THE PROBLEM OF MEASURING PHYSICAL FITNESS FOR HARD WORK:
A COMPARISON OF A FITNESS TEST WITH A MIXED PERFORMANCE TEST.

By Major J. E. Lovelock, M.A., M.R.C.P.,
Royal Army Medical Corps,
Physical Medicine Specialist.
(From The Army School of Physical Training, Aldershot.)

[Received April 17, 1945.]

For centuries the assessment of physical fitness by simple measurements has been sought. These have comprised anatomical and physiological data in varying degrees of complexity. They began with the anthropometric measurements of the Greek period, but in the present century physiological functional tests have been most widely used.

It is now generally agreed that fitness can only be assessed in terms of some specific activity. Its testing requires a measurement of the physiological aptitude for this activity. This implies some directive qualification of "physical fitness" and necessitates the use of more or less of that activity for its measurement. The most reliable test is obviously a performance in which the physical ability to discharge a task is measured by using that task itself. This is seldom practicable, however, but the results of any accepted test should have a close correlation with the performance it is potentially measuring.

The most accessible physiological measurement is the response of the circulation to effort, most easily and reliably measured as the pulse-rate. Literature abounds in references to this in men in training. During the last decade various aspects are discussed and comprehensive summaries are to be found in such works as those of Steinhaus (1933), Cotton and Dill (1935), Hill, Magee and Major (1937), Abrahams (1939) and Jokl et al. (1941). It is generally agreed that of the various simple tests for physical efficiency the rate of cardiac deceleration after intense effort seems the most reliable.

It is understandable that during wartime, when physical fitness is so important, the search for suitable tests should have continued. So varied are the tasks of the fighting soldier, however, that it is sufficient definition to make the aim the testing of the capacity for any hard and enduring physical work. It is this that the Harvard Pack Test, a recently widely accepted one, sets out to gauge.

In this, the pulse deceleration rate is measured after moderately heavy standard exercise. Briefly, it consists of stepping, every two seconds for five minutes, on to a 16-in. high platform; a pack weighing approximately one-third of the body-weight is carried on the back; the hands, shoulder width apart, are kept grasping a horizontal bar at shoulder height. Immediately at the end of this time, or earlier if the rate of stepping lags for more than twenty seconds, the subject sits while the pack is removed. The pulse is measured in half-minute periods, 1–1½ minutes, 2–2½ minutes and 4–4½ minutes after cessation. The index of fitness is calculated as

\[
\text{Index of Fitness} = \frac{\text{Duration of Exercise in Seconds} \times 100}{\text{Twice the Sum of these half-minute Pulses}}
\]

Figures for poor, average, good and superior fitness in normally healthy men are laid down. At this Army School the test was tried out against a hard mixed performance test; the results of the Fitness Index (F.I.) are compared with the Performance Test (P.T.). Special care was taken to eliminate possible variables by the time allocation of the two tests.
The Problem of Measuring Physical Fitness for Hard Work

**METHOD.**

It seemed that three main groups should be taken for this comparison.

1. Highly trained men in a presumably steady state of fitness. These existed in the persons of the permanent staff instructors of the A.S.P.T. Eleven instructors were used in this group.

2. Moderately trained men whose general fitness for work should improve with further training. These existed in the persons of War Course N.C.O.s training for promotion to the A.P.T.C. On these, in addition to the direct correlation between F.I. and P.T., the variation of both over the last four weeks of their course was charted. In fact, the F.I. was recorded three times as compared with the P.T. twice. The results give an interesting criticism of the Pack Test technique. Twelve men were followed over this one month's trials.

3. Initially unfit men on whom the Pack Test might serve as a prognostic test. These do not exist at the A.S.P.T. With this group it would be important to know: (a) how many men of low F.I. are trainable to A.I. on performance; and (b) how many men of high F.I. become poor performers and subsequently break down in routine unit training.

It is hoped that this may later be done.

**DETAILS OF TESTS.**

The Pack Test was carried out strictly in conformity with the details as laid down by the originators.

The Performance Test was designed to try out to their limits the main systems of the body in all-round trials of physiological value and military interest. It was in three parts with no intervening rest periods. The whole performance was over within one hour. It consisted of a four miles run, a two hundred yards carry of a man of similar weight and equipment, and a double ascent of a vertical 15-foot rope. All were done in full battle order.

Marks were awarded mainly on length of time taken in the different parts. This was regarded as the most important single criterion capable of exact measurement; but note was taken of other less easily recorded objective phenomena, such as breathlessness, technique, and degree of effort.

Subjective symptoms were investigated after both tests. Though these were not taken into account in the figures used, helpful information was obtained.

Considered physiologically, this performance test seemed to try out most body systems to the full. The run, in addition to working the large muscle groups in the legs and back, was regarded as a severe test of the cardiovascular and respiratory systems. The carry embraced these, as well as bringing in the large muscle groups of arms and back in an exercise mainly of strength. The rope climbing was more of a trained skill. In addition to the smaller muscles of hand and forearm, it involved the leg and back muscles in practised movements. It forms a usual part of the training at this school and therefore was capable of measurable improvement. This improvement is a reflection of the training of the central nervous system in the execution of skilled movements.

**RESULTS.**

(1) The results of the first group are shown in Table I.

<table>
<thead>
<tr>
<th>(1) C.S.M.I. E.</th>
<th>Fitness Index (F.I.)</th>
<th>Performance Test (P.T.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>78</td>
</tr>
<tr>
<td>(2) C.S.M.I. S.</td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>(3) C.S.M.I. D.</td>
<td></td>
<td>81</td>
</tr>
<tr>
<td>(4) C.S.M.I. B.</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>(5) C.S.M.I. E.</td>
<td></td>
<td>73</td>
</tr>
<tr>
<td>(6) C.S.M.I. L.</td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>(7) Q.M.S.I. N.</td>
<td></td>
<td>87</td>
</tr>
<tr>
<td>(8) Q.M.S.I. R.</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>(9) Q.M.S.I. McN.</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>(10) Q.M.S.I. C.</td>
<td></td>
<td>85</td>
</tr>
<tr>
<td>(11) Q.M.S.I. S.</td>
<td></td>
<td>83</td>
</tr>
</tbody>
</table>
These figures have been plotted and are best considered graphically (Graph 1).

The scatter of the plotted points round the mean line appears to be fairly high. This indicates only a moderate correlation between the Pack and the Performance Tests. Of the four points above and the four below this line, only one on either side shows a marked discrepancy.

Of these, one (8) had a P.T. of 23, considerably higher than his F.I. of 77 would have led us to expect. This fourth highest score was gained by good all-round performances, especially in the run where he finished first. Inquiry showed that he felt "off colour" in the morning's Pack Test and could not concentrate on the timing. He felt equally tired during the afternoon performance but, as he was keen on doing well in the test, scored highly in all three parts. Here, determination conquered indisposition.

The other main divergent (9) had a poor P.T. of 23 compared with his F.I. of 88, the highest registered. He did poorly in the run but better in the carry and ropes. Thus, his endurance work was poorer than his accomplishment of skilled movement.

Of the remaining nine, little comment is needed. Only two, one at either end, merit special note. These are (5) and (7). Both fit well into the line of the graph. The former, just returned from leave, said he was far from fit; the latter thought himself very fit, and both types of test came equally easy to him. Both thought the Pack Test fairly represented their fitness for performance.

(2) The results of the second group of men, extended over a period of one month, are best considered under four heads. These are:

(a) The correlations between the Fitness Index and the Performance Test (i) at the beginning, and (ii) at the end of the month, as already similarly done for the first group.

(b) The changes in value obtained for the F.I. during this month.

(c) The changes in value obtained for the P.T. over this period.

(d) The relation between (b) and (c).

The figures obtained and used in the relevant graphs (2 to 5) are shown in Table II.
The Problem of Measuring Physical Fitness for Hard Work

<table>
<thead>
<tr>
<th>TABLE II.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 24-25</td>
</tr>
<tr>
<td>F.I.</td>
</tr>
<tr>
<td>(1) Cpl. S.</td>
</tr>
<tr>
<td>(2) Bdr. G...</td>
</tr>
<tr>
<td>(3) Cpl. B...</td>
</tr>
<tr>
<td>(4) Cpl. C...</td>
</tr>
<tr>
<td>(5) Cpl. W...</td>
</tr>
<tr>
<td>(6) Bdr. B...</td>
</tr>
<tr>
<td>(7) Bdr. D...</td>
</tr>
<tr>
<td>(8) Cpl. T...</td>
</tr>
<tr>
<td>(9) L/B. B...</td>
</tr>
<tr>
<td>(10) Cpl. B...</td>
</tr>
<tr>
<td>(11) Sjt. C...</td>
</tr>
<tr>
<td>(12) Sjt. C...</td>
</tr>
</tbody>
</table>

(a) The correlations between the F.I. and P.T. on the two dates are shown in Graphs 2 and 3.

(i) Graph 2 shows approximately the same scatter as did the first group (Graph 1). The only two meriting comment are (7) and (10).

The former, with F.I. of 88, would have been expected to give a higher P.T. than 14. That he was short of endurance work was shown by his low score being due mainly to poor time in the four-mile run. In this he was the slowest. In the other two tests he was up to standard.

The latter (10) had a lower F.I. of 87 with an abnormally high P.T. of 22. He was older and stronger, and did well on the endurance work, being placed second equal in the run.

(ii) The second recording (Graph 3) shows a wider scatter. Here, too, the six main divergents might be considered.

One (3) had a lower P.T. of 14 than would be expected from his F.I. of 85. In the latter he felt "better than ever before" but in the former he scored badly in all three parts and was obviously tired. The Pack Test did not try him out enough to show this up. His rather anomalous case is further considered under (6) (i) and (ii).

Another (8) had the highest F.I. of 95, but was fourth lowest in the P.T. at 19. During the morning Pack Test his exuberant fitness had been especially noted. During the afternoon...
performance he looked weary and ill. Inquiry showed that he had strained an abdominal muscle two days previously. Though this had not inconvenienced him in the Pack Test, it had handicapped him in the performances. Yet he considered himself much fitter than a month before. In his case, the Pack Test was a good trial of his cardiovascular fitness for hard work, but it did not show the effect of the localized lesion.

A third (6) had a high P.T. of 23 compared with his rather low F.I. of 82. He thought his Pack Test was very easy, as he was obviously not concentrating on it, and did correspondingly poorly. On the other hand, though he found the performance “fairly hard and tough,” he did well in all parts, especially endurance work. He was a top-scorer in the run. Over the last month his fitness had improved greatly by performance, but not by the test result.

Three others, (2), (7) and (10), diverged widely, but investigation of results and full inquiries threw no light on the causes. To (10) the same remarks apply as were made in the previous section ( (a) (i) ).

(b) The Pack Test was carried out not only at the beginning and at the end of the four-week period in conjunction with the Performance Test, but also intermediate at two weeks without any performances for comparison. The results are shown in Graph 4 and will be discussed in their two main parts.

(i) Over the first two weeks one very significant fact stands out. With the one exception of (7), the second readings show either no change, as in (5) and (1), or else a definite decrease. In the first recordings of the Pack Test it was noted that, in spite of being told to straighten knees and back fully, there was no uniformity in this. When the test was repeated a fortnight later, orders were given that the position of attention was to be taken each time on the top of the bench. This makes the test more severe, and it is believed that this tightening of the technique was the main factor in the fall of the readings.

The two cases with the most marked decrease are (3) and (12). They will be considered separately.

It was particularly noted on the first test that (12) did not straighten his knees fully. His F.I. fell from 98 to 88 in the fortnight. Even after this, in the next fortnight he showed a further slight drop to 87, concomitant with a little further improvement in his straightening.

The drop of F.I. from 88 to 61 in the case of (3) was mainly due to his inability to keep to the required pace the second time; as he was too exhausted to keep the correct step, he was
halted after four minutes instead of the usual five. On inquiry he had been feeling “off colour for some days.” Physical examination showed nothing in any system. His recording was a good indication of his fitness that day for hard physical work.

The others all thought themselves fitter than the fortnight before and presumed they had improved their test results. Their decreased readings almost certainly reflected the tightening of the technique.

(ii) The third reading of the F.I. showed in most cases a marked increase. It is believed that the second half of the graph is a more accurate indication of their true F.I. progress. The first half might well be ignored when comparing with performance.

The most striking increase was recorded by (3) who rose from 61 to 85. He was once again feeling fit and, as far as this test could show, had recovered from his temporary indisposition (cf. (a) (ii) and (c)).

One, (7), showed no change, his F.I. remaining level at 92. On inquiry he “felt about the same” in his last two Pack Tests; but his body straightening had improved all through.

The only marked fall in this fortnight was shown by (9). His variation in F.I. through the month was only 87, 86 and 84. Though he maintained that he was fitter and stronger, he could not understand why the Pack Test felt harder finally than initially. His leg and back straightening had improved as his test score dropped.

(c) Only two recordings were made of the more time-consuming performance tests, at the beginning and at the end of the month’s training. The results are shown in Graph 5.

Of the twelve men, six showed variable but definite improvement, four were stationary, while two declined.

The first followed the expected course. Their improvement was not only objective by the test, but also subjective by their feelings; they all felt fitter and performed more easily.

The stationary four had nothing subjective to add to their recorded results. They thought themselves little changed by the month’s training.

Of the remaining two, one (3) “felt tired on the run” and “had no strength in the carry or ropes.” He was definitely “not as fit as a month ago.” It is probable that this was the aftermath of the indisposition mentioned earlier ( (a) (ii) and (b) (i) ) and recorded in Graph 4. It is well shown in the change of his position from Graph 2 to Graph 3.
The other, (8), dropped from the original high recording of 23 to the fourth lowest of 19. This was accounted for by the abdominal muscle strain already mentioned (a). When the Fitness Indices and Performance Tests (Graphs 4 and 5) are considered together in the light of the previous remarks, particularly remembering that only the second half of Graph IV has for the majority any real significance, the correlation between the two is moderately close. Of the twelve, six may be said to show good correlation, two fair and three poor. That of the twelfth, (3), is a little equivocal but after all allowances are made might best be included as fair. This case has already been discussed, especially under (a) (ii).

The first eight call for no particular comment. It was the result to be expected. Except in the cases of (2) and (12), where the agreement is only fair, there is a striking parallelism between the two sets of lines. This denotes good agreement in the fitness progress as measured by the two types of tests. The three main divergents, however, merit closer consideration.

Though (7) remained level at 92 in the latter half of his F.I. graph, his P.T. rose steeply from 14 to 19. He himself noticed little difference in difficulty in the Pack Test but felt much easier in his second performance than in the first. Especially in his endurance work he considered himself much fitter. The failure of the Pack Test to show this adequately was probably due to the tightening of the test technique (cf. (b) (ii)).

The other, (8), has already been discussed under (a) (ii) and (c). His second performance suffered from the discomfort of his abdominal muscle strain. The Pack Test did not show this in his F.I. results.

In spite of having the most marked rise of all from 16 to 23 in his P.T., (9) showed a slight drop in F.I. from 87 to 84. He thought the Pack Test harder the last time than previously, and it is considered that in the two earlier tests he was not straightening fully. On the other hand, he felt easier in the second performance, especially in the run. In civil life he was a professional footballer. He was always short of endurance work and was very tired at the end of a long, hard game. In this he admitted great improvement after the month's training. His discrepancy appeared due mainly to his improved technique in the test, but also a little to its not being hard enough to show his increased endurance (cf. (b) (ii)).

**Summary of Results.**

(1) A fair correlation existed between the Fitness Index and the Performance Test. In the small series of the eleven men chosen for their known all-round physical fitness, only two showed marked discrepancies. Failure of the Pack Test to forecast lack of endurance or to show interest and concentration seemed the main factors in these.

(2) (a) In a series of twelve men of moderate fitness, ratherless correlation was shown to exist, especially when the double test was repeated after a month's intensive training. At the beginning only three showed marked divergence; endurance seemed the main factor in two of these. At the end, six gave results well away from the mean. To the physical and mental factors, already shown as its main deficiencies, might be added the failure of the Pack Test to disclose a local lesion.

(c) The Performance Test results were in keeping both with the known fitness of the subjects and with their feelings. They gave a good basis for comparison, and showed up the full effects on all-round performance of such things as a localized lesion and a previous indisposition.

(d) When the progress of the Fitness and of the Performance Tests was followed over one month's training and a comparison made, 50 per cent showed good correlation, 25 per cent fair and 25 per cent poor. In two of the last three, one factor which did not seem to be adequately recorded by the Pack Test was endurance. This was better tested as the technique
was tightened. The third was affected by the local abdominal lesion. Though the numbers taken are not enough to be significant statistically, they throw some doubt on the reliability of this test.

(3) Until a series of initially unfit men can be similarly tested over a long period, the prognostic value of the Pack Test must be uncertain.

CONCLUSIONS.

The Harvard Pack Test is of limited value in the testing of physical fitness for hard work. *Stringently applied* it is a good test for the response of the cardiovascular system to work of a not too exhausting nature. It cannot be regarded as a sufficiently hard test for real endurance work for fit men. Its prognostic value is uncertain.

Its discrepancies with performance appear to be caused mainly by the failure to attain technical perfection in its execution. This latter requires careful observation even by a trained observer. Only then is it hard enough to test endurance fully.

It is possibly the best cardiovascular test yet developed, but it is very debatable whether this system, though the most easily measured and recorded, is really the most important in assessing actual or potential fitness. It is certainly not the main limiting factor in training healthy men, where the central nervous system, the endocrine system and the mental volitional factors are more important. A reliable fitness test should include their assessment. It has always been difficult to find any test other than hard performance for such a specific thing as fitness.

ACKNOWLEDGMENTS.

I wish to thank the Commandant, A.S.P.T., for permission both to carry out the tests and to forward this paper, as well as those members of the Staff and the Corps who co-operated so willingly.

REFERENCES.