Reproducibility of cardiorespiratory and performance responses to exercise in hypoxia

Ben J Lee, Charles D Thake

Background. Knowledge of the variance in physiological performance measures between repeated trials can inform whether familiarization sessions are necessary prior to intervention studies. The purpose of this study was to assess the reliability of cardiorespiratory and exercise performance measures during both steady state hypoxic exercise and a preloaded 16.1 km hypoxic time trial. Methods. Eighteen male participants (age, 22 ± 4 years; height, 1.77 ± 0.04 meters; body mass, 76.8 kg; estimated body fat and VO₂peak = 3.50 ± 0.60 L.min⁻¹) were divided into three groups. Reliability of responses (HR, S₉O₂, VO₂, VCO₂, VE and respiratory exchange ratio; RER) to the HST (F(IO₂) 0.14; 15 minutes rest, 60 minutes cycling at 50% normoxic VO₂peak) was assessed across 3 repeat trials (HST 1, 2 and 3, n = 6). Reliability of the preloaded time trial (pTT; 15 min rest, 40 minutes cycling at 50% normoxic VO₂peak, 16.1km time trial) was assessed across 3 repeat normoxic (N; F(IO₂) ≈ 0.21; n=6) and 3 repeat hypoxic (F(IO₂) ≈ 0.14; n = 6) trials. All exercise trials were undertaken at the same time of day, following exercise and dietary controls, 7 days apart. Results. Intra-class correlation coefficients (ICC’s) for mean and peak HR, SpO₂, Vₑ , VO₂, VCO₂ and blood lactate within each trial were improved from HST1 to HST2 (mean data: 0.99, 0.95, 0.75, 0.62, 0.70, 0.90; peak data: 0.98, 0.96, 0.64, 0.69, 0.74, 0.75) to HST2 and HST3 (ICC = 0.99, 0.97, 0.82, 0.85, 0.87 and 0.96 respectively). The reliability for time to pTT completion was improved following one trial, and the CV (test 2 vs. 3) was similar under normoxic (CV = 0.62) and hypoxic conditions (CV = 0.63). Conclusion. Cardiorespiratory responses to the HST were reproducible and the pTT performance time reliable in both normoxia and hypoxia. Since the reproducibility of the measurements in HST trials and reliability of pTT improved between the second and third trials, two familiarization visit are recommended prior to employing these protocols in future studies.
Title: Reproducibility of cardiorespiratory and performance responses to exercise in hypoxia

Authors: Lee BJ\textsuperscript{1,2} and Thake CD\textsuperscript{2}

\textsuperscript{1}: Department of Health, University of Bath, Bath, United Kingdom.

\textsuperscript{2}: Centre for Applied Biological and Exercise Sciences, Coventry University, Coventry, United Kingdom.

Correspondence.

Ben Lee, Ph.D
The University of Bath
Department of Health
Bath
BA2 7AY

Telephone: +44 (0)1225 383029

Email: B.J.Lee@bath.ac.uk
Abstract

Background. Knowledge of the variance in physiological performance measures between repeated trials can inform whether familiarization sessions are necessary prior to intervention studies. The purpose of this study was to assess the reliability of cardiorespiratory and exercise performance measures during both steady state hypoxic exercise and a preloaded 16.1 km hypoxic time trial. Methods. Eighteen male participants (age, 22 ± 4 years; height, 1.77 ± 0.04 meters; body mass, 76.8 kg; estimated body fat and VO₂ peak = 3.50 ± 0.60 L.min⁻¹) were divided into three groups. Reliability of responses (HR, SpO₂, VO₂, VCO₂, V̇ and respiratory exchange ratio; RER) to the HST (FＩO₂ 0.14; 15 minutes rest, 60 minutes cycling at 50% normoxic VO₂ peak) was assessed across 3 repeat trials (HST 1, 2 and 3, n = 6). Reliability of the preloaded time trial (pTT; 15 min rest, 40 minutes cycling at 50% normoxic VO₂ peak, 16.1 km time trial) was assessed across 3 repeat normoxic (N; FＩO₂ ≈ 0.21; n=6) and 3 repeat hypoxic (FＩO₂ ≈ 0.14; n = 6) trials. All exercise trials were undertaken at the same time of day, following exercise and dietary controls, 7 days apart. Results. Intra-class correlation coefficients (ICC’s) for mean and peak HR, SpO₂, V̇, VO₂, VCO₂ and blood lactate within each trial were improved from HST1 to HST2 (mean data: 0.99, 0.95, 0.75, 0.62, 0.70, 0.90; peak data: 0.98, 0.96, 0.64, 0.69, 0.74, 0.75) to HST2 and HST3 (ICC = 0.99, 0.97, 0.82, 0.85, 0.87 and 0.96 respectively). The reliability for time to pTT completion was improved following one trial, and the CV (test 2 vs. 3) was similar under normoxic (CV = 0.62) and hypoxic conditions (CV = 0.63). Conclusion. Cardiorespiratory responses to the HST were reproducible and the pTT performance time reliable in both normoxia and hypoxia. Since the reproducibility of the measurements in HST trials and reliability of pTT improved between the second and third trials, two familiarization visit are recommended prior to employing these protocols in future studies.

Introduction

Many modern sporting scenarios and military operations involve exposure to extreme environmental factors (e.g. altitude or heat) that increase physiological strain and reduce physical
work capacity and performance (Gore et al., 1996; 1997). Knowledge of the extent that specific environmental conditions impact on these response variables can inform the preparation and management of exposed individuals and groups. Therefore it is important to consider the utility of tests used to measure the magnitude of response to defined environmental stressors.

Accordingly, a reasonable degree of confidence that any changes in outcome measures are due to an imposed intervention (e.g. acclimation) rather than measurement error (Atkinson and Nevill, 1998) or inherent high-test variability (Che Jusoh et al., 2015) is important. Fixed workload protocols are often used to provide ‘steady-state’ data in response to a stimulus e.g. the hypoxic stress test used by Lee and colleagues (2014, 2014a) to evaluate the impact of short term heat acclimation on responses to exercise in hypoxia. It is generally assumed that such protocols yield reproducible physiological data under defined conditions although the test-retest coefficient of variation in these measures is seldom reported.

In order to provide an indicator of performance potential, steady-state fixed workload protocols may be extended by continuing the exercise bout until exhaustion. However this approach often suffers from high test – retest variability and low ecological validity. For example, trained cyclists had a test-retest coefficient of variation (CV) of ~27% when cycling to exhaustion at 75% of maximal workload ($W_{max}$; Jeukendrup et al., 1996). Whereas, a performance test (a self-paced 60-minute bout of cycling), conducted with the same participants, returned a test-retest CV of 3.4% (Jeukendrup et al., 1996). This is typical of self-paced tests with a defined end-point (i.e. time trials) that by definition provide more ecologically valid data (Mee et al., 2015; Che Jusoh et al., 2015). Furthermore, under heat stress a comparable CV (3.6%) was reported by Che Jusoh and colleagues (2015) following a 45 minute cycling preload (cycling at 55% $\dot{VO}_2$peak). A similar performance test has been used in hypoxic conditions (Beidleman et al., 2008) although no reliability data was presented for the hypoxic performance test. As hypoxic tolerance before and after an intervention is now commonly reported (Lee et al., 2014; Lee 2014a; Gibson 2015), knowledge of the reproducibility of physiological steady-state conditions during a fixed-work protocol hypoxic stress test (HST), as well as the reliability of a preloaded self-paced cycling time trial cycling test (pTT) could assist data interpretation and study design. Therefore, the purpose of this study was to assess the reproducibility of physiological measures collected during a steady state HST (60 minute fixed load period of cycling) and the reliability of a preloaded (40 minute bout of fixed-intensity cycling) 16.1km TT conducted in moderate normobaric hypoxia.
Methods

Participants

The study received ethical approval from the Coventry University local Ethics Committee (Reference number P2566/P6420) and was conducted in accordance with the declaration of Helsinki. Eighteen healthy and regularly active males (playing team sports at least 3 times per week for a period of over 1 year) accustomed to cycling exercise provided their informed consent to participate. Three groups of 6 participants were formed. Reproducibility of physiological measurements to a fixed relative work-rate hypoxic stress test (HST) where made on the HST group (age, 23 ± 3 years; height, 1.77 ± 0.05 m; body mass, 74.7 ± 6.9 kg; \( \dot{V}O_2 \text{peak} \), 3.50 ± 0.70 L.min\(^{-1}\)); whereas the reliability of a 16.1 km preloaded time trial (pTT) performance test was evaluated using separate groups under normoxic (\( F_{I_O_2} \approx 0.21 \)) conditions (NORM; age, 21 ± 2 years; height, 1.77 ± 0.04 m; body mass, 84.0 ± 17.0; \( \dot{V}O_2 \text{peak} \), 3.17 ± 0.40 L.min\(^{-1}\)), and hypoxic (\( F_{I_O_2} \approx 0.14 \)) conditions (HYP; age; 27 ± 5 years; height, 1.76 ± 0.02 m; body mass, 72.0 ± 7.6 kg; \( \dot{V}O_2 \text{peak} \), 3.82 ± 0.60 L.min\(^{-1}\)).

Experimental design

Each participant visited the laboratory on 4 occasions (1 preliminary and 3 experimental) each separated by at least 7 days (range 7 – 11 days). During the preliminary visit participants gave informed consent and measures of height, body mass, skin folds (4 sites), normoxic lactate threshold and \( \dot{V}O_2 \text{peak} \) were made. Thereafter, on separate visits, participants in the HST group completed 3 fixed work-rate (50% normoxic \( \dot{V}O_2 \text{peak} \), 60 minutes) exercise trials whilst breathing hypoxia (\( F_{I_O_2} \approx 0.14 \)), 6 participants completed 3 preloaded time-trials whilst breathing \( F_{I_O_2} \approx 0.21 \), and a further 6 participants completed 3 preloaded time-trials under hypoxic conditions (\( F_{I_O_2} \approx 0.14 \)).

Preliminary testing

Anthropometric data were collected in accordance with the International Society for the Advancement of Kinanthropometry (ISAK) guidelines (Marfell-Jones et al., 2006). Lactate threshold and peak oxygen consumption (\( \dot{V}O_2 \text{peak} \)) were determined via an incremental exercise
test to volitional exhaustion on calibrated SRM cycle ergometer (Schoberer Rad Meßtechnik, Welldorf, Germany). Resting blood lactate (BLa; Biosen C-Line analyser, EKF Diagnostics, Germany) was determined from a finger capillary whole blood sample following a 10 minute seated rest period. The test began at a workload of 70W for 4-minutes and was then increased by 35W every 4 minutes until a blood lactate value of > 4mmol.L⁻¹ was reached. Thereafter, workload increased 35W every 2 minutes until volitional exhaustion. A cadence of 70 revolutions per minute (RPM) was maintained throughout. Expired gas was collected during the last minute of each exercise stage via a two-way nonrebreathable valve (Harvard Ltd, Eldenbridge, UK) and 1.5 m of 30mm diameter polyvinyl tubing supplying a 200L Douglas bag (Cranlea & Co, Birmingham, UK). Heart rate (Polar FT1, Polar Electro OY, Kempele, Finland) was recorded and perceived exertion (Borg, 1976) sought within the last 10 seconds of each gas collection. Respiratory gas analysis for determination of expired oxygen and carbon dioxide fractions was completed using a Servomex infrared and paramagnetic gas analyzer (model 1400, Servomex, Crowthorne, UK) respectively, and gas volume via a Harvard dry gas meter (Cranlea and Company, Birmingham, UK). Lactate threshold was calculated according to the Dmax method (Cheng et al. 1992). VO₂peak was considered to be achieved if two of the following criteria were met: i) a respiratory exchange ratio of >1.1, ii) a heart rate greater than 95% of age predicted maximum (220–age) and iii) a final blood lactate value in excess of 8 mmol.mL⁻¹ (Winter et al. 2006).

Experimental trials

Laboratory attendance time was consistent for each participant in order to minimize the effects of circadian variation on performance (Drust et al., 2005). Participants were requested to abstain from caffeine (Luo et al., 2008) and alcohol consumption for 72 hours prior to each laboratory visit and required to maintain a food and activity diary as accurately as possible for 3 days prior to each experimental visit (Morton et al., 2006). Participants refrained from all supplementation (i.e. vitamins, ergogenic aids) throughout the study period and were requested to abstain from prolonged thermal exposures (baths, saunas, steam rooms, and tanning devices) and vigorous physical activity for seven days prior to each laboratory visit, and moderate physical activity for 72 hours prior to each visit.
Participants adhered to an overnight fast (Febbraio et al., 2002) and consumed 500 ml of plain water one hour before visiting the laboratory in accordance with the American College of Sports Medicine position stance on hydration (Sawka et al., 2007). Upon arrival (0630 – 0730 hrs), participants voided their bladder to provide a sample for hydration assessment via urine specific gravity (USG; Atago Refractometer, Jencons Pls, Leighton Buzzard, UK) and urine osmolarity (U_osmo; Advanced 3300 Micro-Osmometer, Advanced Inc, Massachusetts, USA). Euhydration was assumed for urine specific gravity values of ≤ 1.020 g·ml⁻¹ and osmolarity values of ≤ 700 mOsm·kg⁻¹ (Armstrong et al., 1994). This control was not violated by any participant during any trial. Following this, participants measured their own nude body mass (Seca 880, Seca, Hamburg, Germany), and fitted a telemetric heart rate monitor around their chest (Polar FT1, Polar Electro OY, Kempele, Finland).

**Fixed work-rate hypoxic stress test (HST)**

Participants in the HST group completed 15 minutes of seated rest in normoxic conditions prior to baseline data collection for all physiological variables. Following normoxic rest, a 15 minute resting ‘wash-in’ period of breathing hypoxic gas (F_\text{O}_2 = 0.14) was then completed. The gas was delivered to the participant via a mouthpiece and 30 mm diameter connector (Harvard Ltd, Eldenbridge, UK) connected to a two-way non-rebreathable valve (Harvard Ltd, Eldenbridge, UK). Ethylene clear vinyl tubing was used to connect the inspiratory side of the valve to a series of 1000L Douglas bags which were filled via a hypoxicator (Hypoxico HYP123 Hypoxicator, New York, USA). At the end of the wash-in period participants began 60 minutes of cycling exercise at an intensity corresponding to 50% normoxic V\text{O}_2\text{peak} (145 ± 16 Watts) on an SRM ergometer (Schoberer Rad Meßtechnik, Welldorf, Germany) at a cadence of 70 RPM. Measures of heart rate (HR), arterial haemoglobin oxygen saturation (Sp\text{O}_2), respiratory variables (\dot{V}_E, \dot{V}_O_2, \dot{V}CO_2) via Douglas bag collections, and fingertip capillary blood samples for determination of glucose and lactate concentrations were taken at the start and end of each 15 minute resting period and every 10 minutes throughout exercise. Perceptual ratings of perceived exertion (RPE; Borg 1976) and thermal sensation (TS) were also noted at 10 minute intervals during the exercise period.

**Preloaded time trial (pTT)**
Participants undertook 30 minutes of seated rest (normoxia $F_tO_2 = 0.21$ throughout; hypoxia 15 min $F_tO_2 = 0.21$ followed by 15 minutes $F_tO_2 = 0.14$ ‘wash in’ period) prior to undertaking a 40 minute period of cycling exercise at an intensity corresponding to 50% normoxic $\dot{VO}_2$peak (144 ± 18 Watts) whilst breathing either $F_tO_2 = 0.21$ or $F_tO_2 = 0.14$. Physiological and perceptual variables were collected every 10 minutes as previously described. At the end of the 40 minute fixed load exercise bout and after 5 minutes of passive recovery the self-paced 16.1km cycling pTT was commenced. The pTT was controlled within the SRM ergometers open-ended mode, which creates a braking force that has a cubic relationship with speed, mimicking the effect of air resistance on a moving bicycle. Participants were instructed to complete distance as quickly as possible. The only feedback available to participants was the distance completed at any given time. During the time trial only measures of HR and power output were collected in order to avoid providing any external time cues.

Statistical analysis

Data are reported as mean ± SD unless otherwise stated. Physiological data were analysed using trial x time ANOVA, with repeated measures over time. In order to assess test-retest reproducibility in the mean physiological measures obtained at rest and during exercise throughout HST, normoxic pTT and hypoxic pTT, the mean difference between trials (change in mean), intra-class correlation coefficient (ICC) and coefficient of variation (CV) was calculated between the 3 repeat conditions in each group (Hopkins, 2000). The technical error of measurement (TEM) between trials was calculated for all participants by calculating the difference scores and standard deviation of difference scores between trials and dividing the square root of the standard deviation of the difference by 2 (Hopkins, 2000). In order to calculate the smallest worthwhile change and therefore the smallest detectable change in a measured variable, the TEM was multiplied by 1.5 (Hopkins, 2000).

Results

All participants adhered here to the activity and nutrition controls specified in the participant information sheet prior to each testing session.

Physiological and perceptual responses to the HST
The cardiovascular responses to each HST are shown in Figure 1. Based on participants mean oxygen consumption, participants were exercising at an intensity eliciting 62 ± 11, 58 ± 9 and 61 ± 16% \( \dot{V}O_2 \) peak at the end HST1, HST2 and HST3 respectively. No main effect for trial (\( p > 0.05 \)) was observed for heart rate, \( S_pO_2, \dot{V}E, \dot{V}O_2, \dot{V}CO_2 \), and RER over the course of the 3 trials. The ICC, technical error of measurement, and coefficient of variation for mean and peak exercise measurements are shown in Table 1. The data indicate that acceptable reliability is achieved after 1 familiarization for cardiovascular responses (mean exercise and peak HR and \( S_pO_2 \)). Respiratory variables in hypoxic conditions are more variable, with improved reliability seen after 2 familiarization sessions for both the mean exercise and peak measurements (Figure 2). Perceptual ratings of perceived exertion averaged 14 ± 1, 13 ± 1, and 13 ± 1, and Thermal sensation averaged during 4.8 ± 0.4, 4.4 ± 0.6 and 4.5 ± 0.6 during HST1, HST2 and HST3 respectively.

**Physiological and perceptual responses to the preload period**

Based on participants mean \( \dot{V}O_2 \) and HR participants in the normoxic group were exercising at an intensity eliciting 51.1 ± 8%, 51.3 ± 11 and 52 ± 10% \( \dot{V}O_2 \) peak, and participants in the hypoxic group at 67 ±, 65 ± 9 and 65 ± % for pTT1, pTT2, and pTT3 respectively. Main effects for time (\( p < 0.0001 \)) but not trial (\( p > 0.05 \)) or time x trial interaction (\( p > 0.05 \)) were observed for HR, \( S_pO_2, \dot{V}O_2, \dot{V}E \) and RER during the 40 minute preload, with reliability indices mirroring those observed during the prolonged 60 minute HST (data not shown). Mean and peak physiological and perceptual data during the preload period are presented in Table 2.

**Time trial performance**

Physiological data from the time trial are shown in Figure 3. Based on mean power output during each TT, the normoxic group completed the distance at 63 ± 5%, 63 ± 6% and 66 ± 9% and the hypoxic group at 60 ± 6%, 62 ± 6% and 61 ± 6% of normoxic \( W_{max} \) for TT1, TT2 and TT3 respectively. Time to complete the 16.1km course was longer in HYP compared to NORM during TT1 (42.3 ± 5.5 versus 39.4 ± 1.1 minutes), TT2 (41.4 ± 5.5 versus 39.1 ± 1.5 minutes) and TT3 (41.8 ± 5.3 versus 38.8 ± 0.6 minutes), although no TT x condition effect was observed (\( p = 0.250 \)). From Table 3 it can be seen that the reliability of the performance time, typically the
main outcome measure obtained from a time trial, was improved when 2 familiarization trials are 
performed. Data regarding the smallest worthwhile change for all variables measured during the 
HST and pTT are shown in Table 4.

Discussion

This study reports that physiological responses to a 60 minute steady-state HST are reproducible 
and that preloaded (40 minutes of steady state cycling) 16.1km TT performance, conducted 
under both normoxic and acute normobaric hypoxic conditions (FiO₂≈0.14), is reliable.
Importantly these data demonstrate that the reproducibility of the measurements in HST trials 
and reliability of pTT performance (under both normoxic and hypoxic conditions) improved 
between the second and third trials. Accordingly when using such protocols in future 
investigations we would advise conducting at least one familiarisation trial where participants 
carry out the full protocol under the proposed experimental conditions, with 2 familiarization 
protocols reducing the technical error of measurement further. Furthermore an awareness of the 
typical variance of each of these tests under the conditions studied can help provide context to 
the effect of an intervention on defined outcome measures.

During each HST trial classic markers of hypoxic physiology (HR, SpO₂, VO₂, VCO₂, VE, RER 
and blood lactate) were measured. Our data indicate that this simple to conduct fixed-workload 
test had good agreement for all the physiological variables studied (Table 1) when tests are 
conducted one week apart, supporting its use in future studies investigating hypoxic tolerance.
Typically, a correlation coefficient of over 0.90 is considered high, 0.70-0.80 moderate, and 
below 0.70 considered too low for a reliable and sensitive physiological test (Vincent, 1995). In 
the present investigation the ICC for each physiological variable was over 0.80 when HST2 and 
HST3 were considered. Based on this observation, the whole body physiological measurements 
of HR, SpO₂, VO₂, VCO₂, RER and VE all have an acceptable level of reproducibility between 
repeated trials in the conditions studied if two familiarisation visits are conducted. With 
reference to the smallest worthwhile change, calculated from the TEM, we suggest that changes 
in mean exercise HR, SpO₂, VE and VO₂ of 2 - 3 beats.min⁻¹, 1%, 3.6 - 4.7 L.min⁻¹, 0.24 – 0.32 
L.min⁻¹ respectively, would constitute a statistically meaningful change in these variables
between HST tests conducted before and after an intervention (e.g. an acclimation period) provided 2 familiarization sessions are completed.

In addition to the fixed workload HST we also examined the reliability of a 16.1km TT conducted immediately following a 40-minute, submaximal preload. This type of experimental design allows researchers to assess both steady state responses, and incorporate a measure of physical performance. As the TEM and CV for normoxia and hypoxia are similar, this indicates that any familiarisation sessions can be completed in normoxic conditions and reduce the potential for conferred acclamatory effect prior to subsequent hypoxic exposure. It should be noted that, while not trained cyclists, all participants in the present investigation were not naive to laboratory testing protocols. To avoid further confounds care was taken to remove the influence of diurnal variation, previous activity, and diet, as these may all effect endurance performance (Atkinson et al., 2005; Rauch et al., 2005). Like other time trial procedures, feedback in the form of distance completed was provided (Jeukendrup et al., 1996; Tyler and Sutherland et al., 2008; Biedleman et al., 2014). This increases the external validity of the test and allows participants to use a pacing strategy to a known end-point (Che Jusoh et al., 2015). One criticism levelled at the use of fixed intensity steady-state preload periods is that by the time participants begin the self-paced performance period they are already near to their physiological limit (Che Jusoh et al., 2015). To minimise this potential confound a workrate of 50% normoxic $\dot{V}O_2$ max was selected as pilot data indicated this allowed the participants to perform exercise below both the normoxic and hypoxic lactate threshold. At the end of the preload period, HR was 140, 142, and 140 for the normoxic group and 156, 156, and 158 for the hypoxic group, which would suggest that while this 40 minute period presents a moderate physiological challenge, participants began the time trial with considerable capacity to increase physiological strain, i.e., from 156 beats.min$^{-1}$ to 180 beats.min$^{-1}$ at the end of the TT. The smallest worthwhile change for performance time following one experimental visit (TT1 and TT2) was between 0:59 and 1:19 minutes and 1:09 and 1:26 minutes for the normoxic and hypoxic time trials. However if two familiarization trials are conducted prior to an experimental trial the smallest worthwhile change is reduced to 0:35 – 0:47 minutes (normoxia) and 0:32 – 0:43 minutes (hypoxia), indicating that a significant change in performance time induced by an intervention would need to be > 43 seconds to be meaningful.
Conclusion

Performance is a key primary outcome variable commonly used when assessing the efficacy of interventions. Thus separating real treatment effects from the typical variance associated with a test and/or participants is important yet seldom reported. Additionally, knowledge of the variance in physiological and performance measures between repeated trials is essential to inform the number of familiarization trials required prior to exposing participants to an intervention. The results of this study show that for moderately trained healthy males, both a steady state HST and a preloaded 16.1km time trial conducted on an SRM ergometer are highly reproducible when performed under laboratory conditions. Using these protocols, practitioners and researchers can determine whether common interventions (e.g. acclimation) reduce physiological strain (HST/preload) during steady state work conditions and improve physical performance (TT time) within a test. In order to reduce the known error and remove confounding learning effects associated with these protocols we recommend 2 familiarization sessions be performed prior to employing these protocols in future studies.

Acknowledgements

The authors wish to thank all participants for their time and effort expended completing these trials. We are also grateful to Mr Roy Petticrew and Susie Wilson for their excellent technical assistance throughout data collection.

References


Table 1 (on next page)

Measures of reliability for mean exercise and peak physiological variables during the HST ($n = 6$).
Table 1. Measures of test - retest reliability for mean exercise and peak physiological variables during the HST ($n = 6$).

<table>
<thead>
<tr>
<th></th>
<th>HST1 – HST 2</th>
<th></th>
<th></th>
<th>HST2 – HST 3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>TEM</td>
<td>CV</td>
<td>ICC</td>
<td>TEM</td>
<td>CV</td>
</tr>
<tr>
<td>Mean HR (bts·min$^{-1}$)</td>
<td>0.99 (0.93 – 0.99)</td>
<td>1.6</td>
<td>0.91</td>
<td>0.99 (0.96 – 0.99)</td>
<td>1.7</td>
<td>0.90</td>
</tr>
<tr>
<td>Peak HR (bts·min$^{-1}$)</td>
<td>0.98 (0.89 – 0.99)</td>
<td>2.44</td>
<td>1.15</td>
<td>0.97 (0.81 – 0.99)</td>
<td>3.92</td>
<td>1.84</td>
</tr>
<tr>
<td>Mean SpO$_2$ (%)</td>
<td>0.95 (0.65 – 0.99)</td>
<td>0.91</td>
<td>0.81</td>
<td>0.97 (0.73 – 0.99)</td>
<td>0.50</td>
<td>0.61</td>
</tr>
<tr>
<td>Peak SpO$_2$ (%)</td>
<td>0.96 (0.69 – 0.99)</td>
<td>0.83</td>
<td>1.15</td>
<td>0.83 (0.50 – 0.97)</td>
<td>1.44</td>
<td>1.37</td>
</tr>
<tr>
<td>Mean $\dot{V}_E$BTPS (L·min$^{-1}$)</td>
<td>0.75 (0.74 – 0.97)</td>
<td>3.92</td>
<td>4.84</td>
<td>0.82 (0.59 – 0.97)</td>
<td>2.37</td>
<td>2.90</td>
</tr>
<tr>
<td>Peak $\dot{V}_E$BTPS (L·min$^{-1}$)</td>
<td>0.64 (-2.7 – 0.95)</td>
<td>6.20</td>
<td>2.23</td>
<td>0.97 (0.83 – 0.99)</td>
<td>1.25</td>
<td>1.53</td>
</tr>
<tr>
<td>Mean $\dot{V}$O$_2$ (L·min$^{-1}$)</td>
<td>0.62 (-3.5 – 0.95)</td>
<td>0.11</td>
<td>4.10</td>
<td>0.85 (-19 – 0.98)</td>
<td>0.16</td>
<td>5.64</td>
</tr>
<tr>
<td>Peak $\dot{V}$O$_2$ (L·min$^{-1}$)</td>
<td>0.69 (-0.61 – 0.95)</td>
<td>0.17</td>
<td>5.75</td>
<td>0.74 (0.72 – 0.95)</td>
<td>0.22</td>
<td>8.06</td>
</tr>
<tr>
<td>Mean $\dot{V}$CO$_2$ (L·min$^{-1}$)</td>
<td>0.70 (-2.2 – 0.96)</td>
<td>0.12</td>
<td>5.65</td>
<td>0.87 (0.15 – 0.98)</td>
<td>0.10</td>
<td>2.41</td>
</tr>
<tr>
<td>Peak $\dot{V}$CO$_2$ (L·min$^{-1}$)</td>
<td>0.74 (-0.41 – 0.96)</td>
<td>0.09</td>
<td>8.61</td>
<td>0.85 (0.33 – 0.97)</td>
<td>0.19</td>
<td>12.04</td>
</tr>
<tr>
<td>Mean BLa (mmol$^{-1}$)</td>
<td>0.90 (0.30 – 0.99)</td>
<td>0.62</td>
<td>15.0</td>
<td>0.96 (0.72 – 0.99)</td>
<td>0.39</td>
<td>8.7</td>
</tr>
<tr>
<td>Peak BLa (mmol$^{-1}$)</td>
<td>0.75 (-0.29 – 0.96)</td>
<td>0.65</td>
<td>19.2</td>
<td>0.98 (0.85 – 0.99)</td>
<td>0.26</td>
<td>3.04</td>
</tr>
</tbody>
</table>

1 ICC = Intraclass correlation coefficient (lower 95% and upper 95% confidence interval. TEM = technical error of measurement. CV = coefficient of variation.

2 HR = heart rate. BTPS = Body temperature pressure saturated.
**Table 2** (on next page)

Measures of reliability in normoxic (n = 6) and hypoxic (n = 6) conditions during the time trial. Data are calculated from the mean TT power output (Watts) and heart rate.
Table 2. Mean ± SD for mean peak exercise physiological, respiratory and perceptual responses during the 40 minute preload period in normoxia (n = 6) and hypoxia (n = 6).

<table>
<thead>
<tr>
<th></th>
<th>Normoxia</th>
<th></th>
<th></th>
<th>Hypoxia</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pTT1</td>
<td>pTT2</td>
<td>pTT3</td>
<td>pTT1</td>
<td>pTT2</td>
<td>pTT3</td>
</tr>
<tr>
<td>Mean HR (bts min(^{-1}))</td>
<td>138 ± 10</td>
<td>137 ± 7</td>
<td>136 ± 8</td>
<td>153 ± 8</td>
<td>154 ± 9</td>
<td>155 ± 6</td>
</tr>
<tr>
<td>Peak HR (bts min(^{-1}))</td>
<td>140 ± 12</td>
<td>142 ± 7</td>
<td>140 ± 10</td>
<td>156 ± 6</td>
<td>156 ± 9</td>
<td>158 ± 4</td>
</tr>
<tr>
<td>Mean SpO(_2) (%)</td>
<td>96 ± 1</td>
<td>97 ± 1</td>
<td>97 ± 1</td>
<td>80 ± 4</td>
<td>79 ± 4</td>
<td>80 ± 3</td>
</tr>
<tr>
<td>Peak SpO(_2) (%)</td>
<td>96 ± 1</td>
<td>97 ± 1</td>
<td>97 ± 1</td>
<td>79 ± 4</td>
<td>79 ± 3</td>
<td>80 ± 2</td>
</tr>
<tr>
<td>Mean (\dot{V}\text{E}) BTPS (L min(^{-1}))</td>
<td>52.5 ± 4.4</td>
<td>54.5 ± 5.1</td>
<td>54.3 ± 4.8</td>
<td>61.2 ± 12.9</td>
<td>61.8 ± 9.8</td>
<td>64.0 ± 6.3</td>
</tr>
<tr>
<td>Peak (\dot{V}\text{E}) BTPS (L min(^{-1}))</td>
<td>53.4 ± 3.7</td>
<td>56.7 ± 5.1</td>
<td>53.4 ± 6.8</td>
<td>63.8 ± 14.7</td>
<td>62.4 ± 8.9</td>
<td>63.1 ± 8.6</td>
</tr>
<tr>
<td>Mean (\dot{V}\text{O}_2) (L min(^{-1}))</td>
<td>2.02 ± 0.11</td>
<td>2.03 ± 0.19</td>
<td>2.04 ± 0.18</td>
<td>1.94 ± 0.36</td>
<td>1.93 ± 0.33</td>
<td>1.96 ± 0.31</td>
</tr>
<tr>
<td>Peak (\dot{V}\text{O}_2) (L min(^{-1}))</td>
<td>2.08 ± 0.10</td>
<td>2.07 ± 0.25</td>
<td>2.09 ± 0.18</td>
<td>2.07 ± 0.42</td>
<td>1.96 ± 0.33</td>
<td>1.96 ± 0.35</td>
</tr>
<tr>
<td>Mean (\dot{V}\text{CO}_2) (L min(^{-1}))</td>
<td>1.88 ± 0.21</td>
<td>1.85 ± 0.16</td>
<td>1.88 ± 0.20</td>
<td>1.92 ± 0.34</td>
<td>1.94 ± 0.30</td>
<td>1.92 ± 0.26</td>
</tr>
<tr>
<td>Peak (\dot{V}\text{CO}_2) (L min(^{-1}))</td>
<td>2.01 ± 0.23</td>
<td>1.88 ± 0.19</td>
<td>1.98 ± 0.22</td>
<td>1.96 ± 0.35</td>
<td>1.96 ± 0.26</td>
<td>1.95 ± 0.28</td>
</tr>
<tr>
<td>Mean RPE</td>
<td>13 ± 1</td>
<td>13 ± 1</td>
<td>13 ± 1</td>
<td>13 ± 2</td>
<td>12 ± 1</td>
<td>12 ± 1</td>
</tr>
<tr>
<td>Peak RPE</td>
<td>13 ± 1</td>
<td>13 ± 1</td>
<td>13 ± 1</td>
<td>14 ± 1</td>
<td>14 ± 1</td>
<td>14 ± 1</td>
</tr>
<tr>
<td>Mean TS</td>
<td>4.9 ± 0.4</td>
<td>4.8 ± 0.4</td>
<td>4.8 ± 0.1</td>
<td>4.6 ± 0.2</td>
<td>4.9 ± 0.4</td>
<td>4.8 ± 0.4</td>
</tr>
<tr>
<td>Peak TS</td>
<td>5.0 ± 0.0</td>
<td>4.8 ± 0.5</td>
<td>4.8 ± 0.5</td>
<td>5.0 ± 0.8</td>
<td>4.8 ± 0.5</td>
<td>4.8 ± 0.5</td>
</tr>
</tbody>
</table>

3 HR = heart rate. BTPS = Body temperature pressure saturated. RPE = Rating of perceived exertion. TS = Thermal sensation.
Table 3 (on next page)

Measures of reliability in normoxic (n = 6) and hypoxic (n = 6) conditions during the time trial. Data are calculated from the mean TT power output (Watts) and heart rate.
Table 3. Measures of reliability in normoxic (n = 6) and hypoxic (n = 6) conditions during the time trial. Data are calculated from the mean TT power output (Watts) and heart rate.

<table>
<thead>
<tr>
<th></th>
<th>TT1 – TT2</th>
<th></th>
<th>TT2 – TT3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>TEM</td>
<td>CV</td>
</tr>
<tr>
<td>Normoxia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.79 (-0.52 – 0.97)</td>
<td>0.66</td>
<td>0.46</td>
</tr>
<tr>
<td>Watts</td>
<td>0.91 (0.35 – 0.99)</td>
<td>5.01</td>
<td>0.45</td>
</tr>
<tr>
<td>HR</td>
<td>0.94 (0.59 – 0.99)</td>
<td>3.0</td>
<td>1.56</td>
</tr>
<tr>
<td>Hypoxia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.98 (0.86 – 0.99)</td>
<td>0.77</td>
<td>1.53</td>
</tr>
<tr>
<td>Watts</td>
<td>0.86 (-0.19 – 0.98)</td>
<td>2.30</td>
<td>2.30</td>
</tr>
<tr>
<td>HR</td>
<td>0.92 (0.47 – 0.99)</td>
<td>3.00</td>
<td>0.43</td>
</tr>
</tbody>
</table>

ICC = Intraclass correlation coefficient. TEM = technical error of measurement. CV = coefficient of variation. HR = heart rate.
Table 4 (on next page)

Reliability measures during the hypoxic tests.

Smallest worthwhile changes for mean exercise data during the HST, pTT preload period, and time trial conducted in hypoxic conditions following 2 familiarization sessions. Data are calculated as 1.5 x the TEM, and 2.0 x the TEM (Hopkins 2000).
Table 4. Smallest worthwhile changes for mean exercise data during the HST, pTT preload period, and time trial conducted in hypoxic conditions following 2 familiarization sessions. Data are calculated as 1.5 x the TEM, and 2.0 x the TEM (Hopkins 2000).

<table>
<thead>
<tr>
<th></th>
<th>Smallest worthwhile change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.5</td>
</tr>
<tr>
<td><strong>HST/pTT preload</strong></td>
<td></td>
</tr>
<tr>
<td>Mean HR (bts min⁻¹)</td>
<td>2</td>
</tr>
<tr>
<td>Mean SpO₂ (%)</td>
<td>1.0</td>
</tr>
<tr>
<td>Mean V̇eBTPS (L min⁻¹)</td>
<td>3.6</td>
</tr>
<tr>
<td>Mean V̇O₂ (L min⁻¹)</td>
<td>0.24</td>
</tr>
<tr>
<td>Mean V̇CO₂ (L min⁻¹)</td>
<td>0.15</td>
</tr>
<tr>
<td>Mean BLa (mmol⁻¹)</td>
<td>0.58</td>
</tr>
<tr>
<td><strong>Time trial</strong></td>
<td></td>
</tr>
<tr>
<td>Performance time (seconds)</td>
<td>30</td>
</tr>
<tr>
<td>Mean power output (Watts)</td>
<td>9</td>
</tr>
<tr>
<td>Mean HR (bts min⁻¹)</td>
<td>5</td>
</tr>
</tbody>
</table>
Figure 1 (on next page)

Heart rate and arterial oxygen saturation throughout the 60 minute HST.

Resting and exercise heart rate (A) and arterial oxygen saturation (B) were no different throughout the 3 HSTs. Line graphs show mean ± SD throughout rest and exercise (n = 6) and box plots display the mean exercise values, and highest and lowest value for HST1, HST2 and HST3.
Heart rate (bts.min⁻¹)

Trial 1
Trial 2
Trial 3

Time (minutes)

SpO₂ (%)

0 10 20 30 40 50 60 70 80 90 100

80 85 90 95 100

HST1
HST2
HST3
Figure 2 (on next page)

Respiratory responses to the HST

Resting and exercise ventilation (A), oxygen consumption (B), and carbon dioxide production (C) were no different between the 3 HSTs. Line graphs show the mean ± SD values and box plots represent the mean exercise values with error bars showing the highest and lowest values recorded.
Figure 3 (on next page)

Physiological responses to the preloaded time trial in normoxia and hypoxia.

Time trial completion time, mean power output throughout each kilometer, and mean heart rate during each kilometer for the normoxic (A, B, C) and hypoxic (D, E, F) time trials. Box plots display the individual performance times (dots), and highest and lowest times recorded. Line graphs display the mean ± SD responses throughout (n = 6).