The Reliability of a Single-Trial Measurement of Maximal Accumulated Oxygen Deficit Determined via Perceptually-Regulated Exercise

Mark Glaister, Adam Liddell, & Kate Estlea

Faculty of Sport, Allied Health, and Performance Sciences, St Mary's University, Strawberry Hill, Twickenham, UK

Corresponding Author:

Mark Glaister

Faculty of Sport, Allied Health, and Performance Sciences

St Mary's University

Strawberry Hill

Twickenham, UK

TW1 4SX

Tel: (+44)208 240 4012

E-mail: mark.glaister@stmarys.ac.uk

ORCID: https://orcid.org/0000-0003-4454-4854

Running head: Reliability of maximal accumulated oxygen deficit

ABSTRACT

PURPOSE: The aim of this study was to evaluate the reliability of a single-trial determination of maximal accumulated oxygen deficit (MAOD) achieved via the aid of perceptually-regulated incremental exercise. METHODS: 14 trained male cyclists (age: 45 ± 8 yrs; height: 1.82 ± 0.06 m; mass: 79.7 ± 6.7 kg; $\dot{V}O_{2max}$: 4.09 ± 0.57 $L \cdot min^{-1}$) performed three trials of a submaximal incremental cycling test followed by a test to exhaustion at 116% of predicted $\dot{V}O_{2max}$. The intensity for each stage of the incremental test was regulated by participants to elicit perceived exertion levels of 9 – 15 on the Borg (6 – 20) scale. Linear regression was used to estimate $\dot{V}O_{2max}$ at a perceived exertion level of 19. MAOD was calculated from the difference between predicted and actual oxygen demand in the test to exhaustion, reported in oxygen equivalents (O_2 eq). A separate incremental test was used to measure $\dot{V}O_{2max}$ directly. RESULTS: Correlation coefficients between perceived exertion and $\dot{V}O_2$ across trials were strong ($r \geq 0.99$), and there were no between-trial differences in predicted $\dot{V}O_{2max}$ (4.03 ± 1.04, 3.76 ± 0.53, and 3.69 ± 0.64 $L \cdot min^{-1}$, respectively; $p = 0.142$) or MAOD (2.75 ± 2.28, 2.50 ± 1.53, and 2.93 ± 1.40 L O_2 eq, respectively; $p = 0.633$). Nevertheless, the coefficients of variation for predicted $\dot{V}O_{2max}$ (14.2%) and MAOD (142.8%) were poor. CONCLUSIONS: The prediction of $\dot{V}O_{2max}$ from perceptually-regulated exercise displays a level of test-retest reliability which prevents its use as a means of evaluating MAOD reliably in a single-trial.

Keywords: anaerobic capacity; sprinting; repeatability; perceived exertion

Introduction

The measurement of anaerobic capacity has been a challenge for exercise physiologists for many years. In contrast to aerobic capacity, which is easily evaluated from the respiratory measurement of maximum oxygen uptake ($\dot{V}O_{2max}$), the evaluation of anaerobic capacity relies on complex and invasive intramuscular assessments of whole-body energy metabolism. As such, researchers have developed non-invasive methods of quantifying anaerobic capacity, the best of which is generally considered to be the maximal accumulated oxygen deficit (MAOD) (Noordhof et al., 2010).

The protocol to evaluate MAOD consists of a submaximal incremental test to establish the $\dot{V}O_2$-power output relationship, an incremental test to exhaustion to determine VO_{2max}, and a continuous supramaximal test to exhaustion (typically 2 – 5 minutes) (Medbo et al., 1988). MAOD is determined from the difference between the predicted and the actual oxygen consumption during the supramaximal exercise bout and is expressed, therefore, in terms of oxygen equivalents (O_2 eq).

Although MAOD has been reported to be able to distinguish between the anaerobic capacity of sprinters and endurance athletes (Scott et al., 1991), there are several issues that limit the routine assessment of MAOD in athletes (Noordhof et al., 2010). Aside from mixed reports about the validity and reliability of the measure (Noordhof et al., 2010), the main limitations appear to lie in: 1) the use of linear regression to establish the intensity of the supramaximal exercise bout; and 2) the practical constraints of having to carry out testing on different days due to the all-out nature of the tests to exhaustion. Regarding the former, although it has been suggested that at exercise intensities greater than lactate threshold, the presence of a $\dot{V}O_2$ slow component may affect the linearity of the submaximal $\dot{V}O_2$-power output relationship (Noordhof et al., 2010), endurance training has been shown to reduce the magnitude of the slow component

(Carter et al., 2000; Lucia et al., 2000; Saunders et al., 2003) and indeed, Muniz-Pumares et al. (2017b) found no influence of the $\dot{V}O_2$ slow component on the linearity of the response in trained cyclists. In contrast, the practicalities of evaluating MAOD via multiple trials continue to limit its routine use.

One strategy that could be used to improve the practicalities of MAOD evaluation is to estimate $\dot{V}O_{2max}$ from perceptually-regulated exercise instead of using the incremental test to exhaustion. Several studies have shown that individuals are able to self-regulate and maintain an exercise intensity corresponding to a prescribed rating of perceived exertion (RPE), and that extrapolation of the linear relationship between $\dot{V}O_2$ and RPE to an RPE of 19 (using the 15-point rating scale developed by Borg (1970)), provides a valid way of estimating $\dot{V}O_{2max}$ (Coquart et al., 2014). In effect, the use of a predicted $\dot{V}O_{2max}$ ($\dot{V}O_{2max\ PRED}$) would reduce the protocol for the assessment of MAOD to a perceptually-regulated submaximal incremental test followed by a continuous test to exhaustion. Moreover, given that MAOD is achieved at an intensity of around 112-120% of $\dot{V}O_{2max}$ (Muniz-Pumares et al., 2017a) and that the test is designed to last > 2 minutes (Medbo et al., 1988) (note: peak MAOD values have been achieved in all-out sprint tests lasting 70 s (Craig et al., 1995)), $\dot{V}O_{2max}$ could be determined from the continuous exercise bout (Poole & Jones, 2017) and used, to verify the predicted value. Therefore, the aim of this study was to evaluate the reliability of a single-trial determination of maximal accumulated oxygen deficit (MAOD) achieved via the aid of perceptually-regulated incremental exercise.

Methods

Participants

Fourteen recreationally-active male cyclists capable of maintaining a pace of at least 30 $km \cdot h^{-1}$ for 20 km (arbitrary inclusion criterion) volunteered to participate in this study. All participants received written and verbal instructions regarding the nature of the investigation and completed a short questionnaire, which indicated that all had been actively involved in cycling for approximately 13.2 ± 9.0 years. Prior to commencement, all participants provided written informed consent. Means ± standard deviation for age, height, and body mass of the participants were: 45 ± 8 years, 1.82 ± 0.06 m, and 79.7 ± 6.7 kg, respectively. Participants were instructed to avoid food and drink in the hour before each trial, and to refrain from strenuous exercise and caffeine consumption for 24 hours before each trial. Prior to the start of Trial 1, participants completed a health screening questionnaire. Participants also completed a 24 hr dietary recall and were instructed to follow that same diet prior to all trials. Ethical approval for the study was granted by St Mary's University Ethics Committee.

Procedures

Participants completed four trials in an air-conditioned laboratory maintained at a temperature of 18°C. All trials were completed on an electromagnetically-braked cycle ergometer (Lode Excalibur Sport; Groningen, Holland), at approximately the same time of day, and with at least 48 hours between trials. The cycle ergometer was fitted with clipless pedals and participants cycled using their own cycling shoes. Prior to Trial 1, the ergometer was adjusted (saddle height and handlebar position) for each participant and the settings were replicated in the remaining trials. Heart rate was monitored at 5 s intervals throughout all trials using a heart rate monitor (Polar s610i; Polar Electro Oy, Kempele, Finland), and $\dot{V}O_2$ was monitored (breath-by-breath) throughout all trials using an on-line gas analyser (Oxycon Pro; Jaeger, Hoechberg, Germany). The gas analyser was calibrated before each trial using oxygen and carbon dioxide gases of known concentrations (Cryoservice; Worcester, UK) and the flowmeter was calibrated using a 3-litre syringe (Viasys Healthcare GmbH; Hoechberg, Germany). During all trials, participants breathed room air through a facemask (Hans Rudolph; Kansas City, MO, USA) that was secured in place by a head-cap assembly (Hans Rudolph; Kansas City, MO, USA). $\dot{V}O_2$ data were filtered to remove errant breaths caused by coughing, swallowing or sighing. A breath was considered to be errant if the value was outside four standard deviations of the local mean (the two breaths preceding and following the breath of interest).

Trials 1 – 3 began with a submaximal incremental test in which participants completed four exercise bouts at intensities corresponding to perceived exertion values of 9 ('very light'), 11 ('light'), 13 ('somewhat hard'), and 15 ('hard') on the 15-point RPE scale. Participants were allowed time prior to the start of the Trial 1 to familiarise themselves with the RPE scale. Participants were given 30 s at the start of the first increment in Trial 1 to establish their preferred cadence and were instructed to maintain the same cadence in all trials. Each increment lasted 4 minutes and participants were given no feedback of elapsed time or exercise intensity. As such, the power output for each increment was adjusted by the researcher, to the required level of perceived exertion, via verbal instructions from each participant in the first minute of each increment.

On completion of each submaximal incremental test, participants were given 15 minutes of rest before completing a continuous bout of exercise to exhaustion at 116% of $\dot{V}O_{2max}$ to determine MAOD. The power output for the continuous exercise bout was determined in two stages. First, $\dot{V}O_2$ data from the submaximal incremental test were averaged over the final minute of each increment to enable the linear relationship between

$\dot{V}O_2$ and RPE to be established, which was then used to predict $\dot{V}O_{2max}$ from the $\dot{V}O_2$ at an RPE of 19 (Coquart et al., 2014). Secondly, linear regression was used on the $\dot{V}O_2$-power output relationship to predict the power output required to elicit 116% of $\dot{V}O_{2max\ PRED}$ (PO_{116}). The only information available to participants during the test to exhaustion was cadence, and verbal encouragement was provided throughout. The test was terminated when participants either stopped voluntarily or when cadence fell by more than 10 revolutions per minute.

MAOD in trials 1 – 3 was calculated from the difference between the predicted oxygen demand (determined from the predicted $\dot{V}O_2$ at PO_{116} and the time to exhaustion (TTE_{MAOD})) and the actual oxygen consumption (determined by integration of the $\dot{V}O_2$ data) during the continuous exercise bout. Peak $\dot{V}O_2$ ($\dot{V}O_{2peak}$) during the continuous exercise bout was determined as the highest 30 s average $\dot{V}O_2$ recorded during the test.

In Trial 4, participants completed an incremental test to exhaustion to determine $\dot{V}O_{2max}$. Each increment lasted 1-minute, with the starting load and the subsequent increments adjusted for each participant (based on the results from trials 1 – 3) to achieve exhaustion in approximately 10 minutes. On completion of the test a 20 μL capillary blood sample was taken from an ear lobe and analysed for blood lactate concentration using an automated analyser (Biosen C-Line; EKF Diagnostic, Ebendorfer Chaussee, Barleben, Germany). $\dot{V}O_{2max}$ was determined as the highest 30 s average $\dot{V}O_2$ recorded during the test provided that at least two of the following criteria had been met: 1) A plateau in $\dot{V}O_2$; as determined by an increase of less than 2 $ml \cdot kg^{-1} \cdot min^{-1}$ over the previous stage; 2) A respiratory exchange ratio (RER) ≥ 1.15; 3) A heart rate within 10 $b \cdot min^{-1}$ of age predicted maximum; 4) A blood lactate concentration ≥ 8 $mmol \cdot L^{-1}$.

Statistical Analysis

All statistical analyses were conducted using the Statistical Package for the Social Sciences (SPSS for Windows; IBM Corporation, Armonk, New York, USA). Measures of centrality and spread are presented as means ± standard deviation. Pearson correlations were used to investigate the $\dot{V}O_2$-RPE and the $\dot{V}O_2$-power output relationships. Between-trial differences in power output at each level of perceived exertion were evaluated using a two-way (Trial × RPE) analysis of variance (ANOVA). Differences between trials 1 – 3 in measures of $\dot{V}O_{2max\ PRED}$, $\dot{V}O_{2peak}$, PO_{116}, MAOD, and TTE_{MAOD} were evaluated using a one-way ANOVA. Reliability across trials 1 – 3 in those same measures was derived from a two-way ANOVA as coefficient of variation (CV) as described by Schabort et al. (1999), with the identity of the participants included as a random effect, and the identity of each trial as a fixed effect. Reliability was also evaluated from the ANOVA as intraclass correlation coefficients (ICC) using the method described by Bartko (1966). Confidence limits (95%) for CV and ICC were calculated using the methods outlined by McGraw and Wong (1996) and Tate and Klett (1959). Differences between measures of $\dot{V}O_{2max}$ (Trial 4) and both $\dot{V}O_{2max\ PRED}$ and $\dot{V}O_{2peak}$ (trials 1 – 3) were evaluated using one-way ANOVAs. α was set at 0.05 for all analyses. Significant effects were investigated using *post hoc* tests with Bonferroni adjustments for multiple comparisons.

Results

The relationship between $\dot{V}O_2$ and RPE was strong across trials 1 – 3 ($r = 0.998 \pm 0.014$, $r = 0.996 \pm 0.005$, and $r = 0.989 \pm 0.022$, respectively), as was the relationship between $\dot{V}O_2$ and power output ($r = 0.995 \pm 0.008$, $r = 0.998 \pm 0.003$, and $r = 0.994 \pm 0.019$, respectively). There was a significant effect of RPE on power output during each trial ($F_{(1.2,15.4)} = 444.5$, $p < 0.001$), with values increasing as RPE increased; however, there was no effect of trial ($F_{(1.1,14.7)} = 0.03$, $p = 0.885$) and no trial × RPE interaction ($F_{(1.5,20.0)} = 2.05$, $p = 0.162$) on power output (Figure 1). The results from trials 1 – 3 for measures of $\dot{V}O_{2max\ PRED}$, $\dot{V}O_{2peak}$, PO_{116}, MAOD, and TTE_{MAOD} are presented in Table 1. There were no between-trial differences in $\dot{V}O_{2max\ PRED}$ ($F_{(1.2,\ 16.0)} = 2.342$, $p = 0.142$), MAOD ($F_{(1.4,\ 17.6)} = 0.340$, $p = 0.633$), or TTE_{MAOD} ($F_{(1.2,\ 16.0)} = 1.397$, $p = 0.263$). There were, however, differences between trials in measures of $\dot{V}O_{2peak}$ ($F_{(1.4,\ 18.4)} = 4.490$, $p = 0.037$) and PO_{116} ($F_{(2,\ 26)} = 6.915$, $p = 0.004$). *Post hoc* tests for $\dot{V}O_{2peak}$ revealed a significant difference ($p = 0.025$) between trials 1 and 2 only (mean difference: 0.50 $L{\cdot}min^{-1}$; 95% likely range: 0.06 – 0.94 $L{\cdot}min^{-1}$). *Post hoc* tests for PO_{116} also revealed a significant difference ($p = 0.029$) between trials 1 and 2 only (mean difference: 35 W; 95% likely range: 3 – 67 W). There were no differences between measures of $\dot{V}O_{2max}$, determined in Trial 4 (4.09 ± 0.57 $L{\cdot}min^{-1}$), and measures of $\dot{V}O_{2max\ PRED}$ ($F_{(1.8,\ 23.8)} = 2.884$; $p = 0.08$). The percentages of $\dot{V}O_{2max}$ that participants were working at (relative to the $\dot{V}O_{2max}$ established in Trial 4) in trials 1 – 3 were: 114.1 ± 23.0, 107.3 ± 11.8, and 105.3 ± 14.9%, respectively. There were significant differences between measures of $\dot{V}O_{2peak}$ obtained in trials 1 – 3 and $\dot{V}O_{2max}$ ($F_{(3,\ 39)} = 6.639$; $p = 0.001$). *Post hoc* tests revealed that $\dot{V}O_{2peak}$ was significantly less that $\dot{V}O_{2max}$ in Trial 1 only (mean difference: 0.67 $L{\cdot}min^{-1}$; 95% likely range: 0.11 – 1.23 $L{\cdot}min^{-1}$). Moreover, whilst all participants achieved a TTE_{MAOD} of ≥ 70 s in trials 2 and 3, four participants had TTE_{MAOD} times that were less than that in Trial 1.

FIGURE 1 NEAR HERE

TABLE 1 NEAR HERE

Reliability statistics for the key outcome measures of the study are presented in Table 2. The reliability of MAOD was based on a sample size of 13, since one participant had a negative result in trials 1 and 2 (predicted oxygen consumption was less than actual oxygen consumption). Although measures of PO_{116} showed a reasonable degree of reliability, the reliability of $\dot{V}O_{2max\ PRED}$, $\dot{V}O_{2peak}$, MAOD, and TTE_{MAOD} was poor.

TABLE 2 NEAR HERE

Discussion

The aim of this study was to evaluate the reliability of a single-trial determination of maximal accumulated oxygen deficit (MAOD) achieved via the aid of perceptually-regulated incremental exercise. The key findings were that while there were no significant differences between trials in $\dot{V}O_{2max\ PRED}$, MAOD, or TTE_{MAOD}, the reliability of those measures was poor.

The MAOD values obtained in the present study are comparable to those of endurance-trained participants in some previous cycling-based studies (Foster et al., 1993; Medbø & Tabata 1993; Pouilly & Busso, 2008), though values of around 4.5 - 5.5 L O_2 eq have been reported by others (Aisbett et al., 2009; Gordon et al., 2011). While these anomalies are likely attributable to methodological differences between studies (Noordhof et al., 2010), they reinforce concerns over the validity of the measure and make it difficult to qualify what a MAOD value means for an athlete. Nevertheless, as highlighted earlier, MAOD has been shown to be able to distinguish between the anaerobic capacity of sprinters and endurance athletes (Scott et al., 1991); moreover, studies have reported changes in MAOD as a result of various experimental interventions designed to improve anaerobic capacity (Jacobs et al., 1997; Medbø & Burgers, 1990; Ravier et al., 2009; Weber & Schneider, 2002).

Previous research examining the reliability of $\dot{V}O_{2max\ PRED}$ from perceptually-regulated incremental exercise has reported test-retest reliability values that were considerably better than those of the present study (Coquart et al., 2014). It is difficult to explain why this would be, though the absence of any significant between-trial differences in any of the submaximal power outputs, or in $\dot{V}O_{2max\ PRED}$, would suggest that reliability statistics were not inflated by any learning effects. Then again, between-trial differences in PO_{116} would support the possibility of a learning effect between trial 1 and 2, and, indeed, it is generally recommended that a familiarization trial of a perceptually-regulated incremental test is performed (Coquart et al., 2014). However, familiarization has also been shown to be unnecessary in active individuals (at least in treadmill exercise) (Eston et al., 2012). Moreover, the inclusion of a familiarisation trial would have prevented MAOD from being determined in a single trial.

The differences between trials in PO_{116}, despite the good level of reliability, must have been large enough to account for the large variability in TTE_{MAOD}. Indeed, the effect of small increases in exercise intensity on time to exhaustion during continuous high-intensity exercise (hyperbolic relationship) underpins the basis for the measurement of the critical power concept (Jones et al., 2010). Moreover, time to exhaustion in a constant-intensity test typically has a large CV (Hopkins et al., 2001; Laursen et al., 2007), even when, unlike the present study, the absolute exercise intensity is consistent between trials. The magnitude of the variability in TTE_{MAOD} would also account for the poor reliability between-trials in $\dot{V}O_{2peak}$ and the fact that $\dot{V}O_{2peak}$ was significantly lower than $\dot{V}O_{2max}$ in Trial 1. It is well established that the time to reach $\dot{V}O_{2max}$ during a supramaximal exercise bout is at least 40 s even in the highest-trained individuals (Jones et al., 2010). Given the sampling time of 30 s used in the present study to determine $\dot{V}O_{2peak}$, all participants would need a TTE_{MAOD} of at least 70 s to enable $\dot{V}O_{2peak}$ to provide an accurate measurement of $\dot{V}O_{2max}$. The fact that four participants had TTE_{MAOD} times which were less than that in Trial 1 would explain the discrepancy between $\dot{V}O_{2peak}$ and $\dot{V}O_{2max}$.

Given that MAOD is reported to plateau in supramaximal exercise at 112-120% $\dot{V}O_{2max}$ it was considered that some error in the estimate of $\dot{V}O_{2max\ PRED}$ would still allow MAOD to be evaluated reliably in a single trial. As such, an exercise intensity midway between those values (PO_{116}) was chosen. However, the actual exercise intensities, relative to $\dot{V}O_{2max}$, showed considerable variability and were, on average, lower than the values based on $\dot{V}O_{2max\ PRED;}$ even though there were no significant differences between $\dot{V}O_{2max}$ and $\dot{V}O_{2max\ PRED}$.

In order to be used as a means of evaluating the effects of various interventions it is important that MAOD is reliable. Measures of $\dot{V}O_{2max}$ have a test-retest reliability of 1.3 – 7.9% (Hopkins et al., 2001). Given that the assessment of $\dot{V}O_{2max}$ is routinely used in laboratory-based tests to evaluate aerobic fitness, it would seem reasonable for MAOD to show a similar level of reliability if used for the assessment of anaerobic capacity. At present, the test-retest reliability of MAOD, assessed using traditional multi-trial approaches, is unclear, with differences in methods and statistical analyses leading to some studies showing good reliability and others not (for a full review of those issues, including guidelines to improve the validity and reliability of MAOD evaluation, see Noordhof et al. (2010)). Either way, the results of the present study show that the reliability (and, by extension, the validity) of MAOD determined in a single trial via the use of $\dot{V}O_{2max\ PRED}$ is poor and is most likely the result of a combination of errors resulting from variability in $\dot{V}O_{2max\ PRED}$ leading to errors in PO_{116} and the knock-on effect on TTE_{MAOD}.

Conclusion

Although MAOD is considered to provide the best non-invasive means of evaluating anaerobic capacity, the practical constraints associated with the traditional multi-trial assessment of MAOD, along with concerns about the validity and reliability of the measure, limit its routine use in laboratory-based fitness evaluations. Whilst the use of $\dot{V}O_{2max\ PRED}$ to replace the physical demands of a laboratory-based assessment of $\dot{V}O_{2max}$ could enable MAOD to be evaluated in a single trial, the results of the present study show that the reliability of MAOD using that approach is poor. As such, for practitioners still wishing to evaluate MAOD, it is recommended that a multi-trial approach is used.

Disclosure of potential conflicts of interest

The authors declare that they have no conflicts of interest

References

Aisbett, B., Le Rossignol, P., McConell, G. K., Abbiss, C. R., & Snow, R. (2009). Effects of starting strategy on 5-min cycling time-trial performance. *Journal of Sport Sciences, 27*(11), 1201–1209. doi:10.1080/02640410903114372

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports, 19*(1), 3-11. doi: 10.2466/pr0.1966.19.1.3.

Borg, G. (1970). Perceived exertion as an indicator of somatic stress. *Scandinavian Journal of Rehabilitation Medicine, 2*(2), 92-98. PMID: 5523831

Carter, H., Jones, A. M., Barstow, T. J., Burnley, M., Williams, C., & Doust, J. H. (2000). Effect of endurance training on oxygen uptake kinetics during treadmill running.

Journal of Applied Physiology, *89*(5), 1744-1752. doi: 10.1152/jappl.2000.89.5.1744

Coquart, J. B., Garcin, M., Parfitt, G., Tourny-Chollet, C., & Eston, R. G. (2014). Prediction of maximal or peak oxygen uptake from ratings of perceived exertion. *Sports Medicine*, *44*(5), 563–578. doi 10.1007/s40279-013-0139-5

Craig, N. P., Norton, K. I., Conyers, R. A., Woolford, S. M., Bourdon, P. C., Stanef, T., & Walsh, C. B. (1995). Influence of test duration and event specificity on maximal accumulated oxygen deficit of high performance track cyclists. *International Journal of Sports Medicine*, *16*(8), 534-540. doi: 10.1055/s-2007-973050.

Eston, R., Evans, H., Faulkner, J., Lambrick, D., Al-Rahamneh, H., & Parfitt, G. (2012). A perceptually regulated, graded exercise test predicts peak oxygen uptake during treadmill exercise in active and sedentary participants. *European Journal of Applied Physiology*, *112*(10), 3459–3468. doi: 10.1007/s00421-012-2326-8

Foster, C., Snyder, A. C., Thompson, N. N., Green, M. A., Foley, M., & Schrager, M. (1993). Effect of pacing strategy on cycle time trial performance. *Medicine & Science in Sports & Exercise*, *25*(3), 383–388. PMID: 8455455

Gordon, D., Hopkins, S., King, C., Keiller, D., & Barnes, R. J. (2010). Incidence of the plateau at $\dot{V}O_{2max}$ is dependent on the anaerobic capacity. *International Journal of Sports Medicine*, 32(1), 1-6. doi: 10.1055/s-0030-1267192

Hopkins, W. G., Schabort, E. J., & Hawley, J. A. (2001). Reliability of power in physical performance tests. *Sports Medicine*, *31*(3), 211-234. doi: 10.2165/00007256-200131030-00005

Jacobs, I., Bleue, S., & Goodman, J. (1997). Creatine ingestion increases anaerobic capacity and maximum accumulated oxygen deficit. *Canadian Journal of Applied Physiology*, *22*(3), 231-243. https://doi.org/10.1139/h97-015

Jones, A. M., Vanhatalo, A., Burnley, M., Morton, R. H., & Poole, D. C. (2010). Critical power: implications for determination of $\dot{V}O_{2max}$ and exercise tolerance. *Medicine & Science in Sports & Exercise*, *42*(10), 1876–1890. doi: 10.1249/MSS.0b013e3181d9cf7f

Laursen, P. B., Francis, G. T., Abbiss, C. R., Newton, M. J., & Nosaka, K. (2007). Reliability of time-to-exhaustion versus time-trial running tests in runners. *Medicine & Science in Sports & Exercise*, *39*(8), 1374-1379. doi: 10.1249/mss.0b013e31806010f5

Lucía, A., Hoyos, J., & Chicharro, J. L. (2000). The slow component of $\dot{V}O_2$ in professional cyclists. *British Journal of Sports Medicine*, *34*(5), 367-374. doi: 10.1136/bjsm.34.5.367

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30-46. https://doi.org/10.1037/1082-989X.1.1.30

Medbø, J. I., & Burgers, S. (1990). Effect of training on the anaerobic capacity. *Medicine & Science in Sports & Exercise*, *22*(4), 501-507. PMID: 2402211

Medbø, J. I., Mohn, A. C., Tabata, I., Bahr, R., Vaage, O., & Sejersted, O. M. (1988). Anaerobic capacity determined by maximal accumulated O_2 deficit. *Journal of Applied Physiology*, *64*(1), 50-60. doi: 10.1152/jappl.1988.64.1.50.

Medbø, J. I., & Tabata, I. (1993). Anaerobic energy release in working muscle during 30 s to 3 min of exhausting bicycling. *Journal of Applied Physiology*, *75*(4), 1654–1660. https://doi.org/10.1152/jappl.1993.75.4.1654

Muniz-Pumares, D., Pedlar, C., Godfrey, R. J., & Glaister, M. (2017a). Accumulated oxygen deficit during exercise to exhaustion determined at different supramaximal work rates. *International Journal of Sports Physiology & Performance*, *12*(3), 351-356. doi: 10.1123/ijspp.2015-0343.

Muniz-Pumares, D., Pedlar, C., Godfrey, R. J., & Glaister, M. (2017b). The effect of the oxygen uptake-power output relationship on the prediction of supramaximal oxygen demands. *Journal of Sports Medicine & Physical Fitness*, *57*(1-2), 1-7. doi: 10.23736/S0022-4707.16.05948-X.

Noordhof, D. A., De Koning, J. J., & Foster, C. (2010). The maximal accumulated oxygen deficit method: A valid and reliable measure of anaerobic capacity? *Sports Medicine*, *40*(4), 285–302. doi:10.2165/11530390-000000000-00000

Poole, D. C., & Jones, A. M. (2017). Measurement of the maximum oxygen uptake $\dot{V}O_{2max}$: $\dot{V}O_{2peak}$ is no longer acceptable. *Journal of Applied Physiology*, *122*(4), 997-1002. doi: 10.1152/japplphysiol.01063.2016.

Pouilly, J.-P., & Busso, T. (2008). Accumulated oxygen deficit during ramp exercise. *International Journal of Sports Medicine*, *29*(1), 16–20. doi:10.1055/s-2007-965068

Ravier, G., Dugué, B., Grappe, F., & Rouillon, J. D. (2009). Impressive anaerobic adaptations in elite karate athletes due to few intensive intermittent sessions added to regular karate training. *Scandinavian Journal of Medicine & Science in Sports*, *19*(5), 687–694. doi:10.1111/j.1600-0838.2008.00807.x

Saunders, M. J., Evans, E. M., Arngrimsson, S. A., Allison, J. D., & Cureton, K. J. (2003). Endurance training reduces end-exercise $\dot{V}O_2$ and muscle use during submaximal cycling. *Medicine & Science in Sports & Exercise*, *35*(2), 257-262. doi: 10.1249/01.MSS.0000048726.30184.90

Schabort, E. J., Hawley, J. A., Hopkins, W. G., & Blum, H. (1999). High reliability of performance of well-trained rowers on a rowing ergometer. *Journal of Sports Sciences*, *17*(8), 627-632. doi: 10.1080/026404199365650.

Scott, C. B., Roby, F. B., Lohman, T. G., & Bunt, J. C. (1991). The maximally accumulated oxygen deficit as an indicator of anaerobic capacity. *Medicine & Science in Sports & Exercise*, *23*(5), 618-624. PMID: 2072841

Tate, R. F., & Klett, G. W. (1959). Optimal confidence intervals for the variance of a normal distribution. *Journal of the American Statistical Association*, *54*(287), 674-682.

Weber, C. L., & Schneider, D. A. (2002). Increases in maximal accumulated oxygen deficit after highintensity interval training are not gender dependent. *Journal of Applied Physiology*, *92*(5), 1795–1801. doi:10.1152/japplphysiol.00546.2001

Figure Legends

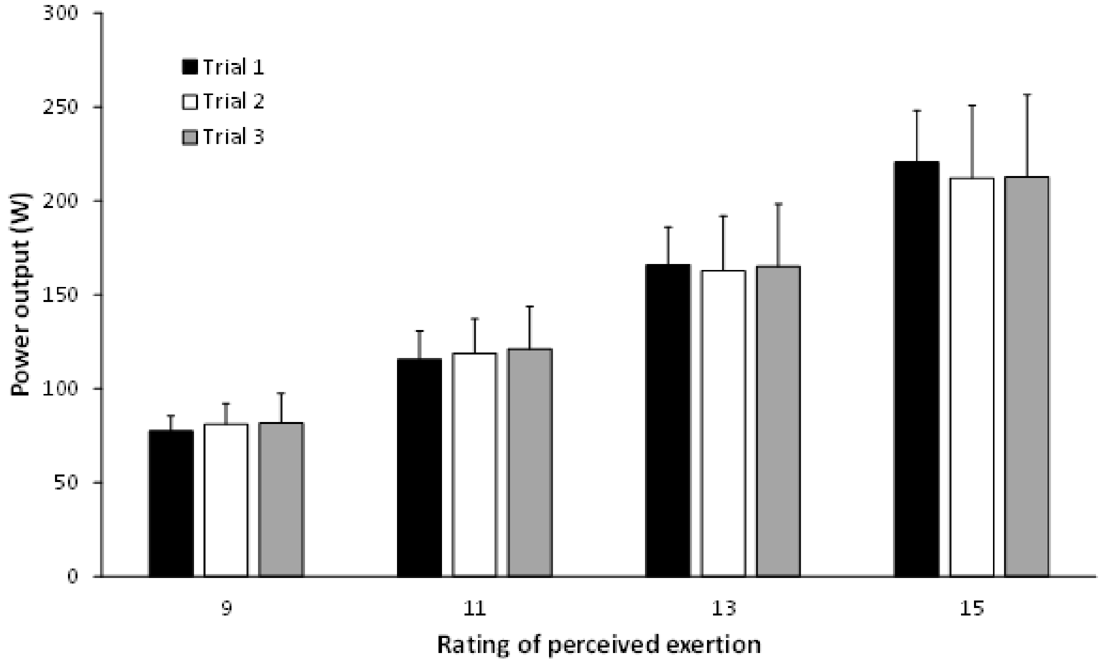

Figure 1. Power output at four different levels of effort perception during three trials of a submaximal incremental cycling test. Values are means; bars are standard deviations.

Table 1. Measures of predicted $\dot{V}O_{2max}$ and the power output required to elicit 116% of predicted $\dot{V}O_{2max}$ from three perceptually regulated incremental tests along with measures of $\dot{V}O_{2peak}$, maximal accumulated oxygen deficit, and time to exhaustion from corresponding continuous exercise bouts at 116% of predicted $\dot{V}O_{2max}$. Values are means ± standard deviations.

	$\dot{V}O_{2max\ PRED}$ (L·min^{-1})	$\dot{V}O_{2peak}$ (L·min^{-1})	PO_{116} (W)	$O_{2\ PRED}$ (L)	$O_{2\ ACC}$ (L)	MAOD (L O_2 eq)	TTE_{MAOD} (mins)
Trial 1	4.03 ± 1.04	3.42 ± 0.68	390 ± 84	10.94 ± 11.37	8.19 ± 12.99	2.75 ± 2.28	2.58 ± 3.12
Trial 2	3.76 ± 0.53	3.92 ± 0.45	355 ± 68	14.78 ± 9.58	12.28 ± 10.67	2.50 ± 1.53	3.52 ± 2.45
Trial 3	3.69 ± 0.64	3.73 ± 0.33	357 ± 75	14.51 ± 7.41	11.58 ± 7.38	2.93 ± 1.40	3.58 ± 2.08

Note: MAOD = maximal accumulated oxygen deficit; $O_{2\ ACC}$ = the actual oxygen accumulated (consumed) during the MAOD test; O_2 eq = oxygen equivalents; $O_{2\ PRED}$ = the predicted oxygen demand of the MAOD test; PO_{116} = the power output required to elicit 116% of predicted $\dot{V}O_{2max}$; TTE_{MAOD} = time to exhaustion at 116% of predicted $\dot{V}O_{2max}$; $\dot{V}O_{2max\ PRED}$ = predicted $\dot{V}O_{2max}$.

Table 2. The reliability of predicted $\dot{V}O_{2max}$ and the power output required to elicit 116% of predicted $\dot{V}O_{2max}$ across three perceptually regulated incremental tests along with reliability measures of $\dot{V}O_{2peak}$, maximal accumulated oxygen deficit, and time to exhaustion across three corresponding continuous exercise bouts at 116% of predicted $\dot{V}O_{2max}$. Values are presented as coefficients of variation and intraclass correlation coefficients with associated 95% confidence limits.

	$\dot{V}O_{2max\ PRED}$	**$\dot{V}O_{2peak}$**	**PO_{116}**	**MAOD**	**TTE_{MAOD}**
CV (%)	14.2	14.4	7.9	142.8*	59.1
CL (95%)	10.6 – 18.7	10.7 – 19.0	5.9 – 10.4	104.9 – 190.4	44.0 – 77.8
ICC	0.56	0.22	0.82	0.22*	0.48
CL (95%)	0.24 – 0.81	-0.09 – 0.59	0.62 – 0.93	-0.10 – 0.61	0.15 – 0.77

Note: CL = confidence limits; CV = coefficient of variation; ICC = intraclass correlation coefficient; MAOD = maximal accumulated oxygen deficit; PO_{116} = the power output required to elicit 116% of predicted $\dot{V}O_{2max}$; TTE_{MAOD} = time to exhaustion at 116% of predicted $\dot{V}O_{2max}$; $\dot{V}O_{2max\ PRED}$ = predicted $\dot{V}O_{2max}$; *based on a sample size of 13.