CYBEX®
Serious about fitness.

The Truth on Fitness:
To Prove or Not to Prove

Paul M. Juris, Ed.D.
Executive Director, CYBEX Institute

The Truth on Fitness: To Prove or Not to Prove

When it comes to matters of finance, cash is king. This simple, yet powerful, phrase suggests that with all of the creative mechanisms for counting assets, the only thing with true intrinsic value is cash in hand.

When it comes to matters of human functioning or performance, the cash equivalent is evidence. In other words, despite the prevailing trends, the relentless dogma, or the protestations of industry "experts" regarding what one should or should not do, the only real currency is that which can be supported through authentic scientific inquest. Everything else is pure unsubstantiated opinion.

On the other side of the evidentiary coin, however, are those who assume that evidence is synonymous with proof. This is also misguided. Proof, in its purest form, really only exists in mathematics, in which there can be only one irrefutable solution to a problem. We could attempt to reach absolute certainty when determining the effects of a research application, but then we would have to measure the entire population of the world. Obviously, aiming for that kind of proof would make any research impossible.

Instead, we gather smaller, representative samples of the population, and use statistics to make inferences about the effects of various treatments. Thus, research science is based on probability, not absolutes. What is the probability that any differences between two groups are due to the different treatments that we apply to those groups, and not to the natural variability within the groups, or to chance alone?

As scientists, we accept no more than a 5% probability that research outcomes are due to variability or chance. Hence, the "p" statistic, written as $p < .05$. If we wish to apply a stricter criterion, we would accept no more than a 1% probability that outcomes are the result of chance ($p < .01$). In either case, we don't prove anything; instead we demonstrate that the outcomes of our experiments are probably not the results of chance (1).

There are some who argue that research is either unnecessary, or to the other extreme, excessive, and that managing variability within sample groups leads to unrealistic outcomes. Perhaps these people don't understand basic statistical concepts, or even more critically, are concerned that the outcomes of any inquiries will fail to support their arguments. Experimental design, after all, is really not that difficult.

Then there are those who continue to espouse a particular point of view, through their research efforts, despite a complete dearth of supporting evidence. One such instance is a recent study by Richards and Dawson (2). The authors postulated that resistance exercises using novel, multi-planar movement arcs would yield greater improvements in strength than similar exercises using stereotypical, uni-planar movements.

Fourteen women from a population of lower extremity power athletes (sprinters, jumpers, etc.) were randomly placed into one of two groups, engaging in six weeks of shoulder abduction and flexion exercises with elastic bands. The first group performed traditional movements, either pure flexion or abduction, within a fixed movement plane. The second group performed the flexion and extension movements, but this time scribing both vertical and horizontal figure-eights rather than the common up and down movements of group one. Both groups were tested for maximum strength, before and after training.

The post-training results demonstrated that both groups experienced significant increases in maximum strength over the six-week training interval ($p < .01$). The data, however, failed to show a significant difference in strength gains between the groups. In other words, the multi-planar movement arcs were no more effective in producing strength gains than the simple, uni-planar arcs.

The authors went on to report that "a consistent trend suggested that the multi-planar group demonstrated a greater percentage improvement in strength than the uni-planar group (116.0% vs. 113.5%)," and that the variable movement scheme produced superior outcomes. The problem, however, is that there is no statistical basis, nor any evidence, in support of this argument, and the authors failed to provide any raw data that would allow someone to draw a similar inference.

The fact is that the recorded difference of 2.5% between the groups was more likely due to the variability within each of these groups than it was to the exercise treatments themselves, particularly when the overall improvement of each group was over 100%. Furthermore, a single insignificant difference between groups does not, in any way, represent a trend. Had the experimenters repeated the study several times, with similar results, then they could have claimed the presence of a trend, despite the lack of statistical significance. Even if they had done so, it wasn't reported in this study.

Very simply stated, the outcomes from each group were the same. But by presenting the data in the manner of their choosing, the authors revealed that they were anticipating, and maybe hoping for, a different outcome. Ironically, Richards and Dawson argued that despite a trend revealing potential differences between the groups at the outset of the study, there was no statistical significance between them, and therefore, they could be considered equal. This finding was absolutely necessary to establish a common starting point for the groups, for differences at the outset would have explained any potential differences in the training effects.

It is perplexing, therefore, that the authors reversed this logic in trying to support their argument of training superiority for the variable movement scheme.

Not all researchers elect to ignore the findings of their own research, in favor of preferred outcomes. Some, rather interestingly, completely fabricate information which they did not study in the first place. Take, for example, the study by Oliver and Dougherty (3), which posed a rather interesting question; what is the difference in levels of muscle activation, if any, as measured by electromyography (EMG), between a traditional, prone hamstring curl, and the more contemporary razor curl?

This study, in contrast to the formerly described paper, involved eight subjects measured against themselves, under two conditions, in what is commonly referred to as a repeated measures design. This is necessary in an EMG study, because these signals cannot be compared across subjects, but are effective in evaluating the same individuals performing different tasks.

The authors examined peak muscle activation in the Biceps Femoris, Medial Hamstrings, Gluteus Maximus, and Gluteus Medius muscles. The results of their analysis revealed no significant differences in activation levels among any of the muscle groups, as a result of the two exercises. They do report that the total combined EMG magnitude of the four muscle groups was greater for the razor curl, but they don't volunteer whether that value was statistically significant. As far as one can discern from their findings, therefore, both exercises were equally effective at stimulating muscle activity.

Once again, however, it appears as though the authors were disappointed with the results, because they continued, in their discussion, to conclude that the razor curl would be more effective in promoting functional outcomes, despite having no greater impact on muscle stimulation, and therefore, is the preferable modality.

This may, in fact, be true. But the problem is that this particular study was limited to an evaluation of muscle contraction, and did not measure functional outcomes. The authors, therefore, had no basis for making such claims. Thus, any attempt by the authors to argue in favor of the razor curl, because of its functional superiority, is pure conjecture. If Oliver and Dougherty wanted to assess the functional differences between prone and razor curls, then they should have designed an experiment to evaluate function, rather than using EMG as a tool to provide evidentiary support for personal biases.

The truth is, that in the hands of skilled statisticians, data can be made to state virtually anything. On the other hand, the true purpose of research is to ask legitimate questions for which we, hopefully, obtain accurate, fair, and unbiased answers. In this way, we are in a position to learn the truth, advance our knowledge and understanding of human functioning, and hopefully, make a positive impact on those with whom we work.

Very often, however, people disguise editorial in a cloak of research, hoping to use scientific methodology to support their personal attitudes and beliefs. This serves only to propagate opinions and myths, and keep us mired in an endless cycle of populist dogma. In this event, it is favorable to side with those practitioners who question the utility of research in the first place.

References

Steele, R.G.D. and Torrie, J.H. (1980). *Principles and Procedures of Statistics: A Biometrical Approach.* 2nd Ed. McGraw-Hill Book Company. New York.

Richards, J.A. and Dawson, T.A. (2009). Optimizing exercise outcomes: the efficacy of resistance training using conventional vs. novel movement arcs. *Journal of Strength and Conditioning Research.* 23(7): 2015-2024.

Oliver, G.D. and Dougherty, C.P. (2009). Comparison of hamstring and gluteus muscles electromyographic activity while performing the razor curl vs. the traditional prone hamstring curl. *Journal of Strength and Conditioning Research.* 23(8): 2250-2255.

WORLD HEADQUARTERS
10 Trotter Drive • Medway • MA 02053 USA • T +1.508.533.4300 • F +1.508.533.5500

CYBEX INTERNATIONAL UK LTD
Oak Tree House • Atherstone Road • Measham • Derbyshire • DE12 7EL UK
T +44.845.606.0228 • F +44.845.606.0227

www.cybexintl.com