Maximal oxygen consumption in college students is reliable following four consecutive trials

OSCAR ROMERO-FALLAS, MARVIN SOTO-ARIAS & JOSE MONCADA-JIMENEZ
Human Movement Sciences Laboratory, School of Physical Education and Sports, University of Costa Rica
COSTA RICA

Published online: December 25, 2012
(Accepted for publication November 25, 2012)

DOI:10.7752/jpes.2012.04069;

Abstract
The purpose of the study was to determine the reliability of the maximal oxygen consumption (VO$_{2\text{max}}$) following four consecutive trials with 10 min rest between trials using the Bruce protocol. Participants were 16 apparently healthy college students (7 females, 9 males), who performed the Bruce protocol and after reaching their VO$_{2\text{max}}$ stepped down the treadmill and rested seated on a chair during 10 min. These procedures were repeated three consecutive times. Results indicated no significant differences in mean VO$_{2\text{max}}$ values between trials ($p > 0.05$), although gender differences were found (Females = 40.77 ± 1.95 ml/kg/min vs. Males = 50.77 ± 1.72 ml/kg/min; $p = 0.002$). Reliability was determined by the intraclass correlation coefficient (ICC), indicating a high reproducibility in the VO$_{2\text{max}}$ values between trials ($r_{\text{ICC}} = 0.96$, CI 95% = 0.92, 0.98). No adverse effects were reported during the 10 min rest between trials. In conclusion, resting 10 min between maximal Bruce protocol treadmill tests does not affect VO$_{2\text{max}}$ in young healthy participants. It is feasible and reliable to perform maximal treadmill tests in a single session without adverse effects for the participant.

Key Words: reliability, maximal oxygen consumption, Bruce Protocol, cardiovascular, college students.

Introduction
Graded exercise tests are commonly used to determine maximal aerobic power (i.e., maximal oxygen consumption [VO$_{2\text{max}}$]). The VO$_{2\text{max}}$ is used as a physiological index for general cardiovascular health and is also used for monitoring training among athletes (Cooper, Baker, Tong, Roberts, & Hanford, 2005). This parameter is defined as the highest velocity an individual can use O$_2$ during exercise (Akkerman, Brussel, Hulzebos, Vanhees, Holders, & Takken, 2010). Theoretically, the VO$_{2\text{max}}$ is reached when there is a plateau in spite of a change in workload (Akkerman et al., 2010).

The criteria used to stop a test and to determine VO$_{2\text{max}}$ is controversial. Some authors suggest new criteria, recommendations or methodologies (Midgley & Carrol, 2009, Midgley, McNaughton, Polman, & Marchant, 2007); however, no single consensus exist at this time. Howley, Bassett and Welch (1995), reviewed criteria to define VO$_{2\text{max}}$ and concluded that “the plateau in the oxygen consumption with an increased workload is the main criteria to achieve VO$_{2\text{max}}$” (p. 1292).

Midgley, Carroll, Marchant, McNaughton and Siegler (2009), suggested that when the VO$_2$ plateau is not evident, then secondary criteria must be used when participant has given a maximal effort. Two secondary criteria commonly used are the respiratory exchange ratio (RER) and heart rate (HR) (Midgley et al., 2009). For instance, Miller, Dougherty, Green and Crouse (2007), used a combination of these criteria to determine VO$_{2\text{max}}$. Participants had to meet at least two out of three possible criteria: a) a plateau in the oxygen consumption with an increased workload; b) to reach the age-predicted maximal HR (HR$_{\text{max}}$); and/or c) RER ≥ 1.15. Similarly, Chia and Aziz (2008), determined VO$_{2\text{max}}$ when participants reached an RER ≥ 1.05; 95% HR$_{\text{max}}$, and/or the participant reached exhaustion. Based on the previous studies, it is evident that some criteria used to determine VO$_{2\text{max}}$ converge and other does not.

Treadmill protocols are commonly used to determine VO$_{2\text{max}}$ because large muscle mass are involved during the tests. The Bruce protocol has been extensively used and is considered a standard test (Miller et al., 2007). VO$_{2\text{max}}$ can be obtained with an indirect calorimetry system along with this protocol, which consists of 8 stages of 3 min of duration. Treadmill speed and incline grade are also part of the stages of this protocol. The Bruce protocol has been used from pediatric to elderly populations (Zijp, Berg, Willensen, Stam, Tibboel, & Ijsselstijn, 2009; Fielding, Frontera, Hughes, Fisher, & Evans, 1997); however, reliability studies are lacking. Reliability allows researchers to repeatability or consistency of values derived from a test, essay or measurement among the same individuals (Hopkins, 2000).

Other reports have estimated reliability of a treadmill-determined VO$_{2\text{max}}$. For instance, Amorim, Byrne and Hills (2009), studies female children, who performed three tests, two of them the same day separated by a 7 h resting period. The authors did not find significant mean differences in the VO$_{2\text{max}}$ between tests, concluding that
treadmill-measured VO\textsubscript{2max} was stable and reproducible (i.e., reliable). The Bruce protocol has also been under scrutiny. Fielding et al. (1997), measured adult women who performed five graded exercise tests on a treadmill using the Bruce protocol. However, a longer resting period (at least 7 days) was given to the participants. The authors reported reliable VO\textsubscript{2max} values between trials.

Protocols using the animal model are also found in the literature. Evans and Rose (1998), determined VO\textsubscript{2max} in horses in three treadmill tests, with no less than 24 h resting period between trials. The researchers also reported reliable VO\textsubscript{2max} values between trials. Similarly, Copp, Davis, Poole and Musch (2009), used the rat model in five treadmill tests. For this study, animals were given 9 to 10 days rest between trials and results indicated non-significant differences in VO\textsubscript{2max} between trials.

Following a careful and extensive literature search, no studies regarding VO\textsubscript{2max} measurement the same day and having short resting periods between trials (e.g., 10 min) were found. A reduced resting time between trials might be useful for instance when technical problems arise during a graded exercise test, when a subject does not give the best effort during a test or when reprogramming is difficult. Therefore, the purpose of the study was to determine the VO\textsubscript{2max} reliability of the Bruce protocol when given 10 min rest between four consecutive trials to college students.

Material & methods

Participants

Participants were 16 (7 females, 9 males) Physical Education students from University of Costa Rica. All were briefed about the purpose of the study and volunteered to participate after reading and signing an informed consent and a medical history questionnaire (American College of Sports Medicine [ACSM], 2009).

Measurement instruments

Participants performed the Bruce protocol on a treadmill (Cosmed, model T-150, Italy). VO\textsubscript{2} was measured by indirect calorimetry on a metabolic cart Quark b\textsuperscript{2} (COSMED, Italy). The oxygen and carbon dioxide sensors were calibrated using gases with known concentrations before each test (CO\textsubscript{2} = 5%, O\textsubscript{2} = 16%, Balance de N\textsubscript{2}). The flow sensors were also calibrated before each test using a 3-L syringe. Heart rate was measured using a telemetric device (Polar®).

Procedures

Participants were given appointments to the Human Movement Sciences Laboratory for initial screening. Medical history and informed consent were recorded and only apparently healthy participants were allowed to participate. Participants were instructed to consume a light breakfast at least 2 h before testing and refraining from strenuous exercise the at least 24 h previous to the tests.

Participants were allowed to warm up on the treadmill for 5 min at a speed of 5 km/h and 0% grade. Following this, the Bruce protocol was performed. The protocol starts at 2.7 km/h and 10% grade. After 3 min, treadmill speed and grade changes to 4.0 km/h and 12 %, then 5.5 km/h and 14 %, 6.8 km/h and 16 %, 8.0 km/h and 18 %, 8.9 km/h and 20 %, 9.7 km/h and 22%, and lastly to 10.5 km/h and 24% grade (ACSM, 2009). Indications to stop the test were to achieve at least two of the following criteria: a) a request to stop the test, b) RER ≥ 1.15, and/or c) a plateau of the VO\textsubscript{2} curve < 2 ml·kg\textsuperscript{-1}·min\textsuperscript{-1} with increased workload (Moncada-Jiménez et al., 2010; Moncada-Jiménez et al., 2009; Moncada-Jiménez et al., 2009).

Once the first test was finished, a cool-down period was allowed for the subject after achieving 80% of the age predicted HR\textsubscript{max}. Following this, the subject was instructed to sit on a comfortable chair for 10 min and to rinse his/her mouth with plain water if necessary. No beverages or foods were allowed to consume during the resting time. After 10 min rest, participants were instructed to stand up and perform the same Bruce protocol following the same protocol described before.

Statistical analysis

Data were analyzed with the Statistical Package for the Social Sciences (SPSS®), version 15.0 for Windows. Data are presented as means (\textit{M}) and standard deviation (+ SD), unless otherwise noted. Mixed factorial 2 (gender) x 4 (trials) repeated measures analyses of variance (ANOVA) were computed to examine dependent variables VO\textsubscript{2max}, HR\textsubscript{max}, the ventilatory equivalent ratio for oxygen and carbon dioxide (VE/VO\textsubscript{2}), and oxygen pulse (VO\textsubscript{2}/HR). Appropriate multiple comparisons Bonferroni post hoc analyses were computed when significant ANOVA interactions or main effects were found. Statistical significance was set \textit{a priori} at \(p \leq 0.05\). The intraclass correlation coefficient (ICC) and the 95% confidence interval (CI\textsubscript{95%}) were computed to determine reliability of the VO\textsubscript{2max} values (Hopkins, 2000).

Results

Nineteen healthy college students participated in the study. Mean (+ SD) age, weight and height are shown in table 1.
Table 1. Descriptive statistic for the participants (Mean ± SD).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Females (n = 7)</th>
<th>Males (n = 9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr.)</td>
<td>22.7 ± 2.6</td>
<td>21.0 ± 1.3</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>55.0 ± 10.6</td>
<td>61.2 ± 3.8</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>159.6 ± 9.1</td>
<td>170.5 ± 3.4</td>
</tr>
<tr>
<td>VO₂ trial 1 (L·min⁻¹)</td>
<td>2.3 ± 0.5</td>
<td>3.1 ± 0.3</td>
</tr>
<tr>
<td>VO₂ trial 2 (L·min⁻¹)</td>
<td>2.2 ± 0.4</td>
<td>3.1 ± 0.3</td>
</tr>
<tr>
<td>VO₂ trial 3 (L·min⁻¹)</td>
<td>2.2 ± 0.4</td>
<td>3.1 ± 0.3</td>
</tr>
<tr>
<td>VO₂ trial 4 (L·min⁻¹)</td>
<td>2.2 ± 0.5</td>
<td>3.2 ± 0.4</td>
</tr>
<tr>
<td>VO₂max trial 1 (ml·kg⁻¹·min⁻¹)</td>
<td>41.6 ± 6.0</td>
<td>50.2 ± 5.5</td>
</tr>
<tr>
<td>VO₂max trial 2 (ml·kg⁻¹·min⁻¹)</td>
<td>39.9 ± 5.8</td>
<td>51.0 ± 4.8</td>
</tr>
<tr>
<td>VO₂max trial 3 (ml·kg⁻¹·min⁻¹)</td>
<td>40.6 ± 4.1</td>
<td>50.3 ± 4.7</td>
</tr>
<tr>
<td>VO₂max trial 4 (ml·kg⁻¹·min⁻¹)</td>
<td>41.0 ± 6.5</td>
<td>51.6 ± 5.4</td>
</tr>
</tbody>
</table>

Mixed factorial 2 x 4 ANOVA revealed no significant interaction between gender and trials in the VO₂max values (p = 0.219). Follow up analyses showed no significant trial effect (p = 0.466); however, main effect gender reached statistical significance (p = 0.002). Independently of the trial performed, mean VO₂max values were higher for males (M = 50.77 ± 1.72 ml⁻¹·kg⁻¹·min⁻¹; CI₉⁵% = 47.10, 54.45 ml⁻¹·kg⁻¹·min⁻¹) than for females (M = 40.77 ± 1.95 ml⁻¹·kg⁻¹·min⁻¹; CI₉⁵% = 36.60, 44.95 ml⁻¹·kg⁻¹·min⁻¹).

Mixed factorial 2 x 4 ANOVA revealed no significant interaction between gender and trials in HRmax values (p = 0.130). Follow up analyses showed no significant gender effect (p = 0.474); however, main effect trial reached statistical significance (p ≤ 0.001). Post hoc analyses indicated that the mean HRmax value for the first trial was 180.72 ± 2.12 beats·min⁻¹, for the second trial was 184.42 ± 2.83 beats·min⁻¹, for the third trial was 186.12 ± 2.08 beats·min⁻¹, and for the last trial was 187.67 ± 2.00 beats·min⁻¹ (Fig. 1).

Mixed factorial 2 x 4 ANOVA revealed no significant interaction between gender and trials in VE/VCO₂ (p = 0.075). However, a significant (p ≤ 0.001) main effect trial was found. Post hoc analyses revealed that mean VE/VCO₂ during the first trial (M = 29.06 ± 0.66) was significantly smaller (p ≤ 0.001) than mean VE/VCO₂ for trials two, three and four (Mtrial 2 = 31.12 ± 0.83; Mtrial 3 = 31.42 ± 0.86; Mtrial 4 = 31.71 ± 0.88) (Fig. 2).
Mixed factorial 2 x 4 ANOVA revealed no significant interaction between gender and trials in the VO₂/HR ($p = 0.577$). However, a significant ($p \leq 0.001$) main effect trial was found. Post hoc analyses revealed that mean VO₂/HR during the first trial ($M = 11.36 \pm 0.48 \text{ ml·min}^{-1}$) was significantly smaller ($p < 0.01$) than mean VO₂/HR for trials two, three and four ($M_{\text{trial } 2} = 10.63 \pm 0.45 \text{ ml·min}^{-1}$; $M_{\text{trial } 3} = 10.44 \pm 0.44 \text{ ml·min}^{-1}$; $M_{\text{trial } 4} = 10.57 \pm 0.49 \text{ ml·min}^{-1}$) (Fig. 3).

Finally, reliability analysis indicated highly reproducible VO₂max values among trials ($r_{\text{ICC}} = 0.96$; $CI_{95\%} = 0.92$, 0.98).
Discussion
The purpose of the study was to determine the reliability of the VO$_{2\text{max}}$ following four consecutive trials with 10 min rest between trials using the Bruce protocol. In this study with apparently healthy college students the Bruce protocol showed a high reliability in the VO$_{2\text{max}}$ attained in the four trials. No other studies were found in the literature that used short resting periods (e.g., 10 min) between trials; some reported resting periods ranging from 7 h to more than 7 days, and in the studies reviewed a high reliability in the VO$_{2\text{max}}$ was reported (Amorin et al., 2009; Fielding et al., 1997; Evans & Rose, 1998; Capp et al., 2009).

The study of reliability consists of repeating a measurement a reasonable number of times in an adequate number of individuals in a given time (Hopkins, 2000). In the present study 16 subjects participated and each one of them performed four trials in a single day (64 measurements total), which reduces variation (e.g., resting, previous exercise, food, ambient temperature, hydration, etc.) if the tests were performed in different days. According to Hopkins, variability within a subject is the most important source of variation studied by researchers to measure reliability because this variation affects estimate precision in the change of a variable in an experimental study. A statistics that captures that notion of random variability of tests performed by an individual is the standard deviation, also known as the typical error. Such an error is influenced by several sources, and one of them is usually when the subject is tested on different pieces of equipment or different test administrators. In the present study, the same measurement equipment was used and every single test was administered by the same operator. Therefore, in this study all possible sources of random variation indicated by Hopkins were accounted for.

The heart rate response showed significant differences between trials in this study. Specifically, the HR attained in the first trial was different from the HR of the next three trials. Fielding et al. (1997), measured adult females and found similar and non-statistically significant HR values between exercise trials. This is likely to be explained by the resting between trials (7 days). According to Wilmore, Costill and Kenney (2008), HR is a direct result of exercise intensity and the heart effort to meet the body energy requirements. Therefore, recovery from exertion is highly dependent of the subject’s fitness and recovery time provided. In the present study only 10 min recovery time was allowed for the participants and significant changes in HR response were observed after the first trial.

The VO$_{2\text{max}}$ is described in Fick’s equation (Q (CaO$_2$ − CvO$_2$)), where Q is the cardiac output of the heart, CaO$_2$ is the arterial oxygen content, and CvO$_2$ is the venous oxygen content (Wilmore et al., 2008). This is also known as the arterial-venous O$_2$ difference (A-VO$_2$). Q is defined as the total volume of blood pumped by the heart, in particular by a left or right ventricle in the time interval of one minute. It is also defined as the product between HR and the stroke volume (SV). In this study, HR was measured and increased significantly after the first trial, and SV (i.e., the volume of blood ejected by the heart during a heartbeat) was not directly measured; however, oxygen pulse (VO$_2$/HR) was recorded. VO$_2$/HR is considered a direct measure of SV (González-Garcia & Maldonado-Gómez, 2000; Padilla, Martínez, Olvera, Ojeda-Cruz & Caudillo-Pérez, 2000).

The A-VO$_2$ is determined by the arterial and venous O$_2$ content exchange at the cellular level (Wilmore et al., 2008). The association between the different gases during a test is analyzed by the metabolic cart. The VE/VCO$_2$ is defined as the ventilatory equivalent ratio for carbon dioxide (i.e., number of breaths/min divided by the CO$_2$ released from the body) (Arena, Humphrey & Peberdy, 2003). In this study, VE/VCO$_2$ increased significantly after the first trial and did not differ between trials two, three and four (Figure 3). This is explained by the fact that during strenuous exercise the CO$_2$ production and PCO$_2$ increases, resulting in a lower body pH. Consequently, ventilation (i.e., hyperventilation) and sodium bicarbonate production increases to buffer the excess of hydrogen ions (H$^+$). In other words, the cardiovascular system is less efficient to mobilize air in to consume each ml of O$_2$.

Another indicator of cardiac efficiency is the VO2/HR. This parameter might explain why the VO2max remained unchanged during the four trials of the present study. The VO2/HR was reduced significantly only after the first trial (Figure 2). The heart’s ability to extract O$_2$ at each heartbeat was reduced (i.e., reduced A-VO$_2$), which was immediately compensated by a higher HR also required to maintain an appropriate Q and finally the observed same VO$_{2\text{max}}$ between trials.

Conclusions
In conclusion, VO$_{2\text{max}}$ obtained by four consecutive Bruce protocol trials with 10 min rest between trials remained unchanged in spite of a higher cardiovascular effort. It is possible to obtain reliable VO$_{2\text{max}}$ values when apparently healthy college students are given 10 min rest between four consecutive maximal exercise tests with no adverse effects.

Conflicts of interest: none.

References

480


