A Step Test for Evaluating the Aerobic Fitness of Children and Adolescents With Mental Retardation

Kenneth H. Pitetti, Bo Fernhall, Nancy Stubbs, and Louis V. Stadler, Jr.

The purpose of this study was to determine if a step test could be feasible, reliable, and valid for youths with educable (EMR) or trainable (TMR) mental retardation. Thirteen males and 11 females (age $M = 14.7 \pm 2.7$ yr) with EMR or TMR participated in this study. Three step tests were employed using one platform height and stepping frequencies of 13, 15, and 17 ascents/min for 3 min. Recovery HR was used to estimate VO$_{2\text{peak}}$. Though significant, correlations between the recovery HR and VO$_{2\text{peak}}$ for the 15 ($r = -0.48$) and 17 ($r = -0.46$) ascents/min were not high enough to be considered valid indicators of VO$_{2\text{peak}}$. The large standard errors of the estimate and total errors suggested systematic errors of prediction. Furthermore, the measured VO$_{2\text{peak}}$ was significantly different from the estimated values at all step rates ($p < .05$). The step-test was relatively feasible, but was not a valid test of VO$_{2\text{peak}}$ in this population.

The preferred method of determining cardiorespiratory fitness involves performing a maximum symptom-limited exercise test on a treadmill or cycle ergometer in conjunction with collection and evaluation of expired gases (1). The use of such tests is often limited due to the cost of highly technical equipment and the expertise needed. In addition, it requires significant cooperation and motivation from the person being evaluated. Motivation can cause a substantial variability of results, especially for persons with mental retardation (MR) (13). Evaluation of children with MR entails additional difficulties in that children may be unfamiliar with perceiving and coping with the discomfort of exercise.

For adults with MR, alternative test procedures have been developed to permit evaluation of the exercise capacity on the basis of standardized submaximal tests (13). For instance, the 1.5 mile (2.4 km) run-walk (6), the 1 mile (1.6 km) Rockport Fitness Walking Test (15), and the Canadian Standardized Test for Fitness (CSTF) modified step test (12) have been shown to be valid and reliable field tests in determining cardiovascular fitness of adults with MR. To date, however,

K.H. Pitetti, N. Stubbs, and L.V. Stadler, Jr., are with the School of Health Science at Wichita State University, Campus Box 43, Wichita, KS 67260-0043. B. Fernhall is with the Exercise Science Programs, School of Medicine and Health Sciences at the George Washington University, 817 23rd St. NW, Washington, DC 20052.
standardized submaximal field tests to determine the cardiovascular fitness of children and adolescents with MR have not been established.

Recently, Francis and Feinstein (7) developed a single stage step test for nonhandicapped children and adolescents (ages 6 to 18 years) that addressed some of the problems associated with earlier fixed-height step test such as the CSTF step-test (18). This step test (7) uses an anatomical model that standardizes the work efficiency of stepping, based on the stature of an individual. Whether this step test can be applied to children and adolescents classified as educable (EMR) or trainable (TMR) mentally retarded (MR) is not known. Therefore, the purpose of this study was to determine the feasibility, reliability, and validity of performing a step-test similar to that established by Francis and Feinstein (7) for children and adolescents with EMR or TMR.

Methods

Subjects

Twenty-four youths (13 males and 11 females, ages 10 to 18 years) classified as either EMR or TMR participated in this study. Eight (4 males and 4 females) had Down syndrome. Subjects were recruited from a summer school program in a Midwestern city, and none of the subjects were institutionalized. Subject height (to the nearest 1/4 inch) and weight (to the nearest 1/4 lb) was measured on a standard physician scale (Detecto) and was converted to the metric system. Body mass index (BMI) was determined by the formula: BMI = body weight (kg)/height² (m). Informed consent was obtained from the subjects' parents or legal guardians. This study was in accordance with university guidelines for research involving human subjects and was approved by the institutional review board.

None of the subjects had any medical contraindications to exercise. Any subject who had a congenital heart defect or who was taking medication affecting heart function was excluded from this study. Furthermore, none of the subjects had any orthopedic or motor limitations that prevented them from walking and jogging on a treadmill or ascending or descending stairs.

Description of the Anatomical Model for Stepping

The anatomical model has been previously described by Francis and Feinstein (7). Step height was determined by measuring the length of the femur in centimeters and multiplying by the constant 0.71. Length of the femur was determined by measuring the distance from the head or greater trochanter to the lateral tubercle of the femur. The constant represented 1 - cos 73°. The hip angle of 73° was used because previous reports of other step test designs indicated that the highest correlation between heart rate response from stepping and measurements of VO2peak occurred when the subjects stepped at platform heights resulting in hip angles closest to 73° (7, 17). The step platform was constructed so it could be raised or lowered in 1.0-cm increments.

Metabolic Determinations

Step Test. The step test developed by Francis and Feinstein (7) is a modification of the 3-min step test described McArdle et al. (10), with the addition of
the height-adjustment model and with different stepping frequencies. The protocol involved three different tests employing one platform height and stepping frequencies of 13, 15, and 17 ascents/min. Each subject stepped up and down on a portable, height-adjusted platform for 3 min to a cadence established by a metronome at either 52 clicks/min (13 ascents), 60 clicks/min (15 ascents), or 68 clicks/min (17 ascents). Immediately after 3 min of stepping, the subjects remained standing while recovery heart rate (HR) was determined. Recovery HR was determined by auscultation for 15 s, starting 5 s into recovery. All three stepping rates were performed on the same day starting with 13 ascents with a 5-min rest period between tests. This protocol was repeated on three different days with a time span of 3–5 days. Tests were performed either between 9:00 and 11:00 a.m. or between 4:00 and 5:00 p.m. and were repeated at the same time for each subject. Prior to data collection, each subject practiced the stepping procedure on a separate day.

The protocol of Francis and Feinstein (7) used stepping frequencies of 22, 26, and 30 ascents/min. It was evident in initial trials that the subjects in this study could not keep pace at these frequencies for 3 min. Practice trials indicated that the majority of subjects could keep pace for 3 min at frequencies of 13, 15, and 17 ascents/min, and these frequencies were therefore chosen for this study. These step rates were similar to those Seidl et al. (16) found reasonable to use with adult women with MR. The predicted VO\(_2\)peak from the step test was calculated from the formula VO\(_2\)peak \(= 105.4 - 1.64 \times (15\text{-s recovery HR})\) from Francis and Feinstein (7). This regression had an \(r\) of .81 and a standard error of the estimate (SEE) of 5.67 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) in their study of nondisabled children.

**Treadmill Test.** The laboratory testing for VO\(_2\)peak followed the procedures described by Fernhall and Tymeson (5) and Pitetti et al. (13). Subjects were familiarized with the laboratory setting, treadmill walking, and use of the headgear prior to data collection. The number of familiarization sessions varied from 1 to 4 sessions, depending on the subject. The subject started at an individualized walking speed between 2.0 and 3.5 mph (i.e., subject's capacity) at 0 % grade, and the grade was increased 4% every 2 min until volitional exhaustion. Each subject completed two treadmill tests, separated by 2–7 days.

Heart rates were collected each minute during the treadmill test using a Polar heart rate monitor. Subjects breathed into a Hans-Rudolph valve, and expired air was collected and analyzed by a Quinton Q-Plex metabolic system. Metabolic data were calculated and displayed in 15-s averages. The metabolic system was calibrated prior to each test. The highest VO\(_2\) attained was recorded as VO\(_2\)peak and was included in the study if two of the following criteria were met: (a) The subject was unable to keep up with the treadmill speed or became too unstable to safely continue the test, (b) a respiratory exchange ratio of >1.05, (c) an increase in HR of <2 bpm with an increase in workload, (d) an increase in VO\(_2\) of <2.0 ml \(\cdot\) kg\(^{-1}\) \(\cdot\) min\(^{-1}\) with an increase in workload. These procedures have been shown to produce valid and reliable measures of VO\(_2\)peak in children, adolescents, and adults with MR (2, 3, 4, 5, 13).

**Statistical Analysis**

Mean and standard deviations were calculated for all variables. Test reliability was assessed through test-retest correlations. The relationship between recovery heart rate and VO\(_2\)peak was evaluated through Pearson product correlations, as an estimate of concurrent validity of the step tests. Estimated VO\(_2\)peak was calculated for
each step test employing the regression equation of Francis and Feinstein (7). The relationship between estimated and measured VO₂peak was assessed through bivariate regression. The prediction error was described by SEE and total error (TE). The TE describes systematic differences between estimated and measured values, and when the TE is substantially larger than the SEE, there is a systematic prediction error (9). The measured and predicted VO₂peak were also compared using t tests on dependent samples. The significance level was set at \( p < .05 \) throughout.

Results

Of the original 24 subjects, 22 completed the 13 and 15 ascents/min, and 19 completed the 17 ascent/min test. Therefore, the descriptive statistics shown in Table 1 are from the 22 subjects who completed at least two of the tests. The 2 subjects who did not complete any of the step tests both had Down syndrome, as did 2 of the 3 subjects who could not complete the 17 ascents/min. For the 4 Down syndrome subjects, physical reasons were the limiting factors. For the subject without Down syndrome, motivation seemed to be the limiting factor.

The reliability of the step test varied with stepping rate. The reliability coefficients for the varying rates were .48, .79, and .83 for 13, 15, and 17 ascents/min, respectively. The reliability coefficients for the 15 and 17 ascents/min were significant (\( p < .05 \)), but the 13 ascents/min was not significant. Utilizing the approach by Francis and Feinstein (7), we correlated the recovery heart rate of the step tests with VO₂peak as a measure of concurrent validity. These correlations are shown in Table 2. Both the 15 and 17 ascents/min tests correlated significantly with VO₂peak. We also applied the regression equation provided by Francis and Feinstein (7) to calculate estimated VO₂peak, and compared those values to actual VO₂peak. These results are shown in Table 3. The regression significantly overestimated the measured VO₂peak, and the estimated VO₂peak correlated marginally with the measured values for 15 and 17 ascents/min (\( p < .05 \)). The SEEs were large and approached 30% of the measured mean values. The TE was almost double the SEE.

Table 1  Means and Standard Deviations for Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>( M )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>14.2</td>
<td>2.1</td>
</tr>
<tr>
<td>Height (inches)</td>
<td>62.3</td>
<td>5.7</td>
</tr>
<tr>
<td>Weight (lb)</td>
<td>130.7</td>
<td>41.5</td>
</tr>
<tr>
<td>BMI</td>
<td>22.8</td>
<td>6.8</td>
</tr>
<tr>
<td>VO₂peak (ml · kg⁻¹ · min⁻¹)</td>
<td>30.65</td>
<td>9.2</td>
</tr>
<tr>
<td>VEpeak (L · min⁻¹)</td>
<td>62.6</td>
<td>23.5</td>
</tr>
<tr>
<td>RER</td>
<td>1.09</td>
<td>0.07</td>
</tr>
<tr>
<td>Peak HR (bpm)</td>
<td>181</td>
<td>11</td>
</tr>
</tbody>
</table>

Note. \( N = 22 \). BMI = body mass index; RER = respiratory exchange ratio; HR = heart rate.
Table 2  Correlations Between VO\(_2\)peak and Recovery Heart Rate During Step Testing at 13, 15, and 17 Ascents/min

<table>
<thead>
<tr>
<th>Variable</th>
<th>VO(_2)peak</th>
<th>13 a/m (n = 22)</th>
<th>15 a/m (n = 22)</th>
<th>17 a/m (n = 19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO(_2)peak</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 a/m (n = 22)</td>
<td>-0.31</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 a/m (n = 22)</td>
<td>-0.48*</td>
<td>-0.88**</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>17 a/m (n = 19)</td>
<td>-0.46*</td>
<td>0.88</td>
<td>0.94</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note. a/m = ascents/min.
*p < .05, **p < .01.

Table 3  Measured Versus Estimate VO\(_2\)peak

<table>
<thead>
<tr>
<th>VO(_2)peak</th>
<th>Measured VO(_2)peak</th>
<th>Estimated VO(_2)peak 13 a/min</th>
<th>Estimated VO(_2)peak 15 a/min</th>
<th>Estimated VO(_2)peak 17 a/min</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>9.2</td>
<td>8.6</td>
<td>7.9</td>
<td>7.6</td>
</tr>
<tr>
<td>r</td>
<td>0.19</td>
<td>0.48*</td>
<td>0.46*</td>
<td></td>
</tr>
<tr>
<td>SEE</td>
<td>10.2</td>
<td>8.7</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>TE</td>
<td>20.8</td>
<td>18.2</td>
<td>16.1</td>
<td></td>
</tr>
</tbody>
</table>

*Note. VO\(_2\)peak is measured in ml · kg\(^{-1}\) · min\(^{-1}\). Values were determined with the formulas of Francis and Feinstein (7).
*aMeasured VO\(_2\)peak < estimated VO\(_2\)peak at all step rates (p < .05).
*b*p < .05.

Discussion

In the past decade a number of studies have been conducted to measure the fitness levels of adults with mental retardation (13). Results of these studies indicated a physical fitness level of this population consistent with adults with a sedentary lifestyle. In an effort to assist organizations concerned with the welfare of adults with MR, researchers have developed simple field test protocols that could help identify those adults with MR in need of fitness training/rehabilitation (13). One of those field tests involved a step test used in the CSTF that was modified for adults with MR (12). However, applying this test to children and adolescents with MR is problematic because of assumptions regarding stepping efficiency. The formula used in estimating aerobic fitness utilizes the oxygen cost of each stepping rate measured on non-MR adult males and females with high levels of aerobic fitness (8). Additionally, the adults in the Montgomery et al. (12) study could not keep pace with the lowest cadence of the CSTF step test and, therefore, were al-
allowed to set their own pace at a lower cadence. Furthermore, children have a higher oxygen cost of work than adults (15).

A valid step test in estimating cardiovascular fitness developed specifically for children and adolescents involving an adjustable step height that was related to body stature was reported by Francis and Feinstein (7). Therefore, the purpose of this study was to determine the feasibility, reliability, and validity of the step test developed by Francis and Feinstein (7) when applied to children and adolescents with EMR and TMR. It became evident following initial trials that a cadence of 22, 26, and 30 ascents/min was too fast for the subjects in this study. When the step cadence was lowered to a frequency of 13, 15, and 17 ascents/min, 22 of the 24 subjects could complete the step test at 13 and 15 ascents/min, and 19 subjects could perform the test at all three frequencies. These step rates were almost identical to those used by Seidl et al. (16) for female adults with MR. Seidl et al. also found that their subjects had difficulty completing faster step rates. Significant reliability coefficients for recovery heart rates in the present study were seen at 15 and 17 ascents/min, with the highest coefficient at 17 ascents/min ($r = .83$). This is similar to test reliability for HR reported in other studies of subjects with MR (3, 13, 14). However, it should be noted that the reliability for measured VO$_2$ or workload is usually greater than that of HR, probably because of lesser variability (3).

Although significant relationships were seen between recovery heart rate at the end of stepping for 15 and 17 ascents/min and peak oxygen consumption, they were not high enough to be considered valid indicators of VO$_2$peak. Furthermore, estimated VO$_2$peak using the regression equation for 22 ascents/min (the lowest step cadence they used) provided by Francis and Feinstein (7) significantly overestimated VO$_2$peak of our subjects (see Table 3). The relationship between estimated and measured VO$_2$peak was also low, although significant. However, the TE was almost twice the SEE, indicating that there was a systematic error in the prediction of VO$_2$peak from the formulas of Francis and Feinstein (7). Thus, our results on children with MR differed substantially from those reported by Francis and Feinstein (7) on children without MR. However, other researchers have reported results similar to those found in this study. For example, Metz and Alexander (11) reported correlation coefficients of only .52 in 12- to 13-year-olds and .42 in 14- to 15-year-olds between VO$_2$max values and those of the Harvard Step Test, but they did not evaluate the TE of the prediction.

It is possible that auscultation measurement of recovery HR might have been inaccurate and, therefore, might have affected the results of this study. To determine if the heart rate measured by auscultation represented the actual heart rate at the end of a 3-min step test, heart rate monitors (Polar) were worn by subjects on 45 tests. Mean recovery HR for actual (HR monitor) and estimated (auscultation method) are found in Table 4. The correlation coefficients indicate that the auscultation method of determining recovery HR was representative of the actual heart rates of the subjects.

Interestingly, the 2 subjects who could not perform the test at any cadence are both individuals with Down syndrome. Although both had low VO$_2$peak as measured on the treadmill (22.0 and 29.6 ml · kg$^{-1}$ · min$^{-1}$), there were subjects with similar VO$_2$peak levels who were able to keep pace. Instead, visual acuity and judgment of distance and space seemed to be the limiting factors. That is, even at the beginning of the test, when aerobic fitness should have little effect on keeping the proper pace, these subjects were hesitant and guarded in their steps. This be-
behavior was also observed as they ascended and descended the stairs leading to the laboratory. Of the 3 subjects who could not keep pace at 17 ascents/min, 2 also had Down syndrome. For the 2 subjects with Down syndrome, fatigue seemed to be the limiting factor, whereas motivation (i.e., failure to concentrate on keeping pace) prevented the subject without Down syndrome from keeping pace. Thus, only 1 of the 24 subjects did not complete the tests for motivational/behavioral reasons.

Another factor that could have affected the results of this study involves motivation of subjects during the treadmill test. In order to control for this, subjects were given ample time to practice the test and become familiar with staff and protocol. Subjects were pushed to exhaustion, and several physiological criteria were used to ensure good peak efforts. Additionally, the treadmill test was repeated on 2 different days with the highest VO$_2$peak chosen for data processing. These procedures have been shown to produce both valid and reliable peak values for children, adolescents, and adults with MR (2, 3, 4, 5, 12, 13, 14). Therefore, it is felt that the VO$_2$peak reached by the subjects in this study represents their best effort. Accordingly, the peak HR must also be indicative of their best effort. However, the peak heart rate seen in this study (181 ± 11 bpm) is lower than age-predicted maximal HR (206 ± 3 bpm). A lower than expected HR is a consistent finding in populations with MR (2, 3, 4, 5, 13), even when a plateau in VO$_2$ has been established. Furthermore, the testing procedures used in this study have produced valid and reliable peak HR data in both children and adults with MR (3, 4, 5). This suggests that the subjects in this study had low HR reserves, and it is possible that their recovery HRs following the step test were truncated and this decreased variability may have affected the correlations. However, the recovery HRs were similar in this group of MR children to those reported by Francis and Feinstein (7) for children without MR, suggesting that the step test provided a greater relative workload for the children with MR.

Another possible factor that could have affected the results of this study concerns variety of stepping efficiency. Seidl et al. (16) demonstrated that women with MR were much less efficient at stair climbing than women without MR. Therefore, standardized values of oxygen expenditure used in stepping tests may be of questionable validity when applied to people with MR. Indeed, a recent

Table 4  Mean and Standard Deviations of Actual Versus Estimated Recovery Heart Rates

<table>
<thead>
<tr>
<th></th>
<th>13 a/m</th>
<th>15 a/m</th>
<th>17 a/m</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 45)</td>
<td>(n = 45)</td>
<td>(n = 45)</td>
</tr>
<tr>
<td>Actual HR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(HR monitor)</td>
<td>130 17</td>
<td>136 17</td>
<td>142 18</td>
</tr>
<tr>
<td>Estimated HR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>asculation</td>
<td>130 17</td>
<td>136 18</td>
<td>142 18</td>
</tr>
<tr>
<td>r</td>
<td>.94</td>
<td>.96</td>
<td>.95</td>
</tr>
</tbody>
</table>

Note. a/m = ascents/min. HR = heart rate.
study by Thomas et al. (18), which measured the oxygen consumption of 121 men and women (ages 16–67 years) while performing the CAFT step test, concluded that variation in oxygen demand of a stepping task can account for a large portion of the error in predicting \( \text{VO}_2\text{max} \) from a submaximal stepping test. Therefore, the ability to predict the oxygen demand of stepping is also limited in populations without MR. The results of this study (18) could explain the variability of the results in our present study, as well as others, and render step-test protocols of questionable validity when estimating aerobic fitness.

In conclusion, the step test used in this study proved feasible and reliable, but although it was statistically valid, the \( r \) value was not high enough for this test to be considered a valid indicator of \( \text{VO}_2\text{peak} \). Furthermore, the \( \text{SEE} \) was too large (approximately 30% of the mean) to justify the use of this step test to predict \( \text{VO}_2\text{peak} \).

References


---

**Physical Activity and Public Health:**
**A Postgraduate Course on Research Directions and Strategies**
**September 16–23, 1997**

Sponsored by the University of South Carolina Prevention Center and the U.S. Centers for Disease Control and Prevention, this 8-day course is directed primarily at postdoctoral personnel and is designed to develop research competencies related to physical activity and public health. A panel of leading researchers on physical activity and public health will make up the faculty of the course. The course will take place at Sea Pines Conference Resort, Hilton Head Island, South Carolina. Space is limited. *Application deadline: May 16, 1997.*

For more information, contact the University of South Carolina School of Public Health. Tel: 803/777-7291; fax: 803/777-8422.