European Journal of Applied Physiology (2022) 122:2385–2392
https://doi.org/10.1007/s00421-022-05020-3

ORIGINAL ARTICLE

# A three-minute all-out test performed in a remote setting does not provide a valid estimate of the maximum metabolic steady state

Ed Maunder[1] · Jeffrey A. Rothschild[1] · Andrius Ramonas[1] · Matthieu Delcourt[2] · Andrew E. Kilding[1]

Received: 4 March 2022 / Accepted: 1 August 2022 / Published online: 10 August 2022


## Abstract

**Purpose** The three-minute all-out test (3MT), when performed on a laboratory ergometer in a linear mode, can be used to estimate the heavy–severe-intensity transition, or maximum metabolic steady state (MMSS), using the end-test power output. As the 3MT only requires accurate measurement of power output and time, it is possible the 3MT could be used in remote settings using personal equipment without supervision for quantification of MMSS.

**Methods** The aim of the present investigation was to determine the reliability and validity of remotely performed 3MTs ($3MT_R$) for estimation of MMSS. Accordingly, 53 trained cyclists and triathletes were recruited to perform one familiarisation and two experimental $3MT_R$ trials to determine its reliability. A sub-group ($N = 10$) was recruited to perform three-to-five 30 min laboratory-based constant-work rate trials following completion of one familiarisation and two experimental $3MT_R$ trials. Expired gases were collected throughout constant-work rate trials and blood lactate concentration was measured at 10 and 30 min to determine the highest power output at which steady-state $\dot{V}O_2$ (MMSS-$\dot{V}O_2$) and blood lactate (MMSS-[La$^-$]) were achieved.

**Results** The $3MT_R$ end-test power ($EP_{remote}$) was reliable (coefficient of variation, 4.5% [95% confidence limits, 3.7, 5.5%]), but overestimated MMSS ($EP_{remote}$, 283 ± 51 W; MMSS-$\dot{V}O_2$, 241 ± 46 W, $P = 0.0003$; MMSS-[La$^-$], 237 ± 47 W, $P = 0.0003$). This may have been due to failure to deplete the finite work capacity above MMSS during the $3MT_R$.

**Conclusion** These results suggest that the $3MT_R$ should not be used to estimate MMSS in endurance-trained cyclists.

**Keywords** Cycling · Threshold · Reliability · Validity · Remote

### Abbreviations

| | |
|---|---|
| 3MT | Three-minute all-out test |
| $3MT_R$ | Three-minute all-out test performed in a remote setting |
| CV | Coefficient of variation |
| $EP_{remote}$ | End-test power output in the remote three-minute all-out test |
| ES | Effect size |
| GET | Gas exchange threshold |
| MMSS | Maximum metabolic steady state |
| MMSS-[La$^-$] | Power output at the maximum metabolic steady state defined by blood lactate concentrations |
| MMSS-[$\dot{V}O_2$] | Power output at the maximum metabolic steady state defined by oxygen uptake kinetics |
| TWD | Total work done |
| $\dot{V}O_2$ | Rate of oxygen consumption |
| $\dot{V}O_2$peak | Peak rate of oxygen consumption |

Communicated by Philip D. Chilibeck.

✉ Ed Maunder
ed.maunder@aut.ac.nz

Jeffrey A. Rothschild
jeffrey.rothschild@aut.ac.nz

Andrius Ramonas
andrius.ramonas@aut.ac.nz

Matthieu Delcourt
matthieudelcourt@gmail.com

Andrew E. Kilding
andrew.kilding@aut.ac.nz

[1] Sports Performance Research Institute New Zealand, Auckland University of Technology, Auckland, New Zealand

[2] Department of Biology, University of Ottawa, Ottawa, Canada

| | |
|---|---|
| W′ | Finite work capacity above critical power |
| $WEP_{remote}$ | Work performed above the end-test power output in the remote three-minute all-out test |

## Introduction

The maximal metabolic steady state (MMSS) demarcates the transition from heavy to severe-intensity exercise (Jones et al. 2019). During severe-intensity exercise, blood lactate concentration, whole-body oxygen consumption ($\dot{V}O_2$), muscle [$H^+$], and muscle [PCr] cannot stabilise, and task failure is characterised by consistent perturbations in these values and attainment of peak $\dot{V}O_2$ ($\dot{V}O_2$peak). In the heavy domain, these muscle and whole-body parameters attain a delayed steady state (Black et al. 2017). The MMSS has been assessed using the critical power model (Jones et al. 2019). Using ~ 3–5 time-to-task failure severe-intensity trials lasting ~ 2–15 min, critical power is identified as the power-asymptote of the hyperbolic power–duration relationship, and the finite work capacity above critical power as the curvature constant (W′) (Jones and Vanhatalo 2017; Jones et al. 2019). Critical power measured in this manner has been shown to discriminate heavy and severe-intensity exercise responses, and therefore estimate the MMSS (Jones et al. 2008; Black et al. 2017). Work output at the MMSS is used in training intensity regulation, training load monitoring, and predicting endurance performance (Coyle et al. 1988; Maunder et al. 2021).

The requirement for multiple time-to-task failure severe-intensity trials led to development of the three-minute all-out test (3MT) for identification of the MMSS (Burnley et al. 2006; Vanhatalo et al. 2007). In the 3MT, an athlete works all-out for three minutes without pacing, depleting W′ in the initial part of the test to ensure the work output is eventually limited to the critical power. Accordingly, average power output during the final 30 s is used to estimate critical power, and the total work performed above the end-test power is used to estimate W′. The 3MT was designed for use on an electromagnetically braked laboratory ergometer in a linear mode, where power output is the product of the linear factor (flywheel resistance) and the square of the cadence; the linear factor is typically applied, such that the power output achieved at the individual cyclist's preferred cadence is the gas exchange threshold power output (GET) plus 50% of the interval between GET and $\dot{V}O_2$peak (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a). The athlete is also supervised and verbally encouraged by a team of researchers, and blinded to power output and time (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a). The 3MT performed in this manner has been shown to produce valid (Burnley et al. 2006; Vanhatalo et al. 2007, 2016) and reliable (Burnley et al. 2006; Johnson et al. 2011) estimates of MMSS and W′ in trained populations.

Since development of the laboratory-based 3MT in 2006, the availability of accurate and reliable power-measuring devices for use during indoor cycle training has increased dramatically (Hoon et al. 2016; Zadow et al. 2016, 2018). As the 3MT only requires accurate measurement of power output and time, it is theoretically possible that the 3MT could be used by athletes in remote settings using personal equipment without supervision for quantification of MMSS and W′. However, it is unknown if the 3MT provides reliable and valid estimates of the MMSS when performed by unsupervised endurance athletes remotely, using typical indoor training set-ups, where there is opportunity to shift gears and therefore resistance to pedalling, and view elapsed time ($3MT_R$). Addressing this gap in the literature is pertinent for remote endurance coaches operating primarily without face-to-face communication with athletes, and for endurance athletes without regular access to laboratory facilities.

Therefore, the primary aim of the present investigation was to determine the reliability and validity of the 3MT for estimation of the MMSS when performed in remote settings by unsupervised athletes using their own indoor cycling setup. It was hypothesised that the $3MT_R$ would produce reliable and valid estimates of the MMSS.

## Materials and methods

### Participants

Fifty-three trained cyclists and triathletes completed the present investigation (32 males, 21 females; age, 39 ± 9 y; self-reported training volume, 10 ± 3 h·week$^{-1}$). Prospective participants were recruited via social media, and any healthy cyclist or triathlete training > 5 h·week$^{-1}$ with access to power-measuring devices (e.g., smart indoor trainer, power pedals, and power cranks) was eligible to take part in the study. After reading a participant information sheet hosted online (Qualtrics, Provo, UT, USA), participants provided informed consent and were directed to a health screening, and, if passed, to a survey to provide details on their anthropometry, basic training history, and training equipment. The survey ended with specific instructions for how to complete the $3MT_R$ trials. Participants were able to complete the trials using their road bicycle mounted to a "rear wheel off" indoor trainer, in which the rear wheel is removed and the bike is attached to the cassette of the static trainer ($N$ = 38), or "rear wheel on" indoor trainer, in which the rear wheel is in contact with a roller ($N$ = 15). The $3MT_R$ trials were completed by participants in their home training set-ups without the researchers present. All procedures were approved by the Auckland University of Technology Ethics Committee (20/137).

## Remote trials

Participants performed the remote $3MT_R$ on three occasions: one familiarisation trial and two experimental trials, each separated by 4–10 days. In advance of the two experimental trials, participants were asked to refrain from vigorous exercise for 24 h and caffeine ingestion for 1 h and repeat any caffeine ingestion within 12 h of the first trial. Participants were asked to complete all trials using the equipment they detailed in the survey and wear a heart rate monitor throughout. These pre-trial controls were designed to simulate what a coach in remote settings could realistically achieve.

Participants were asked to warm up for 10 min at 100 W before commencing each $3MT_R$. The $3MT_R$ was a three-minute all-out effort, in which the participant was asked to produce their maximum power output at every moment of the test. Participants were able to shift gears during the trials. The participant information sheet described the expected pattern of power output vs. time, with power output first rising to a peak before steadily declining and levelling off in the second half of the test. Participants were asked to perform a self-selected cool down following each $3MT_R$. Following each trial, participants emailed output files to the researchers. Familiarisation trial files were screened to ensure that the test was completed appropriately. This included ensuring the overall power output vs. time curve matched the expected profile, and any inexplicable rises in power output late in the test that would identify pacing. In all correspondences, participants were reminded that trials were to be completed in an all-out, unpaced fashion, such that the maximum possible power output was being produced at every moment of the test. Peak, time-to-peak, mean, end-test ($EP_{remote}$; average over the last 30 s), and lowest (average over 6 and 30 s) power output were calculated for each $3MT_R$, along with total work done, work done above $EP_{remote}$ ($WEP_{remote}$), and second-by-second power output and cadence using TrainingPeaks WKO + (Peaksware, LLC, Lafayette, USA).

## Laboratory validity trials

A sub-group ($N = 10$) of locally based participants reported to the laboratory on 4–6 occasions following completion of three $3MT_R$ trials (8 males, 2 females; mass, 71 ± 13 kg; height, 178 ± 10 cm; self-reported training volume 9 ± 3 $h \cdot week^{-1}$; $\dot{V}O_2$peak, 54 ± 7 $mL \cdot kg^{-1} \cdot min^{-1}$). These participants first completed an incremental exercise test for determination of $\dot{V}O_2$ peak. Briefly, participants commenced cycling on an electromagnetically braked ergometer (Excalibur Sport, Lode BV, Groningen, The Netherlands) at 60 W. The work rate increased by 30 W every minute until task failure, with continuous collection of expired gases (TrueOne 2400, ParvoMedics, UT, USA). The $\dot{V}O_2$peak was accepted as the highest 15 s average $\dot{V}O_2$.

Participants returned to the laboratory for three-to-five constant-work rate trials 4–10 days apart, at the same time of day as the $3MT_R$ trials, having refrained from vigorous exercise for 24 h and caffeine for 1 h, and replicated their self-reported 24 h dietary intake, to identify the power output at the MMSS. These trials were completed on each participant's own equipment, and the same equipment they used to complete the $3MT_R$ trials. These trials were completed on an "ergometer mode" on a smart trainer, such that the work-rate was held constant throughout, and participants were instructed to maintain their preferred cadence. The sub-group of participants completing the validity phase of the study used either a Wahoo Kickr ($N = 9$, Wahoo Fitness, Atlanta, USA) or Tacx Neo 2 T ($N = 1$, Garmin®, KS, USA) to measure power output. The reliability and validity of the Wahoo Kickr has been established (Hoon et al. 2016; Zadow et al. 2016, 2018), whereas the Tacx Neo 2 T has to our knowledge not been validated in research; however, the data from this individual participant support credible reliability within this study (within-subject coefficient of variation for $\dot{V}O_2$ during 5–8 min of the standardised warm-up in the validity trials, 2.5%).

Constant-work rate trials began with a 10 min standardised warm-up of 50% of remote end-test power ($EP_{remote}$) for 8 min, followed by 1 min at 60% and 1 min at 70% $EP_{remote}$, after which the main 30 min trial began. Expired gas was continuously measured using a metabolic cart (TrueOne 2400, ParvoMedics, UT). Gas analysis data were initially visually inspected and aberrant points laying more than three standard deviations from the local mean were removed and a three-point moving average filter was applied to the data set. The $\dot{V}O_2$ response kinetics was modelled using exponential and linear fitting to determine the presence (or absence) of a $\dot{V}O_2$ slow component (Eq. 1). The amplitude of slow component was calculated by taking the difference between the steady-state value of the fundamental component and the average $\dot{V}O_2$ in the final 60 s of the constant-work rate trial. Steady-state $\dot{V}O_2$ was defined by a slow component amplitude less than the within-subject coefficient of variation for $\dot{V}O_2$ during minutes 5–8 of the warm-up (3.2%). Additionally, duplicate capillary blood lactate samples were obtained from a finger after 10 and 30 min, with steady-state blood lactate concentrations defined as a rise of < 1 $mmol \cdot L^{-1}$ from 10 to 30 min

$$\dot{V}O_2(t) = \dot{V}O_{2baseline} + A_p\left[1 - e^{-(t-TDp)/\tau p}\right] + S\left[t - TD_s\right], \quad (1)$$

where $\dot{V}O_2(t)$ is the absolute $\dot{V}O_2$ at a given time $t$, $\dot{V}O_2$ represents the mean $\dot{V}O_2$ during 6-8 min of the warm-up, and

$A_p$ and $\tau_p$ represent the amplitude and time constant of the primary component of $\dot{V}O_2$ kinetics, $TD_s$ is a time delay of slow component, and S (slope) is a coefficient of linear regression.

The first trial was performed at $EP_{remote}$. Subsequent trials were completed with power outputs ± 2.5% of $EP_{remote}$ until at least one trial with $\dot{V}O_2$ steady-state and non-steady-state characteristics had been performed. The mid-point of the highest power output at which a $\dot{V}O_2$ steady-state was achieved and the lowest power output exhibiting non-steady-state $\dot{V}O_2$ behaviour was accepted as the maximum $\dot{V}O_2$ steady-state (MMSS-$\dot{V}O_2$), and the mid-point of the highest power output at which steady-state blood lactate concentrations were observed and the lowest power output exhibiting non-steady-state blood lactate concentrations was accepted as the maximum lactate steady state (MMSS-[$La^-$]). The validity of $EP_{remote}$ was assessed in this manner, rather than against the laboratory-based 3MT or critical power derived from a series of severe-intensity constant-work rate trials to task failure, to provide a direct measure of the MMSS.

## Statistical analysis

Data are presented as mean ± standard deviation (SD) unless otherwise stated. Data were assessed for normality using the Shapiro–Wilk test. Simple comparisons were made using paired *t* tests (or non-parametric equivalents). The reliability of $EP_{remote}$, total work done (TWD), and work above end power ($WEP_{remote}$) was assessed using within-SD coefficients of variation (CV) and Pearson's correlation coefficients, both expressed with 95% confidence intervals. Hedges' *g* effect sizes (ES) and associated 95% confidence intervals are presented where appropriate. All statistical analyses were carried out with R version 4.0.3 (The R foundation for Statistical Computing, Vienna, Austria). Significance was inferred when $P \leq 0.05$.

# Results

## *Characteristics of the $3MT_R$*

Power output profiles were consistent between the two $3MT_R$ trials (peak power output, 660 ± 216 vs. 663 ± 238 W; time-to-peak power output, 7 ± 4 vs. 8 ± 6 s; mean power output, 316 ± 73 vs. 318 ± 75 W; $EP_{remote}$, 258 ± 60 vs. 261 ± 63 W; $WEP_{remote}$, 10.4 ± 5.0 vs. 10.1 ± 4.8 kJ; in all cases, $P > 0.05$). Using the mean of each participant's two trials, 97 ± 9% of W′ was depleted in the first 90 s of the $3MT_R$, the lowest 6 s power output was 91 ± 5% of $EP_{remote}$, and the lowest 30 s power output was 97 ± 3% of $EP_{remote}$. Cadence profiles were consistent between the two $3MT_R$ trials (peak cadence, 110 ± 21 vs. 110 ± 23 revs·min$^{-1}$; mean cadence, 90 ± 11 vs. 90 ± 12 revs·min$^{-1}$; end 30 s cadence, 86 ± 12 vs. 87 ± 12 revs·min$^{-1}$; in all cases, $P > 0.05$). Mean second-by-second power output and cadence profiles of the two $3MT_R$ trials are shown in Fig. 1.

## *Reliability of the $3MT_R$*

The $EP_{remote}$ and TWD was acceptably reliable, as evidenced by the lack of systematic variance between the first and second trials, low CV, and strong correlation between-trials; the $WEP_{remote}$ derived from the $3MT_R$ was less reliable (Fig. 2, Table 1). These reliability statistics were largely consistent when considering the overall cohort ($N = 53$), participants who used "rear wheel off" indoor trainers ($N = 38$), and participants who used "rear wheel on" indoor trainers ($N = 15$) (Fig. 2, Table 2). The most common device used to measure power output during the remote 3MT ($N = 18$) was the Wahoo Kickr Core (Wahoo Fitness, Atlanta, USA), which has been validated (Hoon et al. 2016) Reliability statistics performed on this sub-group of participants produced the

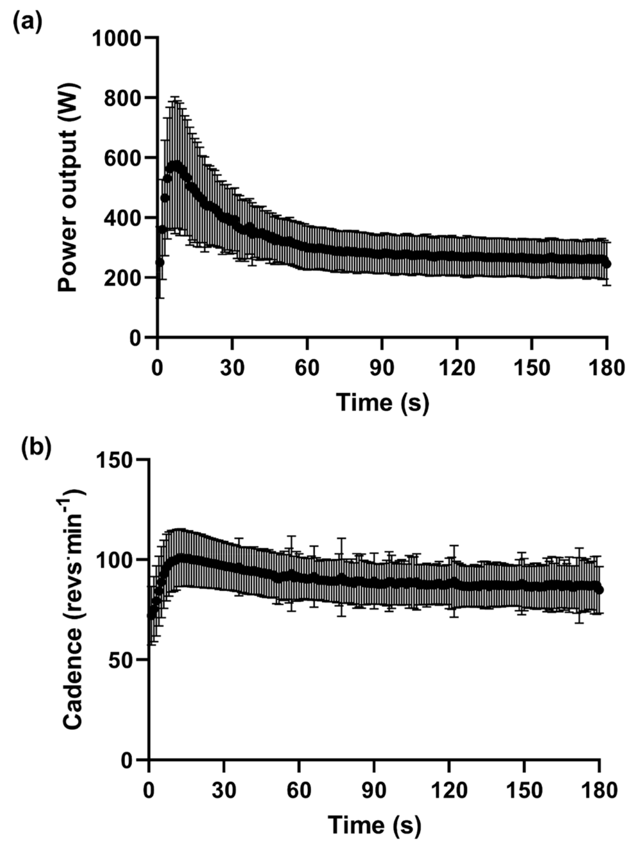


**Fig. 1** Remote three-minute all-out test **a** mean ± SD power (W) vs. time and **b** mean ± SD cadence (revs·min$^{-1}$) vs. time. The mean of each individual's power and cadence from the two trials was calculated for each second-by-second interval. Data presented are from the whole cohort ($N = 53$)

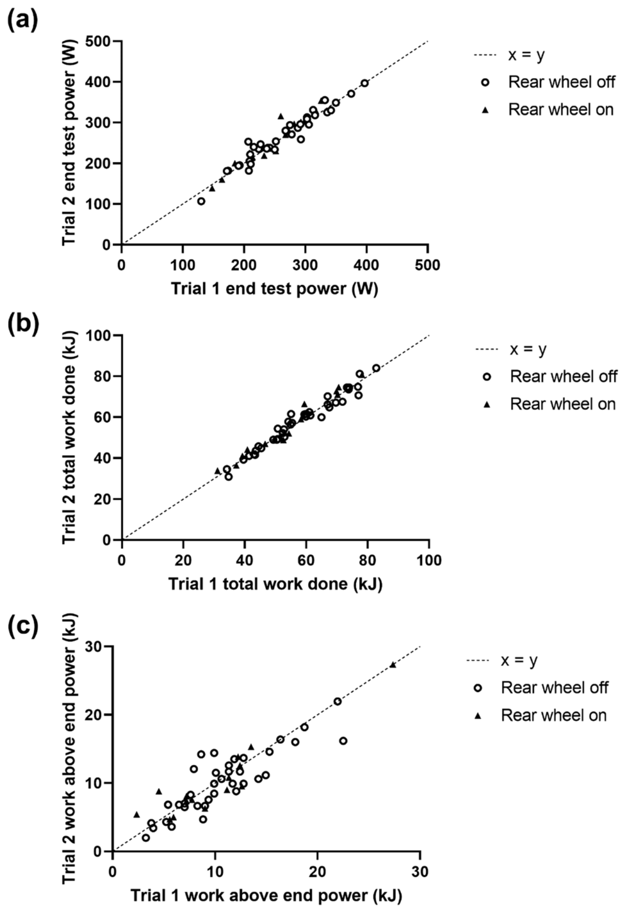


**Fig. 2** Scatter plot of metrics obtained from the two remote three-minute all-out tests (with dashed line for x = y); **a** $EP_{remote}$ (W), **b** total work done (kJ), and **c** W′ (kJ). Participants completing the trials on a rear wheel off indoor trainer are shown with solid squares. Participants completing the trials on a rear wheel on indoor trainer are shown with x symbols

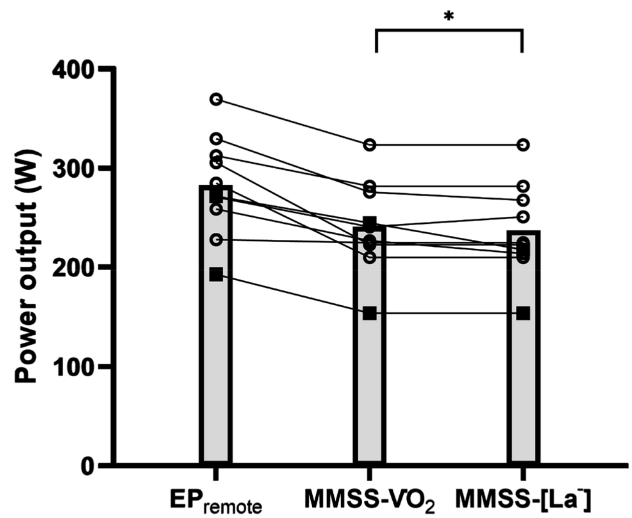


**Fig. 3** Mean (bars) and individual (lines with markers) values for the end power output during the remote three-minute all-out tests ($EP_{remote}$), power output at the maximum metabolic steady state as defined by $\dot{V}O_2$ kinetics during constant-work rate cycling (MMSS-$\dot{V}O_2$), and power output at the maximum metabolic steady state as defined by blood lactate responses to constant-work rate cycling (MMSS-[La$^-$]). Male participants are indicated with clear markers ($N = 8$), and female participants are indicated with solid markers ($N = 2$). '*' denotes $P < 0.05$ vs. $EP_{remote}$

same inferences as the overall cohort for $EP_{remote}$ (264 ± 64 vs. 264 ± 60 W [$P = 0.84$]; CV, 3.3% [2.4, 3.7%]; $r$, 0.98 [0.94, 0.99, $P < 0.0001$]), TWD (59 ± 14 vs. 58 ± 14 kJ [$P = 0.60$]; CV, 3.2% [2.3, 4.6%]; $r$, 0.98 [0.95, 0.99, $P < 0.0001$]), and W′ (11.0 ± 5.1 vs. 10.8 ± 4.5 kJ [$P = 0.73$]; CV, 17.9% [12.5, 26.8%]; $r$, 0.83 [0.60, 0.94, $P < 0.0001$]).

### *Validity of the $3MT_R$*

The sub-group of ten participants who completed the laboratory-based validity component of this study produced similar $3MT_R$ results as the overall cohort (peak power output, 693 ± 255 W; $EP_{remote}$, 283 ± 51 W; $WEP_{remote}$, 9.4 ± 2.6 kJ). None of the participants were able to complete 30 min of constant-work rate cycling at $EP_{remote}$ (time-to-task failure, 11.4 ± 6.5 min; range, 5.0–21.5 min). In all participants, physiological responses to constant-work rate cycling at $EP_{remote}$ were characteristic of the severe-intensity domain (blood [La$^-$] at task failure, 12.3 ± 3.5 mmol·L$^{-1}$, range, 6.0–18.3 mmol·L$^{-1}$; in all cases, the fundamental phase of the $\dot{V}O_2$ response was not completed prior to task failure).

The $EP_{remote}$ overestimated MMSS-$\dot{V}O_2$ (241 ± 46 W, $P = 0.0003$, ES = 1.62 [0.71, 2.67], percent difference, 18 ± 11%) and MMSS-[La$^-$] (237 ± 47 W, $P = 0.0003$, ES = 1.72 [0.78, 2.82], percent difference, 20 ± 12%). The MMSS-$\dot{V}O_2$ and MMSS-[La$^-$] were not significantly different ($P = 0.37$, ES = 0.34 [– 0.27, 0.97], Fig. 3). The $\dot{V}O_2$ and blood lactate responses to constant-work rate trials immediately above and below MMSS-$\dot{V}O_2$ and MMSS-[La$^-$] are shown in Fig. 4.

## Discussion

The primary aim of the present investigation was to determine the reliability and validity of the 3MT when performed in remote settings by unsupervised athletes using their own indoor cycling setup ($3MT_R$). The primary outcomes were that end-test power in the $3MT_R$ was reliable, but overestimated the MMSS. These results suggest that the $3MT_R$, as performed in the present investigation, should not be used to estimate the MMSS in endurance-trained cyclists.

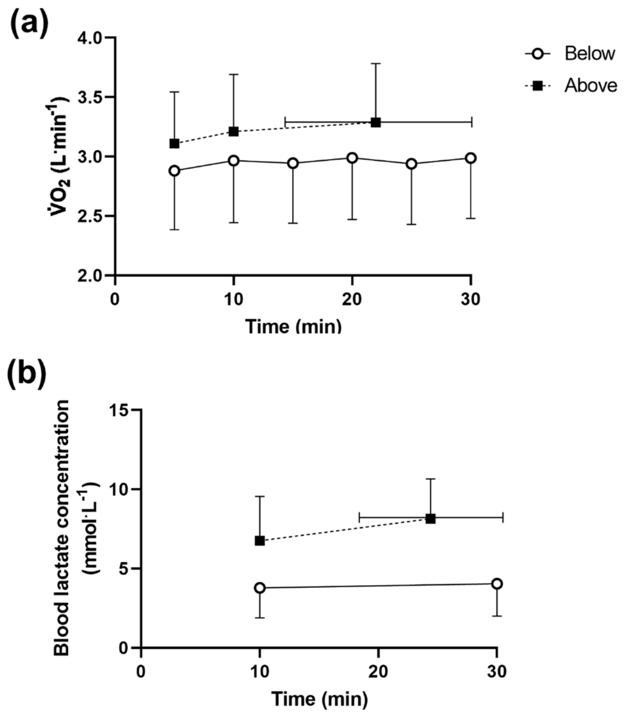


**Fig. 4** **a** $\dot{V}O_2$ and **b** blood lactate concentration responses to constant-work rate cycling performed immediately below (clear markers) and above (solid markers) **a** MMSS-$\dot{V}O_2$ and **b** MMSS-[La$^-$] ($N = 10$).

The reliability of the $3MT_R$ is evidenced by the low CV and high Pearson's correlation coefficients for the primary outcome metrics (Fig. 2, Table 1). The CV values reported for $EP_{remote}$ (~4.0–5.5%) are similar to what has been reported for laboratory-based measurements of end-test power (~3–7%) (Burnley et al. 2006; Johnson et al. 2011). The $WEP_{remote}$ was less reliable than $EP_{remote}$ (CV, ~13.9–16.2%), which is in line with reliability data for laboratory-based estimates (CV, ~20–28%) (Johnson et al. 2011). Therefore, these data suggest that the $3MT_R$ produces similarly repeatable outcomes as the laboratory-based 3MT. However, despite the strong reliability of the $3MT_R$, we present robust evidence that $EP_{remote}$ overestimated the MMSS; this is shown by the severe-intensity physiological responses and short time-to-task failure (11.4 ± 6.5 min; range, 5.0–21.5 min) during constant-work rate cycling at $EP_{remote}$, and the consistency of differences between $EP_{remote}$ and MMSS-$\dot{V}O_2$ and MMSS-[La$^-$] at an individual level (Fig. 3). Additionally, we feel the magnitude by which $EP_{remote}$ overestimated the power outputs at MMSS-$\dot{V}O_2$ (18 ± 11%) and MMSS-[La$^-$] (20 ± 12%) permits this conclusion, even in the context of recent data demonstrating the transition from the heavy to severe-intensity domain is a phased transition (Pethick et al. 2020). Therefore, it is unlikely that the overestimation of MMSS by $EP_{remote}$ is attributable to variability in the results of remote tests.

In the $3MT_R$, total work completed above the end-test power was substantially lower (10.3 ± 4.8 kJ) than has typically been observed during laboratory-based trials (~14–17 kJ) (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a, 2008b, 2016). Importantly, the sub-group of participants included in the validity aspect of this study had similarly low $WEP_{remote}$ (9.4 ± 2.6 kJ). In these participants' work above, the MMSS-$\dot{V}O_2$ power output was similar to work above laboratory-based end-test power observed elsewhere (17.4 ± 3.7 kJ) (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a, 2008b, 2016). Accordingly, it is likely participants in the present investigation failed to fully deplete work output above MMSS in the initial 150 s of the test, and therefore, that $EP_{remote}$ was supplemented by work output above MMSS. This would explain why subsequent constant-work rate trials at $EP_{remote}$ produced clear severe-intensity

**Table 1** Reliability statistics for the remote three-minute all-out test

| | Trial 2–Trial 1 | CV (95% CI) | *r* (95% CI) |
|---|---|---|---|
| Overall cohort ($N = 53$) | | | |
| $EP_{remote}$ (W) | 4 ± 16 ($P = 0.12$) | 4.5% (3.7, 5.5%) | 0.97 (0.94, 0.98, $P < 0.0001$) |
| TWD (kJ) | 0 ± 3 ($P = 0.38$) | 3.3% (2.7, 4.0%) | 0.98 (0.87, 0.99, $P < 0.0001$) |
| $WEP_{remote}$ (kJ) | − 0.3 ± 2.3 ($P = 0.33$) | 15.6% (12.5, 19.8%) | 0.90 (0.82, 0.94, $P < 0.0001$) |
| Rear wheel off ($N = 38$) | | | |
| $EP_{remote}$ (W) | 2 ± 15 ($P = 0.38$) | 4.0% (3.2, 5.1%) | 0.97 (0.94, 0.98, $P < 0.0001$) |
| TWD (kJ) | 0 ± 3 ($P = 0.84$) | 3.1% (2.5, 4.0%) | 0.98 (0.96, 0.99, $P < 0.0001$) |
| $WEP_{remote}$ (kJ) | − 0.5 ± 2.3 ($P = 0.22$) | 16.2% (12.5, 21.3%) | 0.87 (0.77, 0.93, $P < 0.0001$) |
| Rear wheel on ($N = 15$) | | | |
| $EP_{remote}$ (W) | 7 ± 19 ($P = 0.17$) | 5.5% (3.8, 8.1%) | 0.96 (0.89, 0.99, $P < 0.0001$) |
| TWD (kJ) | 1 ± 3 ($P = 0.07$) | 3.7% (2.6, 5.5%) | 0.98 (0.95, 0.99, $P < 0.0001$) |
| $WEP_{remote}$ (kJ) | 0.1 ± 2.0 ($P = 0.83$) | 13.9% (9.1, 23.1%) | 0.94 (0.82, 0.98, $P < 0.0001$) |

*CI* confidence intervals, *CV* within-subject coefficient of variation, $EP_{remote}$ average power during the last 30 s of the three-minute all-out test, *r* Pearson's correlation coefficient, *TWD* total work done, *W′* work above $EP_{remote}$

responses, and therefore, why $EP_{remote}$ significantly overestimated MMSS.

The reason why work above MMSS was not fully depleted in the initial 150 s of the $3MT_R$ could be attributable to several factors. First, laboratory-based trials are supervised and strong verbal encouragement is provided throughout the test (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a, 2008b). Therefore, in the $3MT_R$ where participants were unsupervised and strong verbal encouragement was not provided by the researchers, it is possible the absence of social facilitation resulted in a sub-maximal or paced effort, and therefore failure to fully deplete work above MMSS in the first 150 s. Given that the work above the MMSS-$\dot{V}O_2$ power output was similar to work above laboratory-based end-test power observed elsewhere (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a, 2008b, 2016), it is possible that the $3MT_R$ was paced but still a maximal effort overall (i.e., that total work done across the 180 s was maximal). Had the $3MT_R$ been performed as a sub-maximal overall effort, it is likely that work above MMSS-$\dot{V}O_2$ would have also been noticeably low. The possibility of pacing within the $3MT_R$ is supported by the lowest 6 and 30 s power outputs being substantially lower than $EP_{remote}$ ($91 \pm 5$ and $97 \pm 3\%$ of $EP_{remote}$, respectively). Periods of cycling below MMSS during the $3MT_R$ may have allowed partial recovery of the finite capacity for work above MMSS during the test (Skiba et al. 2012), and thus could have inflated the end-test power value and contributed to the overestimation of MMSS. Pacing may have been made more likely by participants being able to view elapsed time during the $3MT_R$, which is a key difference compared to laboratory 3MTs (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a, 2008b).

A further key difference was the opportunity for altering the resistance to pedalling in the remote test, compared to the fixed linear factor used in previous work (Burnley et al. 2006; Vanhatalo et al. 2007, 2008a). Peak cadence achieved in the $3MT_R$ in the present investigation ($110 \pm 21$ revs·min$^{-1}$) was substantially lower than those typically achieved in the laboratory test (~ 140–155 revs·min$^{-1}$), although end-test cadence was similar (Burnley et al. 2006; Vanhatalo et al. 2008b, a). Therefore, given the similarity in peak power output between our remote and previous studies of laboratory-based 3MTs, it is likely participants in the present investigation self-selected a greater pedalling resistance in the initial part of the test. Differences in the cadence profiles of the $3MT_R$ investigated here and the laboratory-based 3MT reported elsewhere may also contribute to why work above MMSS was not fully depleted in the initial 150 s of the remote test. It has previously been shown that adjusting the linear factor to produce a higher peak cadence ($155 \pm 12$ vs. $148 \pm 15$ revs·min$^{-1}$) resulted in a significant reduction in end-test power and total work done in a laboratory-based 3MT (Vanhatalo et al. 2008a), time-to-task failure at a constant-work rate in the severe domain was reduced when cadence was experimentally increased by 20 revs·min$^{-1}$ (Nielsen et al. 2004), and mean power output during a 30-s all-out sprint was reduced by ~ 15% when performed isokinetically (100 revs·min$^{-1}$) compared to isoinertially with a higher peak cadence ($117 \pm 14$ revs·min$^{-1}$) (Fuentes et al. 2013). As the rate of metabolic energy expenditure is increased at a given power output at higher pedalling cadences (Umberger et al. 2006; Brennan et al. 2019), it is possible the higher cadences achieved during laboratory trials may be necessary to fully deplete work above MMSS in the initial period of a 3MT, and therefore for end-test power to produce a valid estimate of the MMSS. This may explain why other studies have reported lower end-test power output during laboratory-based 3MTs performed at higher than preferred cadences (Wright et al. 2019), and that critical power is greater when cycling at 60 vs.100 revs·min$^{-1}$ (Barker et al. 2006).

Future research may seek to determine if alteration to the $3MT_R$ instructions used in the present investigation would facilitate full depletion of work above MMSS in the initial component of the test, and therefore provide a valid estimate of MMSS. Speculatively, video conferencing in which researchers or a coach view the test and provide verbal encouragement in real time may help to reduce pacing and would also obviate the need for athletes to be able to view elapsed time. Another strategy may be to partially deplete work above MMSS prior to a $3MT_R$ with a planned severe-intensity effort. This approach is similar to what has been investigated recently in a laboratory setting, whereby a ramp test is performed prior to a 3MT (Goulding et al. 2021).

In summary, the present investigation suggests that whilst the three-minute all-out test can be performed reliably by endurance-trained cyclists in remote settings using typical indoor training set-ups, these tests overestimated the MMSS, likely due to failure to fully deplete work above end power in the initial 150 s of the test. Therefore, the $3MT_R$ protocol adopted in the present investigation should not be used for identification of the MMSS. Given that many individual athletes do not have regular access to laboratory facilities (and the on-going possibility of restricted travel to, and use of, laboratories during global pandemics), future research should explore if alterations to the $3MT_R$ protocol utilised here can produce valid estimates of the MMSS.

**Acknowledgements** We would like to thank the participants for their hard work in completing this study.

**Author contributions** EM, JAR, AR, and AEK conceived and designed the research. EM, JAR, and AR conducted experiments and collected

the data. EM, JAR, AR, and MD analysed the data. EM drafted the manuscript. All authors read, revised, and approved the manuscript.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions.

## Declarations

**Conflict of interest** No sources of funding were used in the preparation of this manuscript. The authors have no conflicts of interest related to this work.



## References

Barker T, Poole DC, Noble ML, Barstow TJ (2006) Human critical power-oxygen uptake relationship at different pedalling frequencies. Exp Physiol 91:621–632. https://doi.org/10.1113/expphysiol.2005.032789

Black MI, Jones AM, Blackwell JR et al (2017) Muscle metabolic and neuromuscular determinants of fatigue during cycling in different exercise intensity domains. J Appl Physiol 122:446–459. https://doi.org/10.1152/japplphysiol.00942.2016

Brennan SF, Cresswell AG, Farris DJ, Lichtwark GA (2019) The effect of cadence on the mechanics and energetics of constant power cycling. Med Sci Sports Exerc 51:941–950. https://doi.org/10.1249/MSS.0000000000001863

Burnley M, Doust JH, Vanhatalo A (2006) A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. Med Sci Sports Exerc 38:1995–2003. https://doi.org/10.1249/01.mss.0000232024.06114.a6

Coyle EF, Coggan AR, Hopper MK, Walters TJ (1988) Determinants of endurance in well-trained cyclists. J Appl Physiol 64:2622–2630. https://doi.org/10.1152/jappl.1988.64.6.2622

Fuentes T, Ponce-González JG, Morales-Alamo D et al (2013) Isoinertial and isokinetic sprints: muscle signalling. Int J Sports Med 34:285–292. https://doi.org/10.1055/s-0032-1312583

Goulding RP, Roche DM, Marwood S (2021) The ramp and all-out exercise test to determine critical power: validity and robustness to manipulations in body position. Eur J Appl Physiol. https://doi.org/10.1007/s00421-021-04739-9

Hoon M, Michael S, Chapman P, Areta JL (2016) A comparison of the accuracy and reliability of the Wahoo KICKR and SRM power meter. J Sci Cycl 5:11–15

Johnson TM, Sexton PJ, Placek AM et al (2011) Reliability analysis of the 3-min all-out exercise test for cycle ergometry. Med Sci Sports Exerc 43:2375–2380. https://doi.org/10.1249/MSS.0b013e318224cb0f

Jones AM, Vanhatalo A (2017) The 'critical power' concept: applications to sports performance with a focus on intermittent high-intensity exercise. Sports Med 47:65–78. https://doi.org/10.1007/s40279-017-0688-0

Jones AM, Wilkerson DP, DiMenna F et al (2008) Muscle metabolic responses to exercise above and below the "critical power" assessed using 31P-MRS. Am J Physiol-Regul Integr Comp Physiol 294:R585–R593. https://doi.org/10.1152/ajpregu.00731.2007

Jones AM, Burnley M, Black MI et al (2019) The maximal metabolic steady state: redefining the gold standard. Physiol Rep 7:1–16. https://doi.org/10.14814/phy2.14098

Maunder E, Seiler S, Mildenhall MJ et al (2021) The importance of 'durability' in the physiological profiling of endurance athletes. Sports Med 51:1619–1628. https://doi.org/10.1007/s40279-021-01459-0

Nielsen JS, Hansen EA, Sjøgaard G (2004) Pedalling rate affects endurance performance during high-intensity cycling. Eur J Appl Physiol 92:114–120. https://doi.org/10.1007/s00421-004-1048-y

Pethick J, Winter SL, Burnley M (2020) Physiological evidence that the critical torque is a phase transition, not a threshold. Med Sci Sports Exerc 52:2390–2401. https://doi.org/10.1249/MSS.0000000000002389

Skiba PF, Chidnok W, Vanhatalo A, Jones AM (2012) Modeling the expenditure and reconstitution of work capacity above critical power. Med Sci Sports Exerc 44:1526–1532. https://doi.org/10.1249/MSS.0b013e3182517a80

Umberger BR, Gerritsen KGM, Martin PE (2006) Muscle fiber type effects on energetically optimal cadences in cycling. J Biomech 39:1472–1479. https://doi.org/10.1016/j.jbiomech.2005.03.025

Vanhatalo A, Doust JH, Burnley M (2007) Determination of critical power using a 3-min all-out cycling test. Med Sci Sports Exerc 39:548–555. https://doi.org/10.1249/mss.0b013e31802dd3e6

Vanhatalo A, Doust JH, Burnley M (2008a) Robustness of a 3 min all-out cycling test to manipulations of power profile and cadence in humans. Exp Physiol 93:383–390. https://doi.org/10.1113/expphysiol.2007.039883

Vanhatalo A, Doust JH, Burnley M (2008b) A 3-min all-out cycling test is sensitive to a change in critical power. Med Sci Sports Exerc 40:1693–1699. https://doi.org/10.1249/MSS.0b013e318177871a

Vanhatalo A, Black MI, DiMenna FJ et al (2016) The mechanistic bases of the power–time relationship: muscle metabolic responses and relationships to muscle fibre type. J Physiol 594:4407–4423. https://doi.org/10.1113/JP271879

Wright J, Bruce-Low S, Jobson SA (2019) The 3-minute all-out cycling test is sensitive to changes in cadence using the lode Excalibur sport ergometer. J Sports Sci 37:156–162. https://doi.org/10.1080/02640414.2018.1487115

Zadow EK, Kitic CM, Wu SSX et al (2016) Validity of power settings of the wahoo KICKR power trainer. Int J Sports Physiol Perform 11:1115–1117. https://doi.org/10.1123/ijspp.2015-0733

Zadow EK, Kitic CM, Wu SSX, Fell JW (2018) Reliability of power settings of the wahoo KICKR power trainer after 60 hours of use. Int J Sports Physiol Perform 13:119–121. https://doi.org/10.1123/ijspp.2016-0732