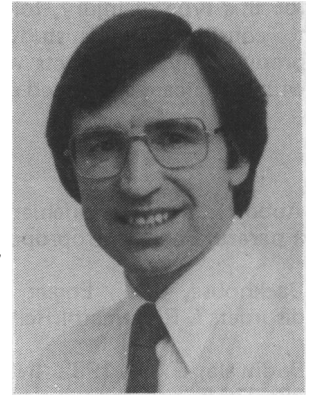


STEP TESTS OF CARDIORESPIRATORY FITNESS SUITABLE FOR MASS TESTING

J. WATKINS, PhD

Scottish School of Physical Education, Jordanhill College of Education, Glasgow, Scotland



J. Watkins

ABSTRACT

Step tests have been widely used to assess cardiorespiratory fitness. Work rate is determined by step height and step frequency, whereas the total amount of work done depends upon the number and duration of work periods. In all step tests which are suitable for mass testing, performance is based on heart rate during recovery. A review of the literature strongly suggests that step tests can give reliable and valid measures of cardiorespiratory fitness provided that (a) the subjects have received adequate instruction and practice in taking the test prior to the criterion test performance, and (b) the work load is above a certain threshold.

Key words: Step test, Cardiorespiratory fitness.

INTRODUCTION

Over the past two decades cardiorespiratory fitness, i.e. the efficiency of the respiratory and cardiovascular systems, has been the subject of a very large number of studies involving endurance-trained athletes. In contrast, it is only recently that the cardiorespiratory fitness of the general adult population has received much attention. This is due to the fact that diseases of the respiratory and cardiovascular systems have become a major cause of adult deaths in Western nations (Shephard, 1977).

The main function of the respiratory and cardiovascular systems is to provide the tissues of the body with oxygen, and cardiorespiratory fitness has been defined by Holmgren (1967) as the oxygen-forwarding capacity of the cardiorespiratory system. The cardiorespiratory system, or oxygen transport system, involves a chain of integrated processes which are concerned with a) the ventilation of the lungs, b) the diffusion of oxygen from the lungs into the blood, c) the transport of oxygen by the blood, and d) the diffusion of oxygen from the blood into the tissues. Limitations may occur at any point in the chain, and may vary from one individual to another (Asmussen, 1969). One of the most valid and widely used methods of assessing cardiorespiratory fitness, especially with endurance-trained athletes, is the maximum oxygen uptake test. The

achievement of maximum oxygen uptake ($\dot{V}O_2$ max), i.e. the maximum rate at which the body can consume oxygen, represents the maximum effective integration of the various processes which make up the oxygen transport system. Direct measurement of an individual's $\dot{V}O_2$ max involves sophisticated laboratory procedures and depends upon the willingness of the individual to work to exhaustion, usually on a treadmill or bicycle ergometer. Since laboratory facilities are not readily available, and maximum physical effort may be dangerous for some people, it follows that the direct measurement of $\dot{V}O_2$ max is not a practical method of assessing the cardiorespiratory fitness of large numbers of the general population. Therefore, for the purpose of mass testing it is necessary to adopt secondary standards of cardiorespiratory fitness which can be assessed by forms of submaximal exercise that can be administered in a wide range of environmental conditions (Shephard et al, 1968).

Endurance training will affect the various processes in the oxygen transport system to a greater or lesser extent. One of the most noticeable and well-documented effects of endurance training is the effect on the heart rate response to exercise; this is manifest in a decrease in the heart rate at rest and during steady-state submaximal exercise, and in a quicker return of the

heart rate to its resting level following exercise. For these reasons, and the fact that heart rate is relatively easy to monitor, two of the most widely used secondary standards of cardiorespiratory fitness are i) heart rate during steady-state submaximal exercise, and ii) heart rate during recovery from steady-state submaximal exercise. In each case, the lower the heart rate the fitter the individual. A number of tests involving one of three forms of exercise (stepping, cycling, running) have been devised for the assessment of cardiorespiratory fitness on the basis of the above secondary standards. The remainder of this article is concerned with a partial review of step tests which are suitable for mass testing.

The work rate in a step test is determined by the product of the step height and the step frequency, whereas the total amount of work done depends upon the number and duration of the work periods. In most step tests, assessment of performance is based either directly or indirectly on heart rate measurements. In some tests, for example, Shephard (1977) and Morton and Docherty (1970) performance is based on heart rate during exercise, the measurement of which requires fairly sophisticated equipment and as such these tests are impractical for mass testing. Step tests which are suitable for mass testing involve the assessment of heart rate during the post-exercise (recovery) period by means of one or more pulse counts. These tests can be divided into two categories, viz., single-stage (single work period) and multi-stage (two or more work periods).

Single-stage tests

The earliest reported step test, which came to be known as the Harvard step test (Brouha et al, 1944), was a single-stage test and was designed for use with undergraduate male students. The test requires the subject to step up and down on a 20" bench at a rate of 120

steps per minute (i.e., 30 cycles/min to a four-beat count) for a period of 5 minutes unless the subject stops from exhaustion before 5 minutes. The pulse is counted from 1 to 1½, 2 to 2½, 3 to 3½ minutes during the recovery period and performance (Physical Fitness Index, PFI) is assessed by means of a formula which incorporates the duration of stepping and the sum of the three pulse counts. Brouha et al (1944) presented PFI norms which were based on results obtained from over 8,000 male students. In its original form the Harvard step test is a very strenuous test and as such is not suitable for use with untrained and aged people (Ryhming, 1953). Furthermore, a number of studies have shown that the test is invalid for use with individuals of short stature, i.e., children and short adults, due to the 20" step which makes the test more a test of local endurance of the legs than a test of cardiorespiratory fitness (Datta et al, 1974). Consequently, a variety of single-stage step tests have been developed from the original Harvard step test model by modification of one or more of the work variables (step height, step frequency, duration of stepping) and the method of scoring; see Table I. With regard to the latter, at least three methods appear in the literature, viz.,

- i) the use of a single pulse count during recovery (e.g. McArdle et al, 1972) or the sum of two or more pulse counts during recovery (e.g. Michael and Adams, 1964).
- ii) a score based on the duration of stepping and the sum of three pulse counts during recovery (Brouha et al, 1944).
- iii) a score based on the ratio of heart rate during recovery to body weight (e.g. Tuxworth and Shahnawaz, 1977).

Normative standards for performance in a number of single-stage step tests, other than the Harvard step test,

TABLE I

Work variables, reliability and validity coefficients for a number of single-stage step tests

Source	Work Variables			n	Age (Yrs)	Rel (r)	Val (r)
	H	F	D				
Fox et al (1973)	20	120	5	135M	17-32		0.38 b
Skubic & Hodgkins (1963)	18	96	3	27F	11-23	0.82	
Michael & Adams (1964)	17	144	1	124M	17-25	0.82	
McArdle et al (1972)	16	88	3	36F	20.8	0.92	-0.75 b
Shapiro et al (1976)	13	100	6	63M	17-19		-0.81 p
Berry (1978)	16	96	3	30M	13-17		-0.67 t
Tuxworth & Shahnawaz (1977)	16	100	5	45M	23-41		-0.87 b

H = step height in inches

F = step frequency in steps per minute

D = duration of stepping in minutes

b, t correlation with directly measured maximum oxygen uptake on a bicycle ergometer and treadmill respectively

p correlation with predicted maximum oxygen uptake from submaximal bicycle ergometer test

have been produced for various groups of people. For example, Skubic and Hodgkins (1963 and 1964) produced norms for school girls and college women for performance in a three minute test involving an 18" bench and a stepping frequency of 96 steps per minute (24 cycles/min). Bailey and Mirwald (1978) presented norms for performance of 11-14 year old children in a six minute test involving a double 8" step and a stepping frequency of 114 steps per minute (19 cycles/min to a six-beat count).

Multi-stage Tests

In a multi-stage test a pulse count is taken in the rest period between successive work periods and the test is terminated when the pulse count reaches a certain level. One of the simplest multi-stage tests in terms of equipment and ease of administration is the Canadian Home Fitness Test which was designed for use in Canadian homes which tend to have staircases made up of 8" steps (Bailey et al, 1976). The tests consists of up to three work periods, each of three minutes duration, in which the subject steps up and down two 8" steps to a six beat count. A thirty second recovery period is given between successive work periods. The stepping frequency, which is increased at the start of the second and third work periods, depends upon the age of the subject. At the end of the first work period, a pulse count is taken during the interval 5-15 seconds of recovery. If the pulse count is above a certain age-related ceiling the test is terminated; otherwise the subject proceeds to a second work period. If the pulse count is still below the age-related ceiling at the end of the second work period, then the subject proceeds to a third work period. The stage at which the test is terminated determines the subject's fitness category, i.e. 'very poor' at the end of the first work period, 'poor' or 'average unfit' at the end of the second work period, and 'average fit', 'good', or 'very good' at the end of the third work period. Normative standards for performance on this test of Canadian subjects 15-69 years of age have been reported by Bailey et al (1976). Bailey and Mirwald (1978) presented normative standards for the performance of 11-14 year old children on a similar two-stage test which also involves the use of a double 8" step.

Kurucz et al (1969) devised a more elaborate multi-stage test which, if completed, consists of six work periods at each of three work rates. The work rates are a) 96 steps per minute on a 15" bench (innings 1 to 6), b) 120 steps per minute on a 15" bench (innings 7 to 12), and c) 120 steps per minute on a 20" bench (innings 13 to 18). Each innings consists of a work period of 30 seconds followed by a recovery period of 20 seconds. A pulse count is taken in the interval 5-15 seconds of each recovery period until a pulse count greater than or equal to 25 (i.e., a heart rate of greater than or equal to 150 b/min) is reached, at which point the test is terminated. A subject's score in the test is

the number of innings completed prior to achieving the criterion pulse count. If a subject completes all 18 innings without achieving the criterion pulse count, he or she is given a score of 19.

Reliability

In general, single-stage and multi-stage step tests have been found to be fairly reliable; see Tables I and II. The reliability of any test will be affected by two kinds of factors, viz., those which apply to all tests, such as the effects of practice, and others which will be test-specific. Practice will affect performance due to learning the habituation, which in turn will affect test-retest reliability. In physical tests, learning leads to an increase in the mechanical efficiency of the individual. The increase in mechanical efficiency is associated with a decrease in heart rate and, therefore, results in an increase in performance when the latter is based on heart rate measurements. With regard to habituation, most individuals will experience a certain amount of anxiety when presented with an unfamiliar test and/or test environment. The level of anxiety will be partly manifested in an increase in heart rate. Habituation refers to the gradual decrease in the level of anxiety and, therefore, decrease in heart rate which occurs as the individual becomes more familiar with the testing situation. Consequently, when performance in a test is based on heart rate measurements, habituation, like learning, will result in an improvement in performance. With regard to submaximal step tests, there is evidence to suggest that after two practice trials, no further improvements in performance due to learning and habituation occur (Shephard et al, 1968).

In relation to tests of cardiorespiratory fitness in which the heart rate during recovery is used to assess performance, there is one factor which will have a considerable effect on reliability, i.e., the accuracy of measuring heart rate during recovery. It is clear from the literature that in such tests, reliability depends to a considerable extent upon the accuracy of heart rate measurements; the more accurate the pulse counting, the higher the reliability of the test and vice-versa. Following a considerable amount of research involving the Canadian Home Fitness test Bailey and Mirwald (1978) report that regardless of age and sex, people who are untrained in pulse counting cannot make accurate pulse counts; however, they also point out that most people, including children, can be taught to make accurate pulse counts and as such, step tests of cardiorespiratory fitness are very suitable for mass testing.

In addition to the effects of practice and the accuracy of pulse counts, the reliability of single-stage step tests would also appear to be affected by the method of scoring. Watkins and Ewing (1984) studied the reliability of a single-stage step test in relation to six test scores which were based on three separate pulse counts.

The pulse-counts involved taking the time for 30 heart beats starting at 5, 60 and 120 seconds after the end of the work period. The first three scores (S1, S2 S3) were calculated by converting the three pulse-count times to heart rates in beats per minute. The other three scores (S4, S5, S6) were derived from the sum of the first and second, second and third, and all three separate scores respectively. The score which resulted in the highest reliability coefficient was S1 (0.94) which was followed by S4 (0.87), S6 (0.79), S2 (0.74), S5 (0.67) and S3 (0.55). The S1 score was based on a pulse count during the first 20 seconds of recovery and it is interesting to note that McArdle et al (1972) also found a high reliability coefficient (0.92) for a single-stage step test when using a score derived from a pulse count taken during the same period of time.

Validity

In terms of human performance, the degree of relationship between any pair of tests will depend upon the similarity between the tests, i.e. the greater similarity, the greater the relationship and vice-versa. With regard to physical tests, the type of movements involved, the degree of physiological stress placed on the subject, and the method of scoring will all have a bearing on the relationship between tests. With this in mind, it is reasonable to assume that step tests which involve sub-maximal effort, and which are assessed on the basis of one component of the oxygen transport process, i.e. heart rate, do not assess the same aspect of cardio-respiratory function as the maximum oxygen uptake test. This assumption would seem to be borne out by the results of studies in which the relationship between maximum oxygen uptake and performance in a variety of step tests has been investigated. In general, the relationship has been found to be fairly low; see Tables I and II. However, it is not possible to form any definite conclusions on the basis of these studies due to differences in experimental procedures. For example, all of the studies cited in Table I involved different step tests (work variables and methods of scoring). Furthermore, the procedures used to obtain VO_2 max values were not the same in all of the studies.

Whereas the relationship between performance in submaximal step tests and VO_2 max is not clear, the relationship between performance in some submaximal multi-stage step tests and a submaximal treadmill test, i.e. the Balke test (Balke and Ware, 1959), has been found to be fairly high. The Balke test involves walking at a progressively increasing work rate, and is scored on the basis of time to achieve a heart rate of 180 b/min. The multi-stage step tests devised by Kurucz et al (1969), Cotten (1971), and Witten (1973) are all based on the same scoring principle as the Balke test, i.e. the time to achieve a prescribed heart rate. For this reason, and the fact that the Balke test and the multi-stage step tests are all submaximal tests, it is, perhaps, not surprising that

fairly high correlations (0.94, 0.85, 0.84) have been found between these multi-stage tests and the Balke test; see Table II.

TABLE II
Reliability and validity coefficients for a number of multi-stage step tests

Source	n	Age (yrs)	Rel. (r)	Val (r)
Bailey et al (1976)	102 M & F	15-69	0.79	
	102 M	15-19		0.49 a
	144 F	15-19		0.61 a
	33 M	60-69		0.76 a
	56 F	60-69		0.40 a
Kurucz et al (1969)	24 M	19-56	0.94	
	30 M	19-56		0.94 b
Cotten (1971)*	34 M	20-30	0.95	0.84 b
	46 M	15-18	0.75	
Witten (1973)*	20 F	18.5		0.85 b
	51 F	18.8	0.90	
Santa Maria et al (1976)**	27 M	22.5	0.77	
	21 M	22.5		0.57 c

* similar test protocol to Kurucz et al (1969); see text.

** same test protocol as Kurucz et al (1969)

a correlation with predicted maximum oxygen uptake.

b correlation with 'time to a heart rate of 180' in a progressive treadmill test.

c correlation with directly measured maximum oxygen uptake on a bicycle ergometer.

It is clear that further studies are needed in order to identify those aspects of cardiorespiratory function which underlie performance in different forms of cardiorespiratory fitness tests. Evidence that different aspects of cardiorespiratory function exist has been reported by Van Gerven and Eynde (1977) who identified, by factor analysis, separate aspects of cardiorespiratory function underlying performance in a step test, a distance-run test and a bicycle ergometer test.

If 'heart rate during recovery from steady-state sub-maximal effort' is an acceptable secondary standard of cardiorespiratory fitness, then the value of step tests lies in their ability to reliably differentiate between individuals with respect to this criterion. A number of the factors which affect the reliability of step tests have been described earlier; however, there is another factor which will affect both the reliability and validity of all submaximal tests of cardiorespiratory fitness, i.e. the intensity and duration of the work period. Åstrand and Ryhming (1954) maintain that the duration of work should be enough to permit the adjustment of circulation and ventilation to the level of the exercise, which

for most people will be between three and four minutes, and that the work rate should result in a steady state heart rate between 125 and 170 b/min. For most untrained adolescents and adults, this intensity of effort can be achieved in a single-stage step test by using a step height of approximately 16" and a stepping frequency of approximately 90 and 100 steps per minute (to a four-beat count) for females and males respectively (Ryhming, 1953; Katch and McArdle, 1977; Berry, 1978). It may be necessary to set higher work rates for fairly fit individuals, in which case it is desirable to increase the stepping frequency rather than the step height since there is evidence to suggest that a step height greater than 16" tends to penalise heavier individuals (McArdle et al, 1972) and to alter the test from one of cardiorespiratory fitness to one of local endurance of the legs (Datta et al, 1974).

SUMMARY

For the purposes of mass testing a reliable and valid

measure of cardiorespiratory fitness can be obtained for healthy adolescents and adults by using a single-stage submaximal step test which is scored on the basis of heart rate during recovery provided that,

- i) the subjects are well practiced in taking pulse counts.
- ii) two practice trials of the step test are given prior to the criterion test performance.
- iii) the step height is approximately 16", the stepping frequency is approximately 90 and 100 steps per minute for females and males respectively, and the duration of stepping is between 3 and 5 minutes.
- iv) the test score is based on a single pulse count taken during the first 30 seconds of recovery.

Multi-stage submaximal step tests which are based on a gradual increase in work rate and are scored on the basis of heart rate during recovery also provide reliable and valid measures of cardiorespiratory fitness, but these tests are not as practical as single-stage tests for the purpose of mass testing.

REFERENCES

- Asmussen, E., 1969 "Some physiological aspects of fitness for sport and work". Proceedings of the Royal Society of Medicine 62: 1160-1163.
- Åstrand, P.-O. and Ryhming, I., 1954 "A nomogram for calculation of aerobic capacity from pulse rate during submaximal work". Journal of Applied Physiology 7: 218-221.
- Bailey, D. A., Shephard, R. J. and Mirwald, R. L., 1976 "Validation of a self-administered home test of cardiorespiratory fitness". Canadian Journal of Applied Sport Sciences 1: 67-78.
- Bailey, D. A. and Mirwald, R. L., 1978 "A children's test of fitness". In: J. Borms and M. Hebbelinck (eds.) Pediatric Work Physiology. Medicine and Sport Vol. 11: 56-64. S. Karger, Basel.
- Balke, B. and Ware, R. W., 1959 "An experimental study of physical fitness of air force personnel". United States Armed Forces Medical Journal 10: 675-688.
- Berry, R. J., 1978 "Estimation of maximal oxygen uptake from step test recovery pulse rates in high school boys". Unpublished thesis, Department of Human Movement Studies, University of Queensland, Australia.
- Brouha, L., Fradd, N. W. and Savage, B. M., 1944 "Studies in physical efficiency of college students". Research Quarterly 15: 211-224.
- Cotten, D. J. 1971 "A modified step test for group cardiovascular testing". Research Quarterly 42: 91-95.
- Datta, S. R., Chatterjee, B. B. and Roy, B. N., 1974 "An improved simple exercise test for evaluation of physical fitness". Ergonomics 17: 105-112.
- Fox, E. L., Billings, C. E., Bartels, R. L., Bason, R. and Mathews, D. K., 1973 "Fitness standards for male college students". Internationale Zeitschrift für Angewandte Physiologie 31: 231-236.
- Holmgren, A., 1967 "Cardiorespiratory determinants of cardiovascular fitness". Canadian Medical Association Journal 96: 697-702.
- Katch, R. I. and McArdle, W. D., 1977. Nutrition, weight control, and exercise. Houghton-Mifflin: Boston.

- Kurucz, R. L., Fox, E. L. and Mathews, D. K. 1969 "Construction of a submaximal cardiovascular step test". *Research Quarterly* 40: 115-122.
- McArdle, W. D., Katch, F. I., Pechar, G. S., Jacobson, L. and Ruck, S., 1972 "Reliability and interrelationships between maximal oxygen intake, physical work capacity and step test scores in college women". *Medicine and Science in Sports and Exercise* 4: 182-186.
- Michael, E. D. and Adams, A., 1964 "The use of a one-minute step test to estimate exercise fitness". *Ergonomics* 7: 211-215.
- Morton, A. R. and Docherty, D., 1970 "A multi-stage continuous step test of cardiovascular fitness". *Australian Journal of Sports Medicine* 3: 4-26.
- Ryhming, I., 1953 "A modified Harvard step test for the evaluation of physical fitness". *Arbeitsphysiologie* 15: 235-250.
- Santa Maria, D. L., Kinnear, G. R., Kearney, J. T. and Martin, T. P., 1976 "The objectivity, reliability and validity of the O.S.U. step test for college males". *Research Quarterly* 47: 445-452.
- Shapiro, A., Shapiro, Y. and Magazanik, A., 1976 "A simple step test to predict aerobic capacity". *Journal of Sports Medicine and Physical Fitness* 16: 209-214.
- Shephard, R. J., Allen, C., Benade, A. J. S., Davies, C. T. M., di Prampero, P. E., Hedman, R., Merriman, J. E., Myhre, K. and Simmons, R., 1968 "Standardisation of submaximal exercise tests". *Bulletin of the World Health Organisation* 38: 765-775.
- Shephard, R. J., 1977. *Endurance fitness*. University of Toronto: Toronto.
- Skubic, V. and Hodgkins, V., 1963 "Cardiovascular efficiency tests for girls and women". *Research Quarterly* 34: 191-198.
- Skubic, V. and Hodgkins, V., 1964 "Cardiovascular efficiency test scores for junior and senior high school girls in the United States". *Research Quarterly* 35: 184-192.
- Tuxworth, W. and Shahnawaz, M., 1977 "The design and evaluation of a step test for the rapid prediction of physical work capacity in an unsophisticated industrial work force". *Ergonomics* 20: 181-191.
- Van Gerven, D. and Eynde, B. V., 1977 "Factor analysis of three cardiorespiratory tests within a group of male and female secondary school students". In: F. Landry and W. Orban (Eds.) *Biomechanics of Sports and Kinanthropometry*. In: *International Congress of Physical Activity Sciences*, Leuven.
- Watkins, J. and Ewing, B. G., 1984 "The effects of practice and method of scoring on performance in a step test suitable for use in schools". *Scottish Journal of Physical Education*, 12 (2): 1984.
- Witten, C., 1973 "Construction of a submaximal cardiovascular step test for college females". *Research Quarterly* 44: 46-50.
-