International Journal of
Sports Science
& Coaching

Original research

International Journal of Sports Science
& Coaching
2025, Vol. 20(6) 2613–2623

DOI: 10.1177/17479541251345575
journals.sagepub.com/home/spo
Sage

Reliability and usefulness of a self-regulated 6-km submaximal ergometer training test set in elite rowers

James R Clark[1], Andrew J McKune[2,3], and Paola S Wood[1]

Abstract

Incorporating tests into regular training sessions permits sport-specific and convenient monitoring of athletes without disrupting team preparation or adding to testing burden. However, few studies have investigated the measurement properties of training test sets for monitoring athletes. This study assessed the reliability and usefulness of a self-regulated 6-km rowing ergometer training-based test (6SRT) as a submaximal monitoring tool in elite rowers. In a repeated measures design, 43 (13 female) rowers at a national training centre were asked to perform the 6SRT according to a 'firm-to-hard'/'threshold' session category goal at 22 strokes·min^{-1} in three consecutive weeks. Mean ± standard deviation 6SRT time, power output (PO), exercise heart rate (HR_{ex}) and rating of perceived exertion (RPE) were 22:18.0 ± 1:21.6 min:s, 70 ± 4% 2-km time trial PO, 87 ± 3% maximal HR and 16 ± 1, respectively. There were no significant between-trial effects ($p \geq 0.101$, $\eta_p^2 \leq 0.05$) and differences between paired trials were trivial (Cohen's d < 0.2) for all variables. Intraclass correlation coefficients ($ICC_{3,1}$: 0.62–0.98) were moderate to excellent for all responses except RPE (0.33). Coefficient of variation (CV: 0.6–3.5%) and typical error (TE: 0.7–4.1%) were low for all variables except HR recovery (HRR). Time, PO and HR_{ex} responses, and PO:HR_{ex} and PO:RPE intensity ratios showed TE ≤ smallest worthwhile change (SWC). The 6SRT shows acceptable reliability and usefulness under stable field conditions to be a valuable training test set for monitoring high-level rowers when maximal tests are undesirable or laboratory tests are inaccessible.

Keywords

Athlete monitoring, heart rate recovery, power output, rating of perceived exertion

Introduction

Monitoring how athletes are responding to training is considered imperative for guiding preparation toward optimal health and performance outcomes,[1] particularly in sports with heavy training demands, like rowing.[2] Two essential measurement properties of any athlete monitoring tool are reliability and usefulness. Reliability refers to the extent to which scores on repeated trials of a test are consistent and free of measurement error,[3,4] and informs on the precision of single measures and the ability to confidently detect changes over time.[5] Usefulness refers to the sensitivity with which a test is able to detect an important, or worthwhile, change, given the measurement error in scores.[6]

Reliable and useful indicators of rowing performance and fitness (e.g., 2-km ergometer time trial [TT], maximal oxygen uptake [$\dot{V}O_{2max}$], incremental peak power output [PPO], lactate threshold [LT] power output [PO], 30-s PO) are well-established.[7] However, the exhaustive efforts or invasive, time-consuming, and costly procedures of these tests make them unsuitable for frequently or rapidly evaluating how rowers are responding to training throughout preparation and competition periods.[8] As an alternative, submaximal rowing ergometer tests[8,9] have demonstrated acceptable reliability for tracking training-related changes

Reviewers: Abdul Rashid Aziz (Singapore Sport Institute, Singapore)
David Cardoso (University of Porto, Portugal)
Lorcan Daly (Technological University of the Shannon, Ireland)

[1]Department of Physiology, Faculty of Health Sciences, University of Pretoria, Pretoria, South Africa
[2]University of Canberra Research Institute for Sport and Exercise (UCRISE), University of Canberra, Canberra, Australia
[3]Discipline of Biokinetics, Exercise and Leisure Sciences, School of Health Science, University of KwaZulu-Natal, Durban, South Africa

Corresponding author:
James Clark, Department of Physiology, Faculty of Health Sciences, University of Pretoria, South Africa.
Email: jimmy.clark@up.ac.za

in rowers. However, the small samples of male participants in those reliability analyses limits broader generalisation and application of the data reported.[5] Moreover, applying those particular protocols in real-world settings to monitor large teams of rowers undergoing heavy training presents several challenges. For example, respiratory gas analysis or blood sampling, and deriving indicators like LT or PO associated with a HR of 170 beats·min^{-1} (PWC_{170})[9] precludes the regular and convenient use of a submaximal test. Also, athletes in a state of overreaching from heavy training commonly display an attenuated exercise HR (HR_{ex}) response,[10] potentially making attainment of a high submaximal HR (e.g., 90% maximal HR [HR_{max}]) required by a test[8] neither possible[11] nor desirable in athletes for whom the tests arguably hold most monitoring value.

'Invisible' monitoring describes embedding assessments within scheduled training,[4,12] implying little to no disruption to preparation from the addition of tests, improved time efficiency by simultaneously assessing multiple athletes, and regular, predictable reassessment when sessions are repeated across microcycles. In this regard, submaximal running drills or small-sided games are popular for tracking short-term fitness and fatigue responses in team sports,[13] and similar training-based monitoring approaches have been explored in cyclists[14,15] and runners.[16–18] In rowing, ergometer training sessions represent convenient, frequent and relatively standardised monitoring opportunities that warrant investigation as training test sets.[19] Notably, submaximal ergometer training sessions have been recommended for monitoring training responses,[20] potentially reflecting a perceived ecological validity among rowing coaches and sport scientists in the field. However, reliability and usefulness data necessary for interpreting repeated ergometer training test set responses have not yet been reported.

Using field experience and working alongside the head coach of a national rowing team, a non-exhaustive ergometer training session was selected to represent both a test that could be feasibly implemented under field conditions, and a training stimulus that could repeatedly (e.g., weekly to monthly) be included in the training programme across preparation periods. The test set comprises 6 km of self-paced ergometer rowing at a fixed stroke rate (SR) of 22 strokes·min^{-1} (6SRT). Although the 6SRT was designed to double as a heavy aerobic[21] training stimulus, rowers regulate intensity individually. Furthermore, administering the 6SRT and measuring rower responses does not require sophisticated equipment or involve complicated procedures. A fundamental prerequisite before adopting a test to systematically monitor athletes is establishing its reliability and usefulness for identifying worthwhile changes.[12,22] Therefore, this study aimed to determine the reliability and usefulness of the 6SRT to evaluate its suitability as a monitoring tool in high-level rowers. This information will aid practitioners and researchers seeking an alternative to existing protocols for evaluating rowers without disrupting training continuity.

Methods

Research design

A repeated measures observational design was used to determine the between-week reliability and usefulness of measures obtained from the 6SRT in a convenience sample of rowers. The participants' training schedule limited replicate observations to single trials each separated by seven days over three consecutive weeks. Procedures were performed independently and unblinded by a single investigator experienced in collecting exercise responses in rowers. The study was approved by the research ethics committee of the University of Pretoria, conducted in accordance with the Declaration of Helsinki, and reported using guidelines for reliability and agreement studies.[23]

Participants

Forty-seven rowers at a national training centre volunteered for this study. Permission was obtained from the head coach and athletes to use training and test data for research and publication purposes, and written informed consent was obtained. Adult male and female lightweight and open category rowers training at the centre for at least one full season prior to the study were eligible. Rowers with illness or injury determined by the team doctor, and those who did not participate in all training and trials during the study period, were excluded. All participants were familiar with the setting, procedures and equipment based on routine exposure to the same athlete monitoring activities. Four participants were excluded from the study due to missed or modified training sessions during the study period (n = 3) and malfunction of the ergometer computer in two of the three trials (n = 1). Therefore, 43 rowers (age: 22.9 ± 3.8 y, stature: 184.2 ± 7.6 cm, body mass: 78.5 ± 10.4 kg) were included in the analysis (Table 1). All were experienced (8.7 ± 3.3 y) rowers, and 36 (84%) had represented their country at international competitions. Seventeen participants (40%) competed in the lightweight rowing category, but none were targeting or maintaining official body mass limits at the time of the study.

General procedures

This study was conducted during the general preparation period (October to December), in which most training was prescribed at a session category goal described as 'low'[21] or 'steady'[24] intensity in zone 1 (Z1) of a three-zone classification model.[25] Weekly training during this study

Table 1. Participant characteristics.

Variable	Total (N = 43)	Male (n = 30)	Female (n = 13)
Age (y)	22.9 ± 3.8	22.8 ± 3.5	23.2 ± 4.6
Stature (cm)	184.2 ± 7.6	187.5 ± 6.0	176.4 ± 4.5
Body mass (kg)	78.5 ± 10.4	83.1 ± 7.8	67.9 ± 7.7
Rowing experience (y)	8.7 ± 3.3	8.7 ± 3.1	8.6 ± 3.8
Open category	26 (60%)	18 (60%)	8 (62%)
Lightweight category	17 (40%)	12 (40%)	5 (38%)
International level	36 (84%)	28 (93%)	8 (62%)
National level	7 (16%)	2 (7%)	5 (38%)
PO_{2km} (W)	370 ± 71	410 ± 38	276 ± 19
HR_{max} (beats·min^{-1})	197 ± 7	197 ± 5	198 ± 10

Data are presented as mean ± standard deviation for continuous variables and count (% category total) for categorical variables. PO_{2km}: 2-km TT mean power output; HR_{max}: maximal heart rate.

comprised thirteen sessions: six specific endurance (on-water and ergometer rowing); four other endurance (running or cycling); and three strength training sessions. The same weekly training routine was replicated throughout this study period, and no unaccustomed training or progression was prescribed. All trials took place without prior exercise between 08:30 and 10:30 on the first day of the week, and no training was prescribed on the day prior to all three trials. Repeated trials were conducted in the same indoor training venue with temperature (21.3 ± 1.9°C) and relative humidity (RH, 54 ± 7%) recorded. Participants were asked to maintain habitual sleep, dietary and recovery practices for the duration of the study, and to follow the same morning routine in terms of food and fluid intake, including caffeine consumption, on all test days. Body mass (Tanita BF-350 electronic scale, Tanita Corporation, Tokyo, Japan) was measured immediately prior to all trials. The recovery-stress questionnaire for athletes (RESTQ-Sport)[26] was completed before each trial to verify that overall recovery status was stable across trials.

6-km submaximal rowing ergometer test set

Consolidation of the 6SRT protocol in the weeks before the study served as familiarisation so that all participants had performed the test on at least one occasion beforehand. All tests were performed on factory-calibrated stationary Concept2 Model D (Concept2, Morrisville, VT, USA) air-braked ergometers—the same ergometers participants used for regular indoor training sessions and periodic performance tests. Ergometer drag factor was controlled by having each participant select their preferred setting within the range of 110–120×10^{-6} N·m·s^2 for males and 95–105×10^{-6} N·m·s^2 for females based on habitual use in training, and then ensuring the same setting was used for all subsequent trials; ergometer foot stretcher position was controlled in the same manner. A familiar and standardised pre-training warm-up for the 6SRT consisted of 3 km of self-paced ergometer rowing at 20 strokes·min^{-1}, followed by 5 min of rowing alternately at 18 and 22 strokes·min^{-1} for 30 s each.

Participants completed the 6SRT as a continuous heavy aerobic endurance training bout according to the familiar session category goal[24] of 'firm-to-hard'[27]/'threshold'[21] within the setting. Rowers were asked to provide an even-paced effort throughout the 6-km set and maintain the same SR (22 strokes·min^{-1}) and rhythm. This fixed submaximal SR[28] is within the range (18–26 strokes·min^{-1}) imposed[9] or self-selected[8] in reliable submaximal rowing ergometer protocols using fixed-intensity formats. To facilitate this, and consistent with routine practice, participants rowed in full view of the Concept2 PM4 (Concept2, Morrisville, VT, USA) rowing ergometer computer, permitting real-time feedback on elapsed time and distance, SR and rowing pace; however, previous results were withheld during repeated trials. Within 30 s of test completion participants were asked to provide a RPE[29] representing their overall degree of exertion. Following 90 s of seated recovery participants initiated a self-selected post-training cool down.

Time and mean PO were automatically stored on the rowing ergometer computer and retrieved after each trial. Heart rate (Polar Team[2], Polar Electro Oy, Kempele, Finland) was recorded continuously during the 6SRT, averaged over 1-s intervals and analysed using Polar Team[2] version 1.4.5 software (Polar Electro Oy, Kempele, Finland). Mean HR_{ex} was determined for the entire 6SRT as well as adjusted to exclude the first 60 s of data (HR_{ex-adj}) to account for the uncontrolled starting HR across trials and the lag time in the rest-to-exercise HR response.[30] Post-exercise HR recovery (HRR) was determined as the absolute difference between HR at test termination and at 60 s of recovery, and expressed as a percentage of HR at test termination to account for potential differences in end-test HR between trials.[31] Since the 6SRT does not use a fixed intensity, simple bivariate ratios among gross external (PO) and internal (HR_{ex} and RPE) intensity measures were determined.[14,15] For descriptive purposes, mean PO and HR_{ex} were expressed relative to 2-km TT mean PO (PO_{2km}) and HR_{max}, respectively. Individual PO_{2km} and HR_{max} data were obtained from routine maximal TT and incremental tests performed earlier in the season.

Statistical analysis

Analyses were conducted using IBM SPSS Statistics version 28.0 (IBM Corp., Armonk, NY, USA) and a custom MS Excel (Microsoft Corp., Redmond, WA, USA) spreadsheet.[32] Variables are descriptively summarised as mean ± standard deviation (SD). Statistical significance was considered as $p < 0.05$. The Shapiro-Wilk test and visual inspection of histograms and normal Q-Q plots were used to determine whether data were normally distributed.

Systematic bias between trials was assessed using repeated measures analysis of variance (ANOVA) or Friedman's test for normally and not normally distributed variables, respectively. Correspondingly, if a significant test statistic was found, a Tukey or Bonferroni post hoc analysis was used to locate pairwise differences. Sphericity was assessed using Mauchly's test and if the sphericity condition was violated, a Greenhouse-Geisser adjustment was used.[33] Partial eta squared (η_p^2) values of 0.01, 0.06, and 0.14 were considered minimum thresholds for small, moderate and large effects across trials.[33] Cohen's d was calculated to express the magnitude of between-trial differences as standardised mean differences (effect size, ES) and interpreted as trivial, small, moderate, large, very large or extremely large for values <0.2, <0.6, <1.2, <2.0, <4.0 or ≥4.0.[34]

Relative reliability was assessed using a two-way mixed effects intraclass correlation coefficient ($ICC_{3,1}$)[35] accompanied by 95% confidence intervals (CI). Thresholds were used to interpret ICC magnitude for single measurements as excellent (≥0.90), good (≥0.75), moderate (≥0.50) or poor (<0.50).[36] Variability across trials was assessed using the CV of participant scores[37] and interpreted as low, moderate or high if CV was <5%, 5–10% or >10% respectively.[38] Absolute reliability was assessed using typical error (TE) determined as the SD of between-trial change scores divided by $\sqrt{2}$.[5] The threshold representing a 'real' effect was considered the minimal detectable change (MDC) at the 95% level of confidence ($MDC_{95\%}$), calculated as TE x $\sqrt{2} \times 1.96$.[35] The smallest worthwhile change (SWC) in each outcome measure was estimated by multiplying the pooled between-participant SD by a standardised ES of 0.2, with thresholds of three-, six- and ten-times the SWC used to delineate the lower boundaries of moderate, large and very large changes, respectively.[34] To facilitate comparison, TE, SWC and $MDC_{95\%}$ were expressed in units of measurement and percentage of mean scores. Practical usefulness for detecting small but important changes was assessed by comparing imprecision with meaningful change using the criteria: TE:SWC <1 ('good'); TE:SWC ~1 ('acceptable'); and TE:SWC >1 ('marginal').[6]

Geographic and financial constraints limited access to other cohorts of rowers competing at national or international level, and therefore no *a priori* sample size calculation was performed. However, a *post hoc* analysis indicated N = 43 and ICC ≈ 0.85 over three trials represented >85% power for detecting ICCs ≥ 0.70 in a reliability study.[39]

Results

There were no significant trial effects for temperature (21.5 ± 1.8°C v 21.1 ± 1.9°C v 21.4 ± 1.9°C, $p = 0.267$, $\eta_p^2 = 0.03$) or RH (53 ± 7% v 54 ± 7% v 55 ± 7%, $p = 0.782$, $\eta_p^2 = 0.01$), with trivial to small between-trial differences ($p \geq 0.427$, $d \leq 0.2$). Body mass showed a mean between-trial change of 0.8 ± 0.4 kg, non-significant ($p = 0.309$, $\eta_p^2 = 0.03$) trial effects and trivial ($p \geq 0.370$, $d \leq 0.1$) between-trial differences. Total stress (15.4 ± 6.0 v 15.0 ± 5.9 v 15.2 ± 6.2 arbitrary units [AU], $p = 0.827$, $\eta_p^2 = 0.01$) and recovery (32.2 ± 7.4 v 33.0 ± 7.1 v 32.4 ± 7.6 AU, $p = 0.469$, $\eta_p^2 = 0.02$) scores were not significantly different across trials, and between-trial differences were trivial ($p \geq 0.560$, $d \leq 0.1$).

An example of the 6SRT heart rate response for a single participant is shown in Figure 1, and group responses across all trials are described in Table 2. Individual mean 6SRT time ranged from 20:30.8 min:s to 25:02.8 min:s, and all trials were completed at 22 strokes·min^{-1}. Overall the 6SRT was completed at 70.1 ± 3.9%PO_{2km} and 86.5 ± 2.8%HR_{max}, and yielded an RPE of 16.0 ± 0.8 and HRR of 37 ± 8 beats·min^{-1}. Individual PO:HR_{ex}, PO:RPE, HR_{ex}:RPE and HRR responses across all three trials are illustrated in Figure 2. None of the 6SRT variables showed significant trial effects ($p \geq 0.101$, $\eta_p^2 \leq 0.05$), with trivial ($d < 0.2$) differences between all trials.

Reliability statistics for 6SRT variables are shown in Table 3. Point estimates of ICCs ranged from moderate to excellent (0.62–0.98) for all variables except RPE (0.33). Relative reliability was highest (0.92 ≤ ICC ≤ 0.99) for 6SRT time, PO, PO:HR_{ex} and PO:RPE. Individual variability was low (CV < 5%) for all variables except HRR (CV: ~7–11%). The lowest TE (0.7–2.2%) was for 6SRT time, PO, HR_{ex}, HR_{ex-adj} and PO:HR_{ex}, with corresponding $MDC_{95\%}$ values (1.8–6.2%) representing small to moderate effects, and TE was highest for HRR (~10%). Time, PO and PO:HR_{ex} displayed usefulness for detecting small worthwhile changes, while other variables demonstrated usefulness for detecting moderate (HR_{ex}, HR_{exadj} and PO:RPE) and large changes (HRR and HR_{ex}:RPE).

Discussion

This study investigated whether responses during a self-regulated rowing ergometer training set based on a 'threshold' session goal displayed adequate reliability and usefulness for use as a submaximal test (6SRT) to monitor rowers under field conditions. The main finding is that time, PO and HR measures, along with gross PO:HR_{ex} and PO:RPE intensity ratios in the 6SRT showed high reliability, low measurement error and 'acceptable' to 'good' usefulness, whereas HRR measures displayed high variability and measurement error, and 'marginal' usefulness. There was also no significant systematic bias indicative of a learning or fatigue effect in repeated 6SRT trials. Results indicate that the 6SRT provides reliable markers and represents a useful ergometer training-

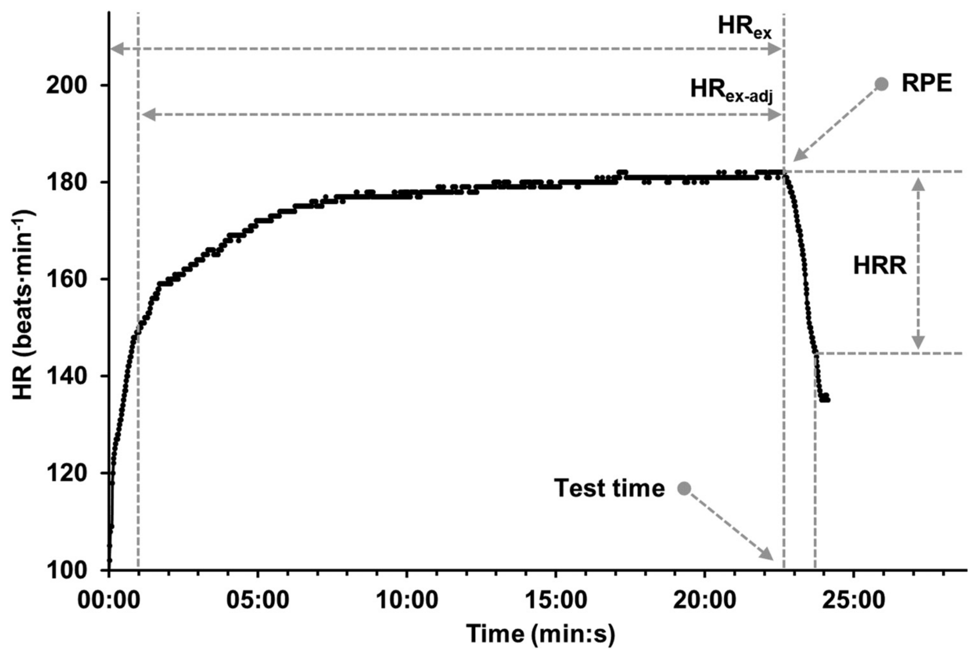

Figure 1. Example of 6-km submaximal rowing ergometer training test set (6SRT) heart rate (HR) response and other outcome variables for one participant.
HR_{ex}: mean exercise heart rate; HR_{ex-adj}: mean exercise heart rate excluding the first 60 s of data; RPE: rating of perceived exertion; HRR: heart rate recovery.

based method for monitoring experienced rowers under stable field conditions.

Submaximal cycling,[40] running[41–43] and rowing[8] protocols with anchored internal intensities (e.g., $\%HR_{max}$, $\%\dot{V}O_{2max}$, RPE) have shown good to excellent reliability (ICC: 0.76–1.00; TE: ~2–6%) for mechanical variables (e.g., average speed or PO). In the current study, 6SRT time and PO showed excellent reliability (ICC: 0.97–0.99, TE: ~1% and ~2% respectively), suggesting these variables are suitable for confidently identifying real and worthwhile changes as small as ~25 s or 14 W in rowers. These results indicate that elite rowers permitted real-time visualisation of ergometer mechanical variables during the 6SRT produce highly reliable training test set time and PO based on training session category goal, without a fixed internal intensity anchor.

Heart rate variables (HR_{ex} and HR_{ex-adj}) in the 6SRT showed good to excellent relative (ICC: 0.87–0.95) and absolute (TE: 2 beats·min^{-1}) reliability, comparable to (ICC: 0.94–0.99; TE: 1–4 beats·min^{-1})[40,43–45] or better than (ICC: 0.50–0.92; TE: 4–9 beats·min^{-1})[9,42] reported in fixed-intensity submaximal protocols in various modalities. The low variability of 6SRT HR measures (CV: ~1%) is comparable to[46] or better than (CV: ~3–5%)[9,47] that reported in controlled submaximal intensity running and rowing protocols. These results support HR as one of the most reliable submaximal response measures,[48] with acceptable usefulness for tracking small but worthwhile changes of ~1%.[49] Additionally, the reliability of HR measures appeared unaffected by using the convenient gross mean exercise HR (HR_{ex}) or the more cumbersome exclusion of the initial 60 s of data (HR_{ex-adj}). This is likely because the rest-to-exercise transition phase accounts for a smaller proportion of 6SRT time than in protocols with short-duration stages.

In this study, the low TE (0.6 units) for RPE compares favourably with results (0.7–1.1 units) in submaximal protocols using intensities anchored using $\%\dot{V}O_{2max}$[43] or $\%HR_{max}$,[8,40] while relative reliability was poor to moderate, and generally lower than previously reported.[8,40] In light of good reproducibility across trials (CV: ~3%), the low ICC observed may be a statistical artefact of the relatively small range (14–17) in RPE ratings among participants relative to within-individual variability.[36,37] Since 6SRT RPE shows usefulness for tracking large (~2 units) individual changes, it should be used in conjunction with other measures when interpreting training responses, and not as the basis for differentiating athletes.[23,35]

Post-6SRT HRR variables showed moderate to good reliability (ICC: 0.67–0.88) but also moderate to high variability and measurement error (~10%), allowing confident identification of only large (~30%) changes. Using the same determination method, studies in well-trained cyclists[40] and rowers[8] have shown good to excellent reliability (ICC: 0.93–0.99; TE: 4–8%) for HRR following submaximal tests. The main reason for the discrepancy in findings is likely the end-test HR between the protocols: fixed and predictable using $\%HR_{max}$ in the studies by

Table 2. 6SRT responses (mean ± SD) for all trials, and magnitude of between-trial differences (Cohen's *d*, descriptor).

Variable	Trial 1	Trial 2	Trial 3	Mean	Trial 1 to 2	Trial 2 to 3	Trial 1 to 3
Time (min:s)	22:21.0 ± 1:26.4	22:16.8 ± 1:20.4	22:16.8 ± 1:18.6	22:18.0 ± 1:21.6	0.05, trivial	0.01, trivial	0.05, trivial
PO (W)	257 ± 46	258 ± 44	258 ± 43	258 ± 44	−0.03, trivial	0.01, trivial	−0.03, trivial
PO (%PO_{2km})	69.7 ± 3.6	70.3 ± 4.0	70.3 ± 4.1	70.1 ± 3.9	−0.16, trivial	0.01, trivial	−0.15, trivial
SR (strokes·min^{-1})	22 ± 0	22 ± 0	22 ± 0	22 ± 0	0.00, trivial	0.00, trivial	0.00, trivial
HR_{ex} (beats·min^{-1})	171 ± 9	171 ± 8	171 ± 9	171 ± 9	−0.01, trivial	0.01, trivial	0.02, trivial
HR_{ex} (%HR_{max})	86.4 ± 2.9	86.5 ± 2.5	86.5 ± 2.9	86.5 ± 2.8	−0.01, trivial	0.01, trivial	0.03, trivial
HR_{ex-adj} (beats·min^{-1})	172 ± 9	173 ± 8	173 ± 9	173 ± 9	−0.01, trivial	0.02, trivial	0.03, trivial
HR_{ex-adj} (%HR_{max})	87.3 ± 2.8	87.4 ± 2.5	87.5 ± 2.9	87.4 ± 2.7	−0.03, trivial	0.03, trivial	0.06, trivial
RPE	16.1 ± 0.8	16.0 ± 0.7	16.0 ± 0.8	16.0 ± 0.8	0.11, trivial	−0.03, trivial	−0.15, trivial
HRR (beats·min^{-1})	36 ± 8	37 ± 8	37 ± 8	37 ± 8	−0.07, trivial	0.04, trivial	0.12, trivial
HRR (%)	20.0 ± 4.8	20.3 ± 4.6	20.5 ± 4.6	20.3 ± 4.7	−0.07, trivial	0.03, trivial	0.10, trivial
PO:HR_{ex}	1.51 ± 0.29	1.52 ± 0.28	1.52 ± 0.28	1.52 ± 0.28	−0.03, trivial	−0.01, trivial	0.02, trivial
PO:RPE	16.02 ± 2.99	16.25 ± 3.09	16.26 ± 3.06	16.18 ± 3.05	−0.07, trivial	0.01, trivial	0.08, trivial
HR_{ex}:RPE	10.64 ± 0.63	10.70 ± 0.55	10.73 ± 0.71	10.69 ± 0.63	−0.10, trivial	0.05, trivial	0.14, trivial

6SRT: 6-km submaximal rowing ergometer training test set; PO: power output, PO_{2km}: 2-km rowing ergometer time trial mean power output; SR: stroke rate; HR_{ex}: mean exercise heart rate; HR_{ex-adj}: mean exercise heart rate excluding the first 60 s of data; RPE: rating of perceived exertion; HRR: heart rate recovery.

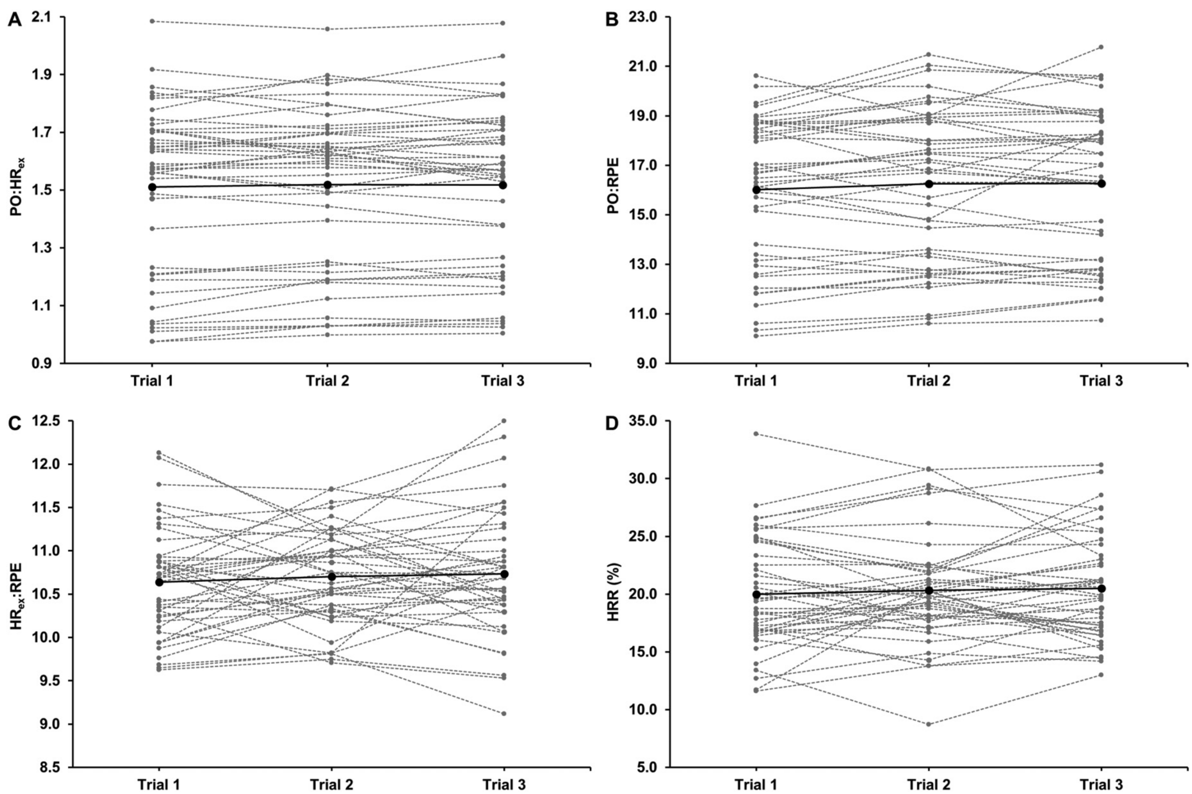

Figure 2. Individual (grey) and mean (black) responses in 6SRT PO:HR_{ex} (A), PO:RPE (B), HR_{ex}:RPE (C) and HRR (D) across three trials.
6SRT: 6-km submaximal rowing ergometer training test set; PO: power output; HR_{ex}: exercise heart rate; RPE: rating of perceived exertion; HRR: heart rate recovery.

Lamberts et al.[40] and Otter et al.,[8] but uncontrolled and variable between trials in the individually regulated 6SRT. Recent reliability data[50] for HRR after a self-paced field test in runners (ICC: 0.64, CV: ~11%) are very similar to our results, supporting this explanation. Individual changes in autonomic restoration between trials may also have contributed to the high variability in HRR.[51] Based on its 'marginal' usefulness for monitoring changes in elite rowers, post-6SRT HRR therefore does not appear to warrant the additional time required for its measurement.

Submaximal intensity ratios between gross external mechanical and internal response measures have been suggested for monitoring athlete fitness and fatigue,[1,52] and have been used in submaximal training-based monitoring of cyclists[14,15] and runners[16,18] where intensity is variable. To date, however, no reliability data for these ratios have been published. This study found excellent reliability (ICC $\geq$ 0.92, TE $\leq$ 4.1%) and low variability (CV $\leq$ 3.5%) for PO:HR_{ex} and PO:RPE, with PO:HR_{ex} showing 'good' and PO:RPE 'acceptable' usefulness for detecting small worthwhile changes. Conversely, reliability and usefulness for HR_{ex}:RPE were 'poor' to 'good' and 'marginal', respectively. These results suggest that PO:HR_{ex} and possibly PO:RPE are practical and useful markers for monitoring individual athletes, and warrant greater investigation. However, since ratios may violate underlying statistical assumptions,[53] simple submaximal bivariate intensity ratios should be avoided for evaluating between-athlete differences or group effects.[54]

The high reliability and usefulness shown for several 6SRT variables may be explained by the level of athletes studied and characteristics of the test. Based on a recent classification framework,[55] participants were Tier 3 (highly trained, national), 4 (elite, international) and 5 (world class) rowers. Experienced, highly-trained endurance athletes are capable of excellent pace judgement and regulation, enabling them to reproduce self-paced efforts in sport-specific tasks with low variability.[56] Also, rowing training sessions are frequently completed as fixed distances or durations with self-regulated intensities guided by a session category goal and SR prescription.[57] Rowers control intensity in large part by varying SR from ~18–20 strokes·min^{-1} in low-intensity training to ~35–45 strokes·min^{-1} in competitive races or TTs.[57,58] Although ergometer rowing permits modulation of effort for a given SR, potentially influencing relationships between PO, HR and RPE,[59] this does not appear to have compromised the reliability of 6SRT responses. Rather, the prescribed SR (22 strokes·min^{-1}) and visualisation of mechanical output during the

Table 3. Reliability and usefulness statistics for 6SRT variables.

Variable	ICC (95% CI)	CV (95% CI)	TE (%)	SWC (%)	$MDC_{95\%}$ (%)	Usefulness
Time (min:s)	0.98 (0.97–0.99)	0.6 (0.4–0.7)	0.15 (0.7)	0.27 (1.2)	0.41 (1.8)	Good
PO (W)	0.98 (0.97–0.99)	1.8 (1.4–2.2)	5 (2.0)	9 (3.4)	14 (5.5)	Good
HR_{ex} (beats·min^{-1})	0.92 (0.88–0.95)	1.2 (1.0–1.4)	2 (1.4)	2 (1.0)	7 (3.8)	Acceptable
HR_{ex-adj} (beats·min^{-1})	0.92 (0.87–0.95)	1.2 (1.0–1.4)	2 (1.4)	2 (1.0)	7 (3.9)	Acceptable
RPE	0.33 (0.14–0.52)	3.2 (2.5–3.9)	0.6 (3.8)	0.2 (1.0)	1.7 (10.6)	Marginal
HRR (beats·min^{-1})	0.78 (0.67–0.86)	9.2 (7.7–10.8)	4 (9.9)	1 (3.7)	10 (27.5)	Marginal
HRR (%)	0.80 (0.70–0.88)	9.5 (7.9–11.1)	2.1 (10.3)	0.8 (4.1)	5.8 (28.5)	Marginal
PO:HR_{ex}	0.98 (0.97–0.99)	2.0 (1.6–2.3)	0.03 (2.2)	0.06 (3.7)	0.09 (6.2)	Good
PO:RPE	0.95 (0.92–0.97)	3.5 (2.9–4.1)	0.66 (4.1)	0.59 (3.7)	1.84 (11.4)	Acceptable
HR_{ex}:RPE	0.62 (0.46–0.75)	3.2 (2.6–3.7)	0.38 (3.5)	0.10 (1.0)	1.04 (9.8)	Marginal

6SRT: 6-km submaximal rowing ergometer training test set; ICC: intraclass correlation coefficient; CI: confidence interval; CV: coefficient of variation; TE: typical error; SWC: smallest worthwhile change; $MDC_{95\%}$: minimal detectable change at 95% level of confidence; PO: power output; HR_{ex}: mean exercise heart rate; HR_{ex-adj}: mean exercise heart rate excluding the first 60 s of data; RPE: rating of perceived exertion; HRR: heart rate recovery.

6SRT may help minimise within-athlete fluctuations in mechanical, physiological and perceptual responses.[60,61] In experienced rowers, the 6SRT therefore represents a familiar, sport-specific task with limited attentional focus, arguably representing a more ecologically valid format than protocols requiring careful adjustment of PO[9] or HR.[8]

An important observation is that the 6SRT was completed without incident and represented a submaximal (non-exhaustive), 'firm-to-hard'[27]/'threshold'[21] training session based on the range for individual mean PO (63–79% PO_{2km}), HR_{ex} (82–92%HR_{max}) and RPE (14–17). These responses align with those described in submaximal cycling,[40] running,[41–43,45] and rowing[8,9] protocols shown to yield reliable responses. It is also worth noting that studies using submaximal tests involving multiple intensities[8,40,42,50,62] have consistently shown better reliability at higher (i.e., 80–90%HR_{max}, RPE: 13–17) than at lower (i.e., 60–70%HR_{max}, RPE: 8–12) submaximal intensities. Those protocols were developed for regular (e.g., daily to weekly) monitoring, with relatively short-duration (≤8 min) exposure to the higher, more reliable submaximal intensities limiting potential effects on training. Conversely, as a ∼20–25-min 'threshold' training test set,[19] the 6SRT is not suitable for daily use. Rather, the 6SRT shows potential as a convenient assessment as part of the training programme on a weekly to monthly basis, which is a popular frequency of submaximal test use in athlete monitoring.[63,64]

Limitations

Limitations in this study merit acknowledgement. First, convenience sampling of participants from a national rowing centre may limit the generalisability of results in rowers at lower competitive or conditioning levels (e.g., Tier 1 and 2[55] rowers). However, we believe the sample is representative of typical athlete groups that practitioners who may utilise the 6SRT will be working with. Second, determining a heavy-to-severe intensity transition marker (e.g., anaerobic threshold, maximal lactate steady state, critical power)[21,65] was not possible during this study, and should be included in future work to more clearly define 6SRT intensity and assist decision-making on test adoption, repetition and scheduling. Finally, 6SRT responses in female participants may have been influenced by menstrual cycle phase,[66] which was not recorded in this study. While individual between-trial responses in our study do not suggest higher variability in female than male participants, results should be considered as the real-world between-week reliability that may be expected in a field setting with a mixed-sex squad of rowers.

Practical applications and future research

Reliability and usefulness results suggest that changes in gross (i.e., unadjusted) 6SRT metrics like time (≥2%), PO (≥6%), HR_{ex} (≥4%), PO:HR_{ex} (≥6%) and PO:RPE (≥11%) in the 6SRT can confidently be interpreted as real and worthwhile in highly-trained rowers. The PO:HR_{ex} and PO:RPE ratios may be particularly valuable given that the 6SRT does not impose a fixed intensity. Furthermore, the 6SRT can conveniently be administered to a large group of rowers simultaneously as part of the training programme. Therefore, the 6SRT shows potential as a reliable tool for coaches and sport scientists to monitor the training response in rowers when additional or exhaustive tests are undesirable, or sophisticated test facilities or resources are limited. Future studies should investigate the predictive validity of the 6SRT for performance, fitness or recovery status, and its sensitivity to longitudinal (e.g., seasonal) changes. An additional avenue of research may involve blinding rowers to mechanical variables for insight into the effect of visual feedback on the reliability of self-regulated training test sets like the 6SRT.

Conclusion

The 6SRT is a convenient, reliable and useful self-regulated submaximal training test set with the potential to be used for monitoring the training response in experienced rowers under stable field conditions. Results may be valuable for confidently identifying changes and guiding training and recovery decisions without disrupting the training schedule. It is hoped that this study may encourage further investigation and refinement of standardised training-based assessments as 'invisible' monitoring methods in rowers and other endurance-trained athletes.

Acknowledgements

Mr R Barrow, Dr D Lincoln and Mr J Young are thanked for supporting this study.

ORCID iDs

James R Clark https://orcid.org/0000-0002-3453-1753
Andrew J McKune https://orcid.org/0000-0002-5479-1544
Paola S Wood https://orcid.org/0000-0002-1242-4842

Ethical considerations

This study was approved by the Research Ethics Committee of the Faculty of Health Sciences, University of Pretoria (633/2019) on 24 October 2019.

Consent to participate

Participants provided written informed consent to be part of this study.

Consent for publication

Participants provided written informed consent for their anonymised data to be published in this study.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability

Data sets generated and analysed in this study are not publicly available, but reasonable requests for access via direct email to the corresponding author will be considered.

References

1. Bourdon PC, Cardinale M, Murray A, et al. Monitoring athlete training loads: consensus statement. *Int J Sports Physiol Perform* 2017; 12: S161–S170.
2. Mäestu J, Jürimäe J and Jürimäe T. Monitoring of performance and training in rowing. *Sports Med* 2005; 35: 597–617.
3. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
4. Weakley J, Black G, McLaren S, et al. Testing and profiling athletes: recommendations for test selection, implementation, and maximizing information. *Strength Cond J* 2024; 46: 159–179.
5. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000; 30: 1–15.
6. Hopkins WG. How to interpret changes in an athletic performance test. *Sportscience* 2004; 8: 1–7.
7. Smith TB and Hopkins WG. Measures of rowing performance. *Sports Med* 2012; 42: 343–358.
8. Otter RT, Brink MS, Lamberts RP, et al. A new submaximal rowing test to predict 2,000-m rowing ergometer performance. *J Strength Cond Res* 2015; 29: 2426–2433.
9. Klusiewicz A, Faff J and Sitkowski D. Reproducibility of the results of the laboratory exercise performed on rowing ergometer. *Biol Sport* 1998; 1998: 145–150.
10. Roete AJ, Elferink-Gemser MT, Otter RTA, et al. A systematic review on markers of functional overreaching in endurance athletes. *Int J Sports Physiol Perform* 2021; 16: 1065–1073.
11. Decroix L, Lamberts RP and Meeusen R. Can the Lamberts and Lambert submaximal cycle test reflect overreaching in professional cyclists? *Int J Sports Physiol Perform* 2018; 13: 23–28.
12. West SW, Clubb J, Torres-Ronda L, et al. More than a metric: how training load is used in elite sport for athlete management. *Int J Sports Med* 2021; 42: 300–306.
13. Shushan T, Norris D, McLaren SJ, et al. A worldwide survey on the practices and perceptions of submaximal fitness tests in team sports. *Int J Sports Physiol Perform* 2023; 18: 765–779.
14. Sanders D, Heijboer M, Hesselink MKC, et al. Analysing a cycling grand tour: can we monitor fatigue with intensity or load ratios? *J Sports Sci* 2018; 36: 1385–1391.
15. Wahl Y, Achtzehn S, Schäfer Olstad D, et al. Training load measures and biomarker responses during a 7-day training camp in young cyclists—a pilot study. *Medicina (B Aires)* 2021; 57: 73.
16. Nuuttila OP, Nummela A, Korhonen E, et al. Individualized endurance training based on recovery and training status in recreational runners. *Med Sci Sports Exerc* 2022; 54: 1690–1701.
17. Nuuttila OP, Uusitalo A, Kokkonen VP, et al. Monitoring fatigue state with heart rate-based and subjective methods during intensified training in recreational runners. *Eur J Sport Sci* 2024; 24: 857–869.
18. Vesterinen V, Hokka L, Hynynen E, et al. Heart rate-running speed index may be an efficient method of monitoring endurance training adaptation. *J Strength Cond Res* 2014; 28: 902–908.
19. Sharma AP and Mujika I. Quantification of training and competition loads in endurance sports: a key to recovery-stress balance and performance. In: M Kellmann and J Beckmann (eds) *Recovery and well-being in sport and exercise*. Oxon: Routledge, 2022, pp.111–128.
20. Rice AJ. *Information for coaches and scientists. 2013-2016 Australian rowing team ergometer protocols. Version 3.0.* Australia: Australian Institute of Sport, 2015.
21. Treff G, Winkert K, Sareban M, et al. The polarization-index: a simple calculation to distinguish polarized from non-

polarized training intensity distributions. *Front Physiol* 2019; 10: 1–6.
22. Riemann BL and Lininger MR. Statistical primer for athletic trainers: the essentials of understanding measures of reliability and minimal important change. *J Athl Train* 2018; 53: 98–103.
23. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol* 2011; 64: 96–106.
24. Seiler KS and Kjerland GØ. Quantifying training intensity distribution in elite endurance athletes: is there evidence for an "optimal" distribution? *Scand J Med Sci Sports* 2006; 16: 49–56.
25. Seiler S. What is best practice for training intensity and duration distribution in endurance athletes? *Int J Sports Physiol Perform* 2010; 5: 276–291.
26. Kellmann M and Kallus KW. *Recovery-stress questionnaire for athletes: User manual*. Champaign, IL: Human Kinetics, 2001.
27. Eston RG and Parfitt G. Perceived exertion, heart rate, and other non-invasive methods for exercise testing and intensity control. In: KI Norton and RG Eston (eds) *Kinanthropometry and exercise physiology*. 4th ed. New York, NY: Routledge, 2019, pp.464–499.
28. Smith B. Monitoring training. In: V Nolte (eds) *Rowing science*. Champaign, IL: Human Kinetics, 2024, pp.279–299.
29. Borg G. *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics, 1998.
30. Fitzpatrick JF, Hicks KM, Russell M, et al. The reliability of potential fatigue-monitoring measures in elite youth soccer players. *J Strength Cond Res* 2021; 35: 3448–3452.
31. Veugelers KR, Naughton GA, Duncan CS, et al. Validity and reliability of a submaximal intermittent running test in elite Australian football players. *J Strength Cond Res* 2016; 30: 3347–3353.
32. Hopkins WG. Spreadsheets for analysis of validity and reliability. *Sportscience* 2015; 19: 36–44.
33. Vincent WJ and Weir JP. *Statistics in kinesiology*. 4th ed. Champaign, IL: Human Kinetics, 2012.
34. Hopkins WG, Marshall SW, Batterham AM, et al. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc* 2009; 41: 3–13.
35. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005; 19: 231–240.
36. Koo TK and Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016; 15: 155–163.
37. Hopkins WG, Hawley JA and Burke LM. Design and analysis of research on sport performance enhancement. *Med Sci Sports Exerc* 1999; 31: 472–485.
38. Turner A, Brazier J, Bishop C, et al. Data analysis for strength and conditioning coaches: using excel to analyze reliability, differences, and relationships. *Strength Cond J* 2015; 37: 76–83.
39. Borg DN, Bach AJE, O'Brien JL, et al. Calculating sample size for reliability studies. *PM&R* 2022; 14: 1018–1025.
40. Lamberts RP, Swart J, Noakes TD, et al. A novel submaximal cycle test to monitor fatigue and predict cycling performance. *Br J Sports Med* 2011; 45: 797–804.
41. Wang SC, Lin HF, Wu CF, et al. Aerobic power assessment by using a 10 min heart rate control running on treadmill. *J Sports Med Phys Fitness* 2010; 50: 32–36.
42. Sangan HF, Hopker JG, Davison G, et al. The self-paced submaximal run test: associations with the graded exercise test and reliability. *Int J Sports Physiol Perform* 2021; 16: 1865–1873.
43. Mann TN, Lamberts RP and Lambert MI. Day-to-day variation in heart rate recovery and excess post-exercise oxygen consumption after a submaximal treadmill protocol. *Int SportMed J* 2014; 15: 352–364.
44. Lamberts RP, Swart J, Woolrich RW, et al. Measurement error associated with performance testing in well-trained cyclists: application to the precision of monitoring changes in training status. *Int SportMed J* 2009; 10: 33–44.
45. Bosquet L, Gamelin FX and Berthoin S. Reliability of postexercise heart rate recovery. *Int J Sports Med* 2008; 29: 238–243.
46. Lamberts RP and Lambert MI. Day-to-day variation in heart rate at different levels of submaximal exertion: implications for monitoring training. *J Strength Cond Res* 2009; 23: 1005–1010.
47. Bagger M, Petersen PH and Pedersen PK. Biological variation in variables associated with exercise training. *Int J Sports Med* 2003; 24: 433–440.
48. Shushan T, McLaren SJ, Buchheit M, et al. Submaximal fitness tests in team sports: a theoretical framework for evaluating physiological state. *Sports Med* 2022; 52: 2605–2626.
49. Buchheit M. Monitoring training status with HR measures: do all roads lead to Rome? *Front Physiol* 2014; 5: 1–19.
50. Nuuttila OP, Uusitalo A, Kokkonen VP, et al. Self-paced field running test in monitoring fatigue and training adaptations in recreational runners. *Int J Sports Physiol Perform* 2024; 19: 1299–1306.
51. Fecchio RY, Brito L, Leicht AS, et al. Reproducibility of post-exercise heart rate recovery indices: a systematic review. *Auton Neurosci* 2019; 221: 1–9.
52. Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med* 2014; 44: S139–S147.
53. Atkinson G and Batterham AM. The use of ratios and percentage changes in sports medicine: time for a rethink? *Int J Sports Med* 2012; 33: 505–506.
54. McLaren SJ, Macpherson TW, Coutts AJ, et al. The relationships between internal and external measures of training load and intensity in team sports: a meta-analysis. *Sports Med* 2018; 48: 641–658.
55. McKay AKA, Stellingwerff T, Smith ES, et al. Defining training and performance caliber: a participant classification framework. *Int J Sports Physiol Perform* 2022; 17: 317–331.
56. Stevens CJ and Dascombe BJ. The reliability and validity of protocols for the assessment of endurance sports performance: an updated review. *Meas Phys Educ Exerc Sci* 2015; 19: 177–185.
57. Thompson P. *Sculling: Training, technique and performance*. Marlborough, Wiltshire, UK: The Crowood Press Ltd, 2005.
58. Ettema G, Haug A, Ludvigsen TP, et al. The role of stroke rate and intensity on rowing technique. *Sports Biomech* 2022; 00: 1–22.

59. Watts SP, Binnie MJ, Goods PSR, et al. Exploring the depths of on-water training in highly-trained rowing athletes. *Eur J Sport Sci* 2024; 24: 597–605.
60. Bell G, Attwood K, Syrotuik D, et al. Comparison of the physiological adaptations to high vs. low stroke rate training in rowers. *Res Sports Med* 1998; 8: 113–122.
61. Hofmijster MJ, Landman EH, Smith RM, et al. Effect of stroke rate on the distribution of net mechanical power in rowing. *J Sports Sci* 2007; 25: 403–411.
62. O'Grady C, Passfield L and Hopker JG. Variability in submaximal self-paced exercise bouts of different intensity and duration. *Int J Sports Physiol Perform* 2021; 16: 1824–1833.
63. McGuigan H, Hassmén P, Rosic N, et al. Training monitoring methods used in the field by coaches and practitioners: a systematic review. *Int J Sports Sci Coach* 2020; 15: 439–451.
64. Taylor K, Chapman DW, Cronin JB, et al. Fatigue monitoring in high performance sport: a survey of current trends. *J Aust Strength Cond* 2012; 20: 12–23.
65. Jamnick NA, Pettitt RW, Granata C, et al. An examination and critique of current methods to determine exercise intensity. *Sports Med* 2020; 50: 1729–1756.
66. McHugh C, Ackerman KE, Foley G, et al. The experiences and perceptions of the menstrual cycle and its impact on performance in female rowers. *Br J Sports Med* 2024; 58: A114.