pretest standard deviations was computed to deal with this problem. The computation involved several steps.

1. The natural logarithm transformation was applied to each standard deviation.
2. The difference between the pre- and post-training natural logarithms of the standard deviation was computed.

Table 1. Training Effect as a Function of Training Status

	-[a]	+[a]	0[a]	z	Sig.	Odds	OR	ES	χ^2	Sig.
Training status										
Control	102	100	14	-.070	.944	.98	–[b]			
Placebo	20	51	7	-3.56	.000	2.55	2.60	.53	10.59	.002
Regular Training	172	443	26	-10.89	.000	2.58	2.63	.53	34.62	<.001

[a]"-" = negative event (i.e., Final *SD* < Initial *SD*), "+" = positive event (i.e., Final *SD* > Initial *SD*), and "0" indicates a neutral event (i.e., Final *SD* = Initial *SD*). A positive effect was consistent with the hypothesis that training increased the variation in test scores. Thus, observations coded "+" indicate the presence of a training effect as defined here.
[b]Reference group for odds ratios.

3. The variances for the log-transformed standard deviations were computed as $\sigma^2 = \frac{1}{2(n-1)}$ (Raudenbush & Bryk, 2002, p. 219).
4. The variance of the difference between the log-transformed standard deviations, $\sigma_{Diff}{}^2 = \sigma_1{}^2 + \sigma_2{}^2 - 2r_{12}\sigma_1\sigma_2$, was computed. The computations employed separate pre-post correlations for each of 10 strength tests that had been administered to ≥20 samples. Other tests had been used too infrequently to obtain a reasonably precise estimate of the post-pre correlation.
5. The inverse of the variance for the difference was the weight variable used to compute the average difference between the log-transformed pre- and posttraining standard deviations.
6. The exponential transformation was applied to the weighted average of the differences to obtain the geometric mean of the ratios of the posttest standard deviation to the pretest standard deviation.

All of the data analyses, except the computation of ORs, were carried out with SPSS-PC, version 15 (SPSS, Inc., Chicago, IL). ORs were computed using Dean, Sullivan, and Soe's (2009) online calculator available at www.OpenEpi.com.

Results

Training Effect

The overall results supported the hypothesis that resistance training increases the variation in strength test scores (Table 1). In control samples, the standard deviation was about equally likely to decrease as it was to increase after training (odds = .98, z = .07, p = .944). In samples that underwent training, the standard deviation was more than 2.5 times as likely to increase as it was to decrease, a clearly significant trend (p < .001).

The ORs clearly differentiated training samples from control samples. OR was statistically significant ($p \leq .002$) whether the study focused on training (OR = 2.58) or on evaluating a supplement (OR = 2.55). The effect size, ES = .53, was moderately large for each comparison.

Study focus did not affect the response to training. The OR comparing training-focused studies to supplement-focused studies showed virtually no difference (OR = 1.01, $\chi^2 = .00$, $p = .971$). Based on this, the results of training-focused studies and supplement-focused studies were combined for subsequent analyses.

Moderator Variables

A moderator is any characteristic of program participants or program design that affects the strength of the training effect. Analyses that were carried out to test for moderator effects indicated that the training effect was a very general phenomenon (Table 2, p. 5). A significant training effect was evident in every subgroup examined in these analyses, although the trend was weak for experienced lifters (odds = 1.33, z = 2.03, $p < .043$).

Although the training effect was present in every subgroup, the ORs in Table 2 indicate that the strength of this effect varied across subgroups. The χ^2 that accompanied each OR was used to determine which differences were statistically significant. Specific findings were:

Gender: The training effect was comparable in samples of men and samples of women, but it was significantly stronger in samples that combined men and women.

Age: The training effect was 3 times larger in older samples compared with younger samples. This statistically significant difference represented a moderate ES.

Experience: The training effect for experienced lifters was half as large as the effect for novices. The statistically significant difference represented a small ES.

Strength Test: Relative to the bench press, the training effect was significantly stronger for 5 of 10 other strength tests. With the exception of a large effect for triceps strength measures, the significant differences were associated with small ES.

Periodization: The training effect in periodized programs was roughly half the effect in progressive programs. The statistically significant difference represented a small ES.

Sessions: The training programs with 2 sessions per week produced a weaker training effect than programs with 3 sessions per week, but the difference was not statistically significant.

Table 2. Effects of Moderator Variables

	Odds	Z[a]	Sig.	OR	χ^2	Sig.	ES[b]	Increase[c]
Gender								
Men	1.55	5.22	<.001	--[d]				1.06
Women	2.12	6.23	<.001	1.37	2.44	.118	.17	1.24
Mixed	5.17	10.16	<.001	3.34	29.37	<.001	.67	1.17
Age								
Young	1.64	7.10	<.001	--[d]				1.08
Old	5.99	11.03	<.001	3.04	31.96	<.001	.61	1.28
Experience								
Novice	2.68	13.01	<.001	--[d]				1.17
Experienced	1.33	2.03	.042	.50	15.01	<.001	-.38	1.02
Strength test								
Biceps curl	2.88	5.35	<.001	2.17	5.60	.018	.43	1.13
Bench	1.32	2.73	.006	--[d]				1.05
Chest Press	4.43	3.64	<.001	3.35	7.86	.006	.67	1.22
Squat	1.55	2.13	.034	1.17	.32	.573	.09	1.04
Knee extension	2.68	4.84	<.001	2.03	5.75	.016	.39	1.16
Lat pull-down	3.00	2.86	.004	2.27	3.22	.073	.45	1.13
Leg curl	2.31	3.26	.001	1.74	2.30	.129	.31	1.18
Leg press	2.83	5.81	<.001	2.14	6.93	.008	.42	1.21
Military press	2.17	2.44	.015	1.64	.92	.339	.27	1.16
Triceps	7.00	3.91	<.001	5.29	10.54	.001	.92	1.31
Other	2.31	5.33	<.001	1.74	3.86	.050	.31	
Periodization								
Progressive	2.63	11.99	<.001	--[d]				1.16
Periodized	1.46	4.19	<.001	.55	11.54	<.001	-.33	1.10
Sessions/week								
2 sessions	2.00	5.56	<.001	.77	1.99	.158	-.14	1.11
3 sessions	2.59	11.20	<.001	--[d]				1.17
Sets/session								
1 set	1.77	2.76	.006	.52	6.16	.013	.36	1.11
3 sets	3.38	11.48	<.001	--[d]				1.17
Repetitions/set								
≤7 per set	1.58	3.15	.002	.44	10.64	.001	-.45	1.09
8-10 per set	3.62	10.29	<.001	--[d]				1.18
11+ per set	2.82	5.20	<.001	.78	.62	.431	-.14	1.22

[a]From Wilcoxon rank test.
[b]ES = ln(OR)/1.81 (Chinn, 2000).
[c]Increase = Weighted Average of (Posttest Standard Deviation/Pretest Standard Deviation).
[d]Reference group.

Table 3. Recreational Lifters Versus Competitive Lifters

	-[a]	+[a]	0[a]	*z*	Sig.	Odds	OR	ES	χ^2
All young lifters									
Recreational lifters	43	65	0	-2.02	.043	1.51	–[b]		
Competitive lifters	31	35	1	-.37	.712	1.13	.75	.86	.355
Young men only									
Recreational lifters	34	45	0	-1.13	.261	1.32	–[b]		
Competitive lifters	31	31	1	0.00	1.000	1.00	.76	-.15	.410

[a]"-"indicates a negative even (i.e., Final SD < Initial SD), "+" indicates a positive event (i.e., Final SD > Initial SD), and "0" indicates a neutral event (i.e., Final SD = Initial SD). A positive effect was one that was consistent with the hypothesized training effect. Thus, observations coded "+" indicate the presence of a training effect as defined here.
[b]Reference group for ORs within the same set.

Sets: Programs with 1 set per session produced roughly half the training effect seen in programs with 3 sets per session. The difference was significant, but the ES was small.

Repetitions: The training effect for programs with ≤7 repetitions per set was less than half the effect seen in programs with 8-10 repetitions per set. The difference represented a small, but statistically significant ES. The training effect for programs with 11+ repetitions did not differ significantly from that for programs with 8-10 repetitions.

The last column of Table 2 expresses the training effect as the geometric mean of the posttraining/pretraining ratio of the standard deviations. The fact that this ratio was >1.00 in every case is further evidence of the generality of the training effect. The wide range of ratios (1.02–1.31) indicates that the magnitude of the effect was influenced by who was trained, how they were trained, and which tests were used to measure how much their strength changed during training.

Further Examination of Experience Effect

The experience effect noted in Table 2 was examined in greater detail to determine whether the extent and nature of the lifter's experience affected the magnitude of the training effect. If experience is a continuum, the effect should be stronger in those who have trained longer. Assuming that competitive lifters have a longer training history than recreational lifters, the analysis produced a nonsignificant trend in this direction (Table 3). The most interesting finding in this analysis was that the training effect disappeared completely in young men lifting competitively.

Effect of Program Length

Training effects might be expected to cumulate over time. If so, the impact of resistance training on strength test score variability should be larger in longer programs. The correlation of program length with training effect was computed to test this hypothesis. Separate correlations were computed for different demographic groups to avoid confounding duration effects with the influence of gender, age, or training history. The average correlation in these analyses was r = .063 for program length, a trivial ES in Cohen's (1988) classification. While the raw correlations clearly varied about the average, the variation was too small to indicate significant differences. Variation about the average correlation was not related to gender, age, or training history (χ^2 = .55, 7 df, p > .999).

Discussion

This study was undertaken to evaluate the hypothesis that resistance training increases the variability of strength test scores. The results presented in this report lead to two general observations regarding this hypothesis. First, the hypothesized training effect does exist. Second, the magnitude of the effect depends on trainee characteristics and details of the training program.

Resistance training increases the variability of strength test performance. The odds that the posttraining standard deviations for a strength test scores would be larger than the pretest standard deviation were 2.5:1 for samples that underwent a resistance training program. The odds dropped to 1:1 in control groups. The inference that training increases the variability of strength test scores follows from this difference.

The training effect depended on trainee characteristics and how they were trained. With regard to trainee characteristics, the training effect was much stronger for older people than for younger people. The training effect was much weaker for experienced lifters than for novice lifters. Closer examination suggested that the training effect might disappear all together among competitive lifters. The gender comparison was atypical because the training effect was about equally strong for men and women. With regard to how they were trained, the training effect was weaker in periodized programs than in progressive programs, in programs with 1 set per session than in programs with 3 sets per session, and in programs with ≤7 repetitions per set compared with programs with 8+ repetitions per set. These differences generally were modest; the ORs converted to small ESs.

The training effect was robust despite its dependence on who was trained and how they were trained. The effect was evident to some degree in every demographic group and every type of training program. The moderator effects described in the previous paragraph only represent variation in the strength of the effect.

Longer programs did not produce larger training effects. This result could be evidence that the training effect occurs early in training and then remains constant. If so, it would be reasonable to suggest that neuromuscular adaptations that are known to occur early in resistance training are the basis for the training effect. This explanation also could account for the impact of experience on the training effect. Experienced lifters would be expected to have developed the requisite neuromuscular mechanisms for lifting as part of their prior training. This expectation would apply with special force for competitive lifters assuming they have a long history of intense training. Therefore, a neuromuscular adaptation effect could account for two important trends in the results. A restriction of range for program duration would be another possible explanation because the evidence is limited to relatively brief programs. This restriction will reduce the strength of correlations between program length and other variables (Sackett & Yang, 2000). However, the observed program length-training effect correlations were so weak that correcting for the restriction of range would have little effect.

The practical importance of the training effect described in this paper is uncertain. The effect will not change conclusions about the overall effectiveness of resistance training programs. The change in the average score on strength tests translates into ESs that typically fall between 1.00 and 2.00. Sometimes these effect sizes are computed using the pretraining standard deviation. Using the posttraining standard deviation in the computations would reduce ES by 23% in the worst case (i.e., a 30% increase in the standard deviation during training). The revised range of ESs describing the increase in strength would be 0.77 to 1.54. Strength gains of this magnitude would be statistically significant in most cases. The actual shrinkage generally would be less than this worst-case scenario. Taking a 15% increase in test score variation as a more representative value, the shrinkage in ES would be 13%. Applied to the typical ES values for resistance training programs, this shrinkage would so the revised range of ESs would be .87 to 1.74. With reasonable sample sizes, studies of effects this large will have substantial statistical power, so it is unlikely that the null hypothesis will be incorrectly retained in any individual study.

The training effect could be important for other applications. The training effect indicates that training increases the spread of test scores. Accurate representation of the spread of the scores will be important when determining the ability to meet some absolute criterion. For example, the question might be “How many people will meet a minimum job strength criterion after training?” The job strength would set a standard that divided job applicants into those who were qualified and those who were not. If the training goal is to increase the number of qualifiers, the cases of interest would be found in the lower tail of the test score distribution. Fewer people will fall below the criterion if the relatively narrow spread of scores implied by the pretraining standard deviation is used to estimate the failure rate than if the broader spread of scores implied by the posttraining standard deviation is used. The same problem arises if a high strength criterion is set to determine who is qualified for an exceptionally demanding job. However, in this case, the use of the pretraining standard deviation will result in an underestimation of the number of qualified applicants. Either type of error could be important in some evaluations of resistance-training programs.

The training effect could also be important in matching people to resistance-training programs that will maximize their outcomes. A typical resistance-training program produces large changes in average strength test scores. The this change is so great that the evidence implies that nearly everyone increases his or her strength during training. Suppose, for the purposes of discussion, that a training program is carried out that increases the strength test scores for every participant. If none of the participants are negatively affected by training (i.e., no one gets weaker), the standard deviation of strength test scores could only increase if some program participants improved more than others. Understanding the reasons for differential responses to the training program could provide insights into how to revise the program to improve its overall benefits. Understanding the reasons for differential responses to the training program might also suggest guidelines for assigning some people to different programs that would benefit them more. This latter possibility would be even more important if the large average increase in strength test scores hides the fact that some people either do not benefit from training or even lose strength as a consequence of training period. Large average strength gains do not rule out these possibilities. Average strength would improve if the individuals who do not benefit represent only a small part of the program participants and/or their negative responses to training were small relative to the gains made by other subjects. In either case, the increased standard deviation indicates the presence of individual differences in response to the training program. Understanding the basis or bases for those differences could be important for matching people to training programs to maximize the benefits of training.

Every study has strengths and weaknesses. In this case, the strengths include broad coverage of training populations. Differences in age, gender, and training experience have been examined. A wide range of strength tests has been covered. The brevity of most training programs is a weakness because it limits the opportunity to draw stronger conclusions about the effect of program length. Also, the reported statistical significance levels should be viewed with caution. Each test administered to a sample was treated as an independent observation. These observations clearly are not independent, so the sample size for the significance tests has been overstated to some extent. Overstating the sample size will result in overestimating the extremity of the p values for significance tests. However, the important effects were highly significant ($p < .001$), so of the statistically significant trends identified in this report probably would remain significant even with a substantial reduction in the degrees of freedom.

In summary, resistance training increases the variability of strength test scores. This training effect is a robust phenomenon even though the magnitude of this effect depends on who is being trained and how they are trained. The training effect may have practical importance for some types of program evaluations, but it will be unimportant for others. As a consistent phenomenon, some consideration should be given to the expected increase in posttraining test score variation when evaluating the effects of resistance training studies.

References

Chinn S. (2000). A simple method for converting an odds ratio to effect size for the use in meta-analysis. *Stat Med* 19:3127-31.

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

Dean AG, Sullivan KM, Soe MM. OpenEpi: Open Source Epidemiologic Statistics for Public Health, Version 2.3. www.OpenEpi.com, updated 2009/20/05, accessed 2009/06/04.

Falk B, Tenenbaum G. 1996. The effectiveness of resistance training in children. A meta-analysis. *Sports Med* 22:176-86.

Payne VG, Morrow JR, Jr., Johnson L, Dalton SN. 1997. Resistance training in children and youth: a meta-analysis. *Res Q Exerc Sport* 68:80-8.

Peterson MD, Rhea MR, Alvar BA. 2004. Maximizing strength development in athletes: a meta-analysis to determine the dose-response relationship. *J Strength Cond Res* 18:377-82

Raudenbush SW, Bryk AS 2002. Hierarchical linear models (2nd ed.). Thousand Oaks, CA: Sage Publications.

Rhea MR, Alderman BL. 2004. A meta-analysis of periodized versus nonperiodized strength and power training programs. *Res Q Exerc Sport* 75: 413-22.

Rhea MR, Alvar BA, Burkett LN. 2002. Single versus multiple sets for strength: a meta-analysis to address the controversy. *Res Q Exerc Sport* 73: 485-8.

Rhea MR, Alvar BA, Burkett LN, Ball SD. 2003. A meta-analysis to determine the dose response for strength development. *Med Sci Sports Exerc* 35:456-64.

Sackett PR, Yang H. 2000. Correction for restriction of range: an expanded typology. *J Appl Psychol* 85:112-8.

Siegel S. (1956). *Nonparametric statistics*. NY: Wiley.

Vickers RR Jr., Hervig LK, Barnard, A. Program Length as a Factor in the Effectiveness of Resistance Training: A meta-analysis. Manuscript in preparation

Wolfe BL, LeMura LM, Cole PJ. 2004. Quantitative analysis of single- vs. multiple-set programs in resistance training. *J Strength Cond Res* 18:35-47.

REPORT DOCUMENTATION PAGE

1. Report Date (DD MM YY)
08 06 09

2. Report Type
Technical Report

3. DATES COVERED (from - to)
Jan 09 – Jun 09

4. TITLE AND SUBTITLE
Resistance Training Increases the Variability of Strength Test Scores

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

6. AUTHORS
Vickers, Ross. R., Jr.; Barnard, Amanda C.; Hervig, Linda K.

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

8. PERFORMING ORGANIZATION REPORT NUMBER
09-29

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

Commander
Navy Medicine Support Command
P.O. Box 140
Jacksonville, FL 32213-0140

10. Sponsor/Monitor's Acronyms(s)
NMRC/NMSC

11. Sponsor/Monitor's Report Number(s)

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT (maximum 200 words)

Informal observations made while conducting a meta-analysis of resistance training programs suggested that the between-person variation in strength test scores is greater after training than before. This study treated the informal observation as a hypothesis to be evaluated. The odds were 2.5:1 that the standard deviation after training would be larger than the standard deviation before training if a sample underwent resistance training, compared with 1:1 odds for control groups in training studies. This difference supported the study hypothesis. Extending the analysis to subcategories based on age, gender, training experience, and training program characteristics, the training effect was present in all subgroups, but it was significantly stronger in some than others (e.g., older or novice lifters). The training effect did not increase with program length. This fact and the weak effect in experienced lifters suggested that the training effect might be a product of neuromuscular adaptations occurring early in training. Whatever its source, the training effect seldom will change inferences about whether resistance training has increased the strength of program participants, but it could be important when predicting how training affects the ability to meet basic performance standards.

14. SUBJECT TERMS
resistance training, strength tests, score variation, meta-analysis

16. SECURITY CLASSIFICATION OF: a. REPORT	b.ABSTRACT	b. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	18a. NAME OF RESPONSIBLE PERSON
UNCL	UNCL	UNCL	UNCL	12	Commanding Officer

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