TRAINERS JOURNAL
SECTION
The NATIONAL ATHLETIC TRAINERS ASSOCIATION

Tests of Body Function as Measures of Progress in a Physical Training Program

By W. W. Tuttle, Ph.D.
Department of Physiology, University of Iowa

CHANGING the so-called "soft" American youth into a tough performer, whether in sports or battle, has been efficiently and quickly done. Success in this endeavor is due to the fact that American youth was not as "soft" as some may have thought, and to the knowledge which trainers had acquired during years of experience, as to the requirement of an adequate training program. That our training programs have worked well can not be denied since the accomplishments of the participants have been noteworthy, even spectacular, all over the world.

One of the problems confronting administrators of training programs is to show periodically, the progress which is being made. They may wish to show what improvement, if any, is being made in physical capacity and condition, as well as in athletic performance. In order to do this, various tests are employed. When tests are correctly used, and when their limitations are recognized, they can be of great value. Only in the light of the limitations of a test can the results secured by it be interpreted fairly. If this is not done, unwarranted conclusions may be drawn, and the whole training program discredited or overrated. It is the purpose here to consider some of the difficulties that may be encountered in using and interpreting certain tests of body function.

In the first place, let us all be suspicious of spectacular results. When such results are proclaimed, one can be assured that they will be checked and rechecked by a multitude of investigators. No one can afford to stake his reputation on spectacular findings, until he is quite certain that they can stand up under cross-examination. Large amounts of money have been spent, and trainers have proceeded on false assumptions through no fault of their own, simply because careless researchers have promulgated false ideas.

In the second place, we should not claim a cause-and-effect relationship, unless a control experiment fails to show the result in question. It is generally recognized, for example, that vital capacity may be increased as a result of certain types of physical training. It is not advisable, however, to use vital capacity scores to demonstrate the effectiveness of a training program. The reason for this is that, by practicing deep breathing, the vital capacity of a control group not participating in any other type of exercise also becomes progressively greater. A similar situation exists in riding a bicycle ergometer. By doing no other exercise than riding the bicycle, one experiences a progressive increase in capacity for this work. There are several examples in literature in which a progressive increase in the capacity to ride a bicycle was attributed to other factors, when in reality the cause was nothing more than repetition in bicycle riding.

It is never safe to conclude that, by adding something to a diet, capacity to do work is phenomenally improved, unless a matched group is run parallel to the experimental group, as a control. If the control group fails to show the improvement, and the experimental group does show it, then perhaps, the dietary factor is responsible. If the experiment can be repeated with similar results, this is additional evidence of the authenticity of the findings. And when it can be universally repeated with similar findings, then the conclusions can be stated without reservations.

It was reported to me recently that, by adding accessories to the diet of athletes, a phenomenal health record was set. Knowing the general physiologic effects of the food elements concerned, I was prone to reject the idea that the alterations in diet were responsible for the health record. In

this instance the effects were apparently there but it is doubtful that they were due to the new dietary regime. The person who made such a claim has no proof since the findings in his experimental group were not checked against a control group. How did the investigator know what health records would be established by the experimental group during the period in question, had the diet accessories been omitted? In this instance, the question can never be answered. The available experimental evidence, however, leads us to believe that in this case the accessories used had nothing to do with either health or the capacity to do work.

The effect of practice may be evident in some tests of strength as well as in bicycle riding. For example, in testing a group of twenty-five college women for leg strength by means of a belt dynamometer, there was a marked improvement from week to week over a period of three weeks. The group average for the first set of scores was 570 pounds. The second week the average was 673 pounds, an increase of 103 pounds. Complaint concerning the way the belt fit the hips prompted the securing of another belt before collecting the third set of scores. The third weekly test gave a group average of 837, an increase of 164 pounds. During this period, other strength scores (grips, push, pull, and back strength) changed but little. Had these girls been indulging in a program for strengthening leg muscles, and had this particular leg strength test been used as an index of improvement, the conclusion might have been that the program was responsible for a progressive increase in leg strength. However, in this instance, the group was indulging in no training program of any kind. In the light of this fact and the result of the other strength tests, one was forced to believe that some factor other than an increase in leg strength was operating. No doubt, part of the increase in the leg strength score was due to the fact that by practice one learns how to apply strength to a dynamometer more effectively. Furthermore, in this instance, the first belt tended to hurt the hips when strength was applied to it. The new belt alleviated the difficulty, permitting a greater effort on the part of the girls.

Considerable time and effort have been spent in attempting to develop adequate tests of physical condition. One of the approaches has been through the use of cardiovascular tests. If one is to obtain a complete picture of cardiovascular changes as related to physical condition, every cardiovascular element would have to be measured. To undertake this would be folly because it is either difficult or impossible to measure some of these elements. For this reason, we must resort to sampling the cardiovascular elements, using only those that can be readily measured. Although the sampling attack must be employed, it falls far short of yielding the results that would be obtained if all the items involved in cardiovascular response could be included as test items.

Since heart rate and blood pressure are known to vary with many factors such as age, sex, emotional states, and physical activity, it is a mistake to interpret cardiovascular scores solely in terms of physical condition. A standard exercise may be more strenuous for one individual than for another, because of differences in body size and strength. Obviously then, cardiovascular test scores will not correlate very highly with independent measures of physical condition. This means that cardiovascular tests are limited with respect to validity. It does not mean that such tests are to be discarded as being of little or no value, rather that test scores should be interpreted with the fact of low validity in mind.

Physical condition is a term that is difficult to define because it describes the relationship between the actual and the potential capacity of an individual to do work without injury or undue fatigue. An eight-year-old boy may not be able to do as much work as a mature individual, but both may be in excellent physical condition. The same thing holds for a big man and a little one. Physical condition can not be measured directly in terms of work output; it is work output in relation to what the individual can be expected to do. For this reason, measures or ratings of condition are in themselves fallible and lacking in validity. Therefore, the correlation between cardiovascular test scores and available measures of condition are spuriously low. It is for this reason that certain cardiovascular tests should be retained, even though the scores appear to have little significance from a statistical point of view.

The development of tests of condition, endurance, and physical efficiency, as well as other tests, has involved a rather wide use of certain relatively simple statistical procedures. Unless a person has a good working knowledge of elementary statistics, he is severely handicapped in analyzing test results and in understanding the work that others are doing along this line. For most purposes it is sufficient to know what is meant by correlation, probable error, standard deviation and measures by which differences are evaluated. If one will apply himself for a reasonable length of time to the task of learning about these simple elements of statistics, one will be able to understand quite completely, most scientific reports. This falls within the ability of almost everyone capable of doing research at all, and it is certainly essential for those who wish to comprehend the reports of other investigators.

In general, the intelligent use of tests of body function in the physical training program is to be encouraged. If, on the basis of information gained by testing, the effectiveness of a program can be materially improved, time spent in analyzing the results will have been well spent.