

A Review of Various Step Test Protocols for Use in Assessing Aerobic Fitness in Schools

不同台階測試用以量度學童之心肺耐力

Peggy PY CHEUNG

*Department of Health and Physical Education,
The Hong Kong Institute of Education, HONG KONG*

張佩儀

香港教育學院
健康與體育學系



Abstract

Among the commonly used field tests, the step test is a practical field assessment that provides acceptable validity and resolves the concerns on space and time requirement for implementation in Hong Kong schools. The present paper reviews various step test protocols and discusses the concerns on step test implementation. Recommendation on step test implementation in assessing children's aerobic fitness in school setting was discussed.

Keywords: field test, school children, physical fitness

摘要

在眾多的場地測試中，台階測試是其中一項適合作為應用在學校之心肺耐力測試的方法，本文旨在介紹不同台階測試的模式，以及討論有關研究及應用台階測試之相關事項。

Introduction

There have been numerous reasons given to justify the conduction of school-based physical fitness tests. The rationale for the application of fitness testing on children includes program evaluation, student motivation and recognition, selection of sport talents, and promotion of cognitive and affective learning (Pate, 1989). Among the components in health-related physical fitness, the measurement of aerobic fitness has been considered as one of the major components in the youth fitness test battery, such as FITNESSGRAM (Welk & Meredith, 2008) in the United States and Australian Fitness Education Award for students (ACHPER, 1996). The selection of aerobic fitness test in the test battery was based on the belief that improved aerobic function is related to positive health

and, therefore, should be of national concern (Jackson, 1976). Fitness testing in children, nevertheless, can provide valuable quantitative data for the educators, with which to evaluate programs aimed at improving the fitness level of children.

Facing the constraints on the practicability of conducting direct measurements of children's aerobic fitness, various sub-maximal field-tests (e.g., run test, walk test, step test) with acceptable validity and reliability have been developed. Owing to concerns regarding the amount of space, time required and the expenditure of effort involved in the field tests, the step test, is considered an applicable alternative that can be used within school environment. This paper reviews different step test

protocols as well as a discussion of studies on the issues related to step test implementation on assessing individual's aerobic fitness.

Step Test Protocols

The implementation of the step test requires individuals to perform a simple step up and step down movement from a specific bench height for a pre-determined stepping cadence and duration. The stepping cadence is calculated as steps per minute. To ensure proper stepping, a metronome or a tape recording of metronome frequency is used to regulate the stepping rhythm. The equipment for the test is easy to obtain because the step platform is portable and available in most school environments.

There are two types of step test protocols, namely single-stage and multiple-stage tests. The single-stage step test requires the participants to perform one given workload within a specific time period. The multiple-stage step test requires participants to perform incremental workloads by altering the stepping cadence. The test is terminated when the participants' heart rate (HR) reaches a pre-determined level. With the pre-determined time set for implementation, the single-stage step test is considered more suitable for school environment in the form of mass testing. The followings are descriptions on common single-stage step test protocols.

Astrand-Ryhming Step Test

The Astrand-Ryhming Step Test (Astrand & Ryhming, 1954) was first developed as a sub-maximal work test, in which the prediction of maximum oxygen consumption (VO₂max) was made from the exercise intensity at the HR between 125 and 170 beats/minute. In this test protocol, the height levels of the bench are 40, 27, and 33 cm for young males, for males aged 40–70 and for females, respectively. The stepping cadence is designed at 22.5 steps/minute, and the total duration of the test is 5 minutes. The HR measurement is then made between 15 and 30 seconds immediately after the test. The scores of the test are then converted into VO₂ using a nomogram.

The Astrand-Ryhming Step Test can be used as an initial test to identify individuals with less than optimum levels of aerobic fitness. The test can also be used to monitor their improvement as a result of conditioning activity. However, few studies on the application of this test exist, thus limiting the popularity of the test.

Harvard Step Test

The Harvard Step Test was developed during World War II and was originally used for the selection of combat officers (Brouha, 1943). The test protocol requires participants to step up and down a 50 cm-height bench at a stepping cadence of 30 steps/minute for 5 minutes. Three 30-second measured HRs after the tests are used to reflect the individual levels of fitness while the participant is in a sitting position. The 30-second HR is measured at three phases: 1–1.5, 2–2.5, and 3–3.5 minutes of the recovery phase. A Physical Efficiency Index (PEI) is then computed using the following formula:

$$PEI = \frac{\text{Duration of exercise in seconds} \times 100}{2 \times \text{sum of pulse counts in recovery}}$$

If the participants are unable to complete the 5-minute step test, the duration of the exercise used will be the number of seconds elapsed when the test terminated. Based on about 8,000 tests of college men, norms have been established wherein a PEI over 90 is considered excellent and that below 54 is poor in terms of aerobic fitness.

The Kasch Pulse Recovery Step Test (Kasch ST)

One of the major differences between various step test protocols is the utilization of different phrases of post-exercise heart rate (PHR) interval to reflect individual aerobic fitness. The Kasch Pulse Recovery Step Test (Kasch & Boyer, 1968) provides an example, which adopts a recovery HR of one whole minute after the step test. The original test protocol was designed for participants ranging from 6-year old children to adults. The test uses a 12-inch (30.5 cm) bench and a stepping cadence of 30 steps/minute. After the exercise, a 1-minute HR is counted with the participant in the sitting position. Suggested norms are available to categorize the children's aerobic fitness. For the age group of 6–12, 1-minute post-exercise HR values of 74 (for boys) and 82 (for girls) are graded as superior.

The YMCA Step Test is a modification of the Kasch Step Test; it uses the same step height but a different stepping cadence of 24 steps/minute (Goldings, Myers & Sinning, 1989). Normative value is available to categorize the individual aerobic fitness. However, there has been no reported validity of the test between measured VO₂ max and HR response.

The distinctive characteristic of both the Kasch Pulse Recovery Step Test and the YMCA Step Test is the application of the recovery HR after the test. The test is not intended to predict individual $\dot{V}O_2$ max from the recovery HR, but to categorize the aerobic fitness based on the recovery HR response.

Hodgkin's and Skubin Step Test (SH Step Test)

The SH Step Test (Skubic & Hodgkins, 1963) is a 3-minute step test assessed on young girls and older women population. The test protocol consists of a 45.7-cm high bench at a rate of 24 steps/minute. After 3 minutes, the participants are allowed to rest in a sitting position for 1 minute, followed by a recording of a 30-second recovery pulse. Cardiovascular (CV) efficiency is calculated as follows:

$$\text{CV efficiency} = \frac{\text{Duration of exercise (sec)} \times 100}{\text{Recovery pulse} \times 5.6}$$

Individual fitness based on CV efficiency rates scores of 71–100 as excellent and scores of below 27 as very poor and is used to distinguish between trained, active, and sedentary participants. The SH Step Test utilizes the recovery HR to validate the measured $\dot{V}O_2$ max and the scoring with moderate correlation (-.64) was reported.

Queens College Step Test (QC Step Test)

The QC Step Test has been reported as a fairly valid indicator of the aerobic capacity of college female students (-.75) (McArdle, Katch, Pechar, Jacobson & Ruchk, 1972). The test protocol required the participants to perform a 3-minute step test at a cadence of 22 steps/minute utilized a bench height of 41.3 cm, which was actually the height of a bleacher step at Queens College. In the study, 41 female undergraduate students performed a 3-minute step test and a maximal treadmill test to obtain $\dot{V}O_2$ max. At the completion of the step test, a 15-second recovery HR was measured at the 6th–20th sec (first recovery HR) and the 60th–75th sec (second recovery HR) after the test, with the participants in a standing position.

Compared with the recovery HR correlation on the $\dot{V}O_2$ max, a higher validity coefficient (-.75) was obtained on the first recovery HR. In addition, lower recovery HRs are associated with higher $\dot{V}O_2$ max. Regression equation is generated to estimate individual $\dot{V}O_2$ max as follows:

$$\begin{aligned} \text{(Women) } \dot{V}O_2 \text{ max} &= 65.81 - (0.1847 \times \text{HR}), \text{ and} \\ \text{(Men) } \dot{V}O_2 \text{ max} &= 111.33 - (0.42 \times \text{HR}). \end{aligned}$$

The QC Step Test utilized the benches available at the college which ease the practicality issue, and the protocols demonstrated better correlations between PHR and measured $\dot{V}O_2$ max in college women ($r = -.75$) than the SH Step Test (-.64) (McArdle et al., 1972).

Height-Adjusted Step Test

The height-adjusted step test was initially developed for female adults aged 19–35 (Francis & Cuippeper, 1988). It was subsequently validated for male adults aged 18–47 (Francis & Brasher, 1992) and children aged 6–18 (Francis & Feinstein, 1991). In all these studies, a height-adjustable step platform has been adopted. The platform height is determined by the geometric relationship between individual stature and femur length. The step height determination is based on the level of the foot when the hip is flexed at 73.3°. Based on the height-adjusted model research done on children (Francis & Feinstein, 1991), the results of the test demonstrated satisfactory coefficients between the 15-second PHR and the direct measured $\dot{V}O_2$ max in all stepping frequencies (.81, .79 and .80 for 22, 26 and 30, steps/minute, respectively). The regression equation for estimation of 6–18 years old children $\dot{V}O_2$ max is as follows:

$$\dot{V}O_2 \text{ max} = 103.42 - (1.588 \times \text{15-second Recovery HR})$$

Although satisfactory validity has been obtained for the height-adjusted step test for children between $\dot{V}O_2$ max and PHR, the test protocol suffered from a major constraints for mass testing, i.e., the requirement of using individualized step platforms.

The different single-stage step test protocols consisted of the modifications among four related tested variables. These were the platform height, the stepping cadence, duration of stepping and the HR measurement. The resulting protocols reflected diversity in bench heights ranging from 30 to 50cm, stepping cadences ranging from 18 to 36 steps/min, and the stepping duration that ranged from 3 to 5 min. Nevertheless, most tests utilized the post-exercise HR response at varying intervals from 15 to 120 seconds to estimate individual aerobic fitness. Table 1 summarizes the single-stage step test protocols.

Table 1. Summary of Single Stage Step Test Protocols.

Step Test (ST)	Step Height (cm)	Stepping Cadence (Steps/min)	Duration of test	HR measurement / position
Astrand-Rhyming (Astrand & Rhyming, 1954)	young male 40 male 27 female 33	22.5	5 min	between 15 and 30 sec after the test/ not specify
Harvard (Brouha, 1943)	50.8	30	5 min	1-1.5 min, 2-2.5min, 3-3.5 min after the test / sitting
Kasch (Katch & Boyer, 1968)	30.8	24	3 min	1-2 min after the test / sitting
YMCA (Goldings, Myers & Sinning, 1989)	30.8	24	3 min	1-2 min after the test / sitting
Skubic & Hodgkins (Skubic & Hodgkins, 1963)	45.7	24	3 min	1-1.5 min after the test / sitting
Queens' College (McArdle et al., 1972)	41.3	22	3 min	15 sec at 6th-20th sec after the test / standing
Height-Adjusted (Francis & Feinstein, 1991)	According to the knee angle produced when the feet flat on the step	22 26 30	3 min	15 sec at 6th-20th sec after the test / standing

Considerations on the Step Test Protocols

Studies on the validity of field tests to assess individual aerobic fitness was examined by a) the correlation between directly measured $\dot{V}O_2$ max and the $\dot{V}O_2$ max estimated from physiological responses obtained from the field tests, such as the heart rate (HR), or b) the correlation between directly measured $\dot{V}O_2$ max and test performance. The application of step test, particularly, assessed individual aerobic fitness based on the inverse relationship between the measurement of HR during or immediately after a stepping exercise and the participant's maximal aerobic power (Rowland, 1996). There are three assumptions in the $\dot{V}O_2$ max estimation from the sub-maximal exercise. First, it is assumed that there is a linear relationship between HR and $\dot{V}O_2$ for various intensities of light to moderate exercises. In a standard work task, such as the step test, individuals with lower PHR are estimated to have higher $\dot{V}O_2$ max and vice versa. Second, it is assumed that there is a similar maximum HR at a given age. The step test utilizes the

PHR to assess the individual's aerobic fitness, whereas maximum HR has no influence on the prediction. Based on this assumption, variation of maximum HR has no effect on the step test. Third notes that this is assumed that constant mechanical efficiency is assumed in the performance of the exercise. The performance of the step test has been standardized by the pre-determined cadence, which minimizes the discrepancy in the mechanical efficiency of the exercise (McArdle, Katch & Katch, 1986).

Although the above assumptions constitute a certain degree of prediction error on $\dot{V}O_2$ max by sub-maximal exercise, the step test is still an effective method for field testing of individual aerobic fitness when considering the purpose of categorizing individual aerobic fitness and the concerns of practicability (McArdle et al., 1986). In comparison with different field tests on aerobic fitness, concerns on the step test protocol application, such as HR measurement, platform height, and stepping cadence is discussed.

HR measurement

PHR is a major and practical physiologic variable used to estimate individual aerobic fitness during step test. Different measurement on PHR had been reported to have acceptable validity, such as on the 5th and 20th seconds of PHR (.75) (McArdle et al.,1972) and on the 15-second PHR (.70) (Francis & Cuiepper, 1988). The 15-second PHR was reported to have high correlation on VO₂ max with the stepping frequencies of 22 (.72) and 26 steps/minute (.81) (Francis & Brasher, 1992). The PHR to estimate VO₂ max was also valid on children population study. For example, the 15-second PHR produced high correlation on directly measured VO₂ max levels are .80 at 30 steps/minute, .79 at 26 steps/minute, and .81 at 22 steps/minute on children aged 6-14 (Francis & Feinstein,1991).

Participants with higher PHR tended to have lower VO₂ max, whereas those with a lower PHR tended to be associated with a relatively higher VO₂ max. HR responses after stepping correlate well with the directly measured VO₂ max. The reported satisfactory validity further demonstrates that PHR is a reasonable variable for estimating individual aerobic fitness. A summary of studies on correlation of VO₂ max and HR response after step test is shown on Table 2.

Table 2. Correlation of VO₂ max and HR response after Step Test.

Studies	Subjects	HR measurement	Correlation between VO ₂ max and PHR
McArdle et.al., 1972	41 college female	15 sec post exercise	R=.75 (22 steps/min)
Francis & Cuiepper, 1988	30 female aged 19-35	15 sec post exercise	R=.70 (22 steps/min)
Francis & Brasher, 1992	43 males aged 18-47	15 sec post exercise	R=.72 (22 steps/min) R=.81 (26 steps/min)
Francis & Feinstein, 1991	6-14 children	15 sec post exercise	R=.81 (22steps/min) R=.79 (26 steps/min) R=.80 (30 steps/min)
Hui & Cheung, 2004	7-12 children	15 sec post exercise	R=.76 (22steps/min) R=.78 (26 steps/min) R=.72 (30 steps/min)

Platform height

The step platform should be at a reasonable height level to produce valid relationship between VO₂ max and HR response. When the platform is too low, individual HR is not elevated sufficiently for assessment of aerobic fitness. Conversely, too high a platform height either affects the stepping action or subjects are unable to finish the test. For instance, the height used in the Harvard Step Test (50.8 cm) is problematic that participants stop the test because of premature fatigue in their lower limbs instead of the cardiorespiratory deficiency (Bandyopadhyay, 2007). In an earlier study of a group of 75 healthy young men performed the Harvard Step Test, 22 subjects failed to complete the 5-minute test (Keen & Sloan, 1958). Similarly, only 30% of the young adult participants

were able to finish the whole Harvard Step Test among a group of 12 young adults (Datta, Chatterjee & Roy, 1974). The inability to complete the test, will highly affect the validity and feasibility of the test.

In the belief that platform height affects the performance of the participants in the step test, a height-adjusted platform has been suggested (Francis & Feinstein, 1991). However, study reported that similar correlation has been found between HR and VO₂ max by using fixed-height platform (.75) and a height-adjusted platform (.72) (McArdle et al., 1972). As such, the prediction ability of individual aerobic fitness performed by a height-adjusted platform is not necessarily better than that of a fixed-height platform.

Concerning the platform height issue on step test protocols, different studies and recommendations had been reported, such as a double 20 cm step platform for children (Jette, Ashton & Sharratt, 1984), a 8-inches platform (for age 8-10), a 12-inches platform (for age 10-12) (Montoye, 1983), a 25-cm platform (Shapiro, Shapiro & Magazanik, 1976), and a 40-cm high platform (Chin & Kuei, 2007).

There are no consistent findings on the influence of step height on the predictive ability of step test on $\dot{V}O_2$ max, the selection of step platform height is suggested to base on convenience and practicality. For practical reason, a 12-inches (30.5-cm) high platform, which can easily accessed, is suggested to be adopted as the step test protocols in local schools environment.

Stepping cadence

The calculation of $\dot{V}O_2$ in stepping involved all phases in a stepping cycle, including upward displacement, downward displacement, and horizontal displacement. The authors of that work have proposed that the relative oxygen requirement in stepping is primarily dependent upon stepping rate and—to a lesser extent—on stepping height, rather than on the absolute workload. Hence, for a given work rate, the lower the stepping cadence resulted in the lower $\dot{V}O_2$ of stepping (Nagle, Balke & Naughton, 1965). Nevertheless, the stepping cadence is believed to be more important than stepping height for meeting the oxygen requirements of the task. Consequently, it is hypothesized that the increase in $\dot{V}O_2$ brought about by increased stepping cadence is due to the need to lift the leading leg in excess of the bench height more frequently for the given work load (Shahnawaz, 1978).

Previous studies had conducted to compare different stepping cadences in performing step test in children. By utilizing a height-adjusted step platform to examine the correlation coefficient between PHR and $\dot{V}O_2$ max, all stepping cadences produced similar correlation, with 22 steps/minute (.81), 26 steps/minute (.79) and 30 steps/minute (.80) (Francis & Feinstein, 1991) and a stepping cadence of 22 steps/minute protocol is recommended. In another study using 12-inches high step platform for children, the results indicated that the criterion-related validities of the three stepping protocols were similar, with the protocol of 26 steps/minute (.78) being slightly better than the protocol of 22 steps/minute (.76) and 30 steps/minute (.72) (Hui & Cheung, 2004).

Increasing stepping cadence elevates the work intensity, but too high a rate in the stepping cadence affects movement patterns and causes problems in maintaining balance and movement coordination. The participant might not be able to stand upright when both feet on the platform are stepping at higher cadence rates. The stepping cadence, therefore, should be properly monitored to ensure the performance of the step test.

Conclusion and Recommendation

Physical fitness testing alerted teachers, parents, and educators in gaining a better understanding of the physical fitness levels of children. One of the major concerns in selecting a field test for evaluating children's aerobic fitness is practicability. This is especially critical when strenuous effort is required to test school-aged children. Neither student nor researcher is enthusiastic about forcing maximal exertion to exhaustion as it often discourages children from participation in future tests (Francis & Feinstein, 1991). The step test is advantageous because by maintaining reasonable validity, it requires a moderate level of intensity throughout the test. The sub-maximal effort requirement of the test made the test feasible for people with different fitness level. Particularly, the test is also useful in epidemiological study and measurement on the fitness status among different population (e.g., China National Physical Test Battery).

Considering the implementation of aerobic fitness test in a group of students, the space and time requirement is always a prime concern in local school environment. The step test protocol that can be completed within 3-5 minutes in a small space or even in indoor provides another alternative for PE teachers. The step test, therefore, is recommended as a physical fitness test to measure children's aerobic fitness in a school setting due to its reported validity and its advantages in time economy, ease of administration, and feasibility for mass testing. Recommendations for implementation of step test in school setting (e.g., during PE classes), include: 1) selection of appropriate step test protocol and consistently apply the same protocol for the purpose of comparison on fitness level; 2) proper measurement of the HR by teaching students about palpation of pulse, which also can be an excellent educational process on students health and fitness knowledge; 3) emphasis on the proper stepping skill, with participants stand upright when both feet are on the platform.

The present review showed that the step test protocols design froze in the last 20 years, yet continuous studies have impeded on the validity of the step test in assessing individual aerobic fitness. The study on the step test application is, therefore, an area worth further study.

Reference

- Astrand, P.O., & Ryhming, I. (1954). A nomogram for the calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *Journal of Applied Physiology*, 7, 218-221.
- Australian Council for Health, Physical Education and Recreation [ACHPER]. (1996). *User's manual & curriculum ideas: Australian fitness education award*. Australia: Author.
- Bandyopadhyay, A. (2007). Queen's college step test - an alternative of harvard step test in young indian men. *International Journal of Applied Sports Sciences*, 19(2), 1-6.
- Brouha, L. (1943). The step test: A simple method of measuring physical fitness for muscular work in young men. *Research Quarterly for Exercise and Sport*, 14, 31-36.
- Chin, ML., & Kuei, F.L. (2007). Estimation of VO₂ max: A comparative analysis of post-exercise heart rate and physical fitness index from 3-minute step test. *Journal of Exercise Science and Fitness*, 5(2), 118-123.
- Datta, S., Chatterjee, B., & Roy, B. (1974). An improved simple exercise test for evaluation of physical fitness. *Ergonomics*, 17, 105-112.
- Francis, K., & Brasher, J. (1992). A height-adjusted step test for predicting maximal oxygen consumption in males. *Journal of Sports Medicine and Physical Fitness*, 32, 282-287.
- Francis, K., & Cuipepper, M. (1988). Validation of a three minute height-adjusted step test. *Journal of Sports Medicine and Physical Fitness*, 28, 229-233.
- Francis, K., & Feinstein, R. (1991). A simple height-specific and rate-specific step test for children. *Southern Medicine Journal*, 8, 169-174.
- Golding, L.A., Myers, C.R., & Sinning, W.E. (Eds.). (1989). *Y's way to physical fitness*. (3rd ed.). Champaign, IL: Human Kinetics.
- Hui, SC., & Cheung, PY. (2004). Comparison of the effects of three stepping cadences on the criterion-related validity of a step test in Chinese children. *Measurement in Physical Education and Exercise Science*, 8(3), 167-179.
- Jackson, A. S., & Coleman, A.E. (1976). Validation of distance run tests for elementary school children. *Research Quarterly*, 47(1), 87-94.
- Jette, M., Ashton, N.J., & Sharratt, M.T. (1984). Development of a cardiorespiratory step-test of fitness for children 7-14 years of age. *Canadian Journal of Public Health*, 75, 212-217.
- Kasch, F., & Boyer, J. (1968). *Adult Fitness: Principles and Practice*. Palo Alto CA, National Press Books.
- Keen, E.N., & Sloan, A.W. (1958). Observations on the Harvard Step Test. *Journal of Applied Physiology*, 13, 241-243.
- McArdle, W.D., Katch, F.I., & Katch, V.L. (Eds.). (1986). *Exercise physiology : energy, nutrition, and human performance*. (2nd ed). Philadelphia: Lea & Febiger.
- McArdle, W.D., Katch, F.I., Pechar, G.S., Jacobson, L., & 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*, 4, 182-186.
- Montoye, H. J. (1983). Risk indicators for cardiovascular disease in relation to physical activity in youth. In R.A. Binkhorst., H.C.G. Kemper., and W.H.M. Saris (Eds.). *Children and Exercise*, XI. (pp. 3-25). Champaign, IL : Human Kinetics.

- Nagle, F.J., Balke, B., & Naughton, J.P. (1965). Gradational step tests for assessing work capacity. *Journal of Applied Physiology*, 20, 745-748.
- Pate, R. R., (1989). The case for large scale physical fitness testing in American youth. *Pediatric Exercise Science*, 1, 290-294.
- Rowland, T. W. (1996). *Developmental Exercise Physiology*. Champaign, IL : Human Kinetics.
- Shahnawaz, H. (1978). Influence of limb length on a stepping exercise. *Journal of Applied Physiology*, 44, 346-349.
- Shapiro, A., Shapiro, Y., & Magazanik, A. (1976). A simple test to predict aerobic capacity. *Journal of Sports Medicine and Physical Fitness*, 16, 209-214.
- Skubic, V. & Hodgkins, J (1963). Cardiovascular Efficiency Test for Girls and Women. *Research Quarterly*, 34, 191.
- Welk, G. J., Meredith, M.D. (Eds.). (2008). *Fitnessgram / Activitygram Reference Guide*. Dallas, TX: The Cooper Institute.

Correspondence

Dr. Peggy Cheung
Department of Health and Physical Education,
Hong Kong Institute of Education,
Tai Po, N.T., Hong Kong.
Email: cheungpy@ied.edu.hk