International Journal of EXERCISE SCIENCE

Original Research

Test-Retest Reliability of a 4-Minute All-Out Critical Force Test in Rock Climbers

ZACHARY J. McCLEAN[†#1], KEENAN B. MacDOUGALL[†#1], JARED R. FLETCHER[‡2], SAIED J. ABOODARDA[‡1], and BRIAN R. MACINTOSH[‡1]

[1]Human Performance Laboratory, Faculty of Kinesiology, University of Calgary, Calgary, AB, CANADA; [2]Department of Health and Physical Education, Mount Royal University, Calgary, AB, CANADA

[†]Denotes graduate student author, [‡]Denotes professional author, [#]Denotes co-first authorship

ABSTRACT

International Journal of Exercise Science 16(4): 912-923, 2023. The purpose of this study was to assess the test-retest reliability of a 4-minute all-out critical force test in well-trained rock climbers. Thirteen rock climbers (n=4 females) completed a familiarization session and two all-out critical force tests on different days. During each trial, participants completed 24 repetitions of 7s right-handed, maximal effort hangs from a 20mm edge interspersed with 3 s rest. The end-test force (EF; i.e., critical force), impulse above EF (IEF), and peak force achieved durin g the test were analyzed with paired t-tests to determine differences between trials. Intraclass correlation coefficient (ICC), coefficient of variation (CV), and Bland-Altman analysis were performed to quantify the relative and absolute reliability of the measure, respectively. The level of significance for this study was set at $p<0.05$. There were no significant differences between trials for any of the reported variables ($P\geq0.455$). For EF, IEF, and peak force, ICC was 0.848, 0.820, and 0.938, respectively; and CV was 21.0%, 13.2%, and 5.6%, respectively. Bland-Altman analyses showed a mean relative bias of -2.3%, -2.8%, and -1.3%, with 95% limits of agreement (LoA) of -62.6% to 58.1%, -40.5% to 30.9%, and -17.2% to 14.6% for EF, IEF, and peak force, respectively, however linear regression revealed a significant proportional bias for EF ($p = 0.026$, $R^2 = 0.377$). The reliability of this protocol was good to excellent for all parameters; however, there was larger intra-individual variability for EF and IEF. This study suggests that when using the 4-min all-out critical force test in rock climbers, coaches and athletes should be aware that there may be a trade-off between the test's practicality and the precision of its results.

KEY WORDS: Finger flexor performance; critical force assessment; reliability

INTRODUCTION

Rock climbing is a sport that heavily relies on the strength and endurance of the finger flexors, as the ability to maintain sufficient finger flexor force during fatiguing climbs plays a large role in climbing success (18). As research interest in rock climbing continues to grow, finding valid and reliable methods to assess the physiological characteristics underpinning rock climbing performance is critical for both the prescription of exercise, as well as monitoring the effectiveness of training. The rules of competitive sport climbing require climbers to climb as high as they can on a wall within a 6 min time limit, where points are awarded based on the highest point reached, and the amount of time to do so. Thus, the climber who can sustain a higher intensity (i.e., perform more difficult moves) for a longer duration than their competitors, should hold the advantage.

The relationship between exercise intensity and duration is often modeled as a hyperbolic function which contains two parameters (25). The asymptote represents a critical intensity (force, power, speed) and the curvature constant (often termed "W′") dictates a fixed amount of exercise (e.g., impulse, work) that can be performed above the critical intensity. It is generally assumed that the critical intensity can be supported by aerobic energy, and that the fixed amount of exercise above this intensity is supported by non-aerobic energy. Theoretically, once this energy is depleted, continued exercise can only be supported at or below the critical intensity.

Conventionally, the critical intensity and W′ are estimated by having participants perform multiple (e.g., 3-5), constant-intensity exercise bouts until task failure in the range of approximately 2-15 minutes (25). By plotting the exercise intensity as a function of the time to task failure, or a mathematically equivalent transform, these two parameters can be estimated. The critical intensity concept has been successfully utilized in various exercises (e.g., cycling (26), running (12), swimming (28)), and due to the increase in rock climbing popularity and associated development of high-performance sport climbing, researchers have adopted the critical intensity approach in the quantification of finger flexor performance (16,17). As the critical intensity concept may offer some predictive ability of endurance performance in the time range of typical sport climbs, it provides an intriguing lens to quantify climbing capability.

While the conventional method of estimating the critical intensity is useful for both normalizing and prescribing exercise intensities, it is also time consuming, requiring multiple laboratory visits across different days (19). More recently, investigators have utilized 'all-out' tests as a more time efficient means of estimating this threshold (5,6,9,17,21,27). The theoretical rationale for the all-out tests is that, once W′ is fully depleted, the critical intensity should be the highest force (torque or power) that can be sustained, despite continued maximal effort. Thus, the critical intensity should correspond to the output at the end of an all-out trial when the performance has reached a plateau, and indeed such tests have been shown to agree very well with conventional methods (5,27). Recently, Giles *et al.* (17) developed a 4-min all-out critical force test for rock climbers, using a 7:3 s work-to-rest ratio and an instrumented 20mm hangboard rung (17). Although these authors reported good relative reliability with a sub-sample of participants (n = 7) (intraclass correlation coefficients (ICC) of 0.96 and 0.87 for critical force and W′, respectively), the 95% limits of agreement (LoA) for W′ was substantial, and these authors did not report the coefficient of variation (CV) of either measure. To the best of our knowledge, no study has explored this reliability with a more complete statistical analysis, and verified the findings of Giles *et al.* (17) with a larger sample. Moreover, the instrumented hangboard used by Giles *et al.* (17) is limited in its utility, as it only allows testing on a single type of hold, and its price may be prohibitive for many coaches and athletes. We recently validated an inexpensive alternative measurement device marketed towards rock climbers, (Entralpi force plate, Entralpi, Montreal, Canada), which provides the advantage of implementing various types of climbing holds during testing.

While an attractive methodology due to its improved time efficiency, the reliability of an all-out rock climbing specific critical force test must first be assessed prior to being implemented in either research or practical settings. Therefore, using a novel and inexpensive experimental setup, the purpose of this study was to verify the pilot testing of Giles *et al.* (17), and to establish the test-retest reliability of a 4-min all-out critical force test in rock climbers.

METHODS

Participants

A power analysis using the R package ICC.Sample.Size (R Version 3.5.2) determined that a minimum of 10 participants were needed for a power of 0.80 and an α = 0.05, based on the reported ICC for critical force of 0.96 from Giles *et al.*,(17) and a minimum threshold of 0.75. For this study, thirteen recreationally active rock climbers (9 males, 4 females, mean ± standard deviation: age = 30.3 ± 8.3 years, weight = 69.4 ± 7.8 kg, height = 175.0 ± 12.4 cm) were recruited using convenience sampling by sending study information to local climbing training centres, with inclusion criteria stating that participants must have been engaged in rock climbing training regularly for at least the last 2 years. To compare the climbing ability of participants in our study to that of previous studies, participants self-report their highest red-point lead climbing grade achieved in the previous six months (the highest difficulty route climbed after the route has been previously rehearsed) (23), which has been validated for lead climbing ability (10). Grades were then converted to a universal scale recommended by the International Rock Climbing Research Association (IRCRA)(11). Participant climbing ability ranged from 14 to 25 on the IRCRA scale (mean = 18.4 ± 3.8). All participants cleared the Physical Activity Readiness Questionnaire (PAR-Q+) and were free of any health conditions or injuries that could prevent them from performing maximal contractions on a hangboard. Participants were informed of the benefits and risks of the study and provided written informed consent, which was approved by the University of Calgary Conjoint Health Research Ethics Board (REB21-1503). The study was conducted in accordance with the Declaration of Helsinki (without registration). The investigation meets the guidelines set forth by the International Journal of Exercise Science (24).

Protocol

Participants visited the laboratory on three separate occasions, separated by at least 48 hours (mean of 8 days), and all sessions for a given participant occurred at approximately the same time of day (i.e., within 2 hours of the time that the familiarization trial occurred). Each of these visits were identical, consisting of a standardized warmup followed by the 4-min all-out critical force test (see below). The first session served as a familiarization trial, while the second and third sessions were assessed for their test-retest reliability. All participants were instructed to avoid any alcohol and caffeine for 12 hours prior to the testing, and to refrain from strenuous exercise in the 24 hours preceding the testing. Additionally, all participants were instructed to

adhere to their typical dietary habits 48 hours prior to each experimental session, and arrive at each session sufficiently hydrated to engaged in high intensity exercise.
At the beginning of each session, participants completed a standardized finger flexor warmup utilizing a half-crimp position with their dominant hand maintaining an approximately 90-degree angle at the proximal interphalangeal joint of the index finger (warm up protocol: 5 s on, 5 s off at 2×25%, 2×50%, 1×75%, 1×100% of peak force measured during the familiarization trial) performed on a 20 mm wooden edge (Lattice Training, Chesterfield, UK). During the trial, participants were positioned with the shoulder of their dominant hand flexed to ~ 180 degrees overhead, with their elbow extended to ~ 180 degrees, and instructed to hang as much of their bodyweight as they could while maintaining a half-crimp hand position. Of note, no participants could hang their entire bodyweight from the rung; therefore, the addition of external weight was not required. Participants completed 24 repetitions of 7 s maximal hangs with 3 s recovery between repetitions. Strong verbal encouragement was provided throughout. Participants were instructed that they could lower their arm between contractions to apply chalk, but could not shake their hands, as this has been shown to improve intermittent handgrip performance (2).

Data acquisition: Participants stood on a calibrated force plate, which recorded net force (relative to body weight) throughout the trial (i.e., force decreased as bodyweight was lifted off the force plate). Instantaneous net force signals, as well as audio and visual prompts for signaling the beginning and end of each contraction were displayed on a screen in front of participants via the Entralpi online software application. Data for the Entralpi force plate were collected in the Entralpi application. The Entralpi force plate was validated against the Pasco Pasport 2-Axis force platform (model PS-2142, Roseville, California, USA) and showed excellent agreement in all force measurements during the all-out tests (average difference of 0.02% and 0.5% for peak force and total impulse, respectively).The day-to-day reliability of the force plate was confirmed by a static force measurement over two days using standard weight plates across a range of ~16 to ~91 kg (ICC= 1, CV = 0.03%).

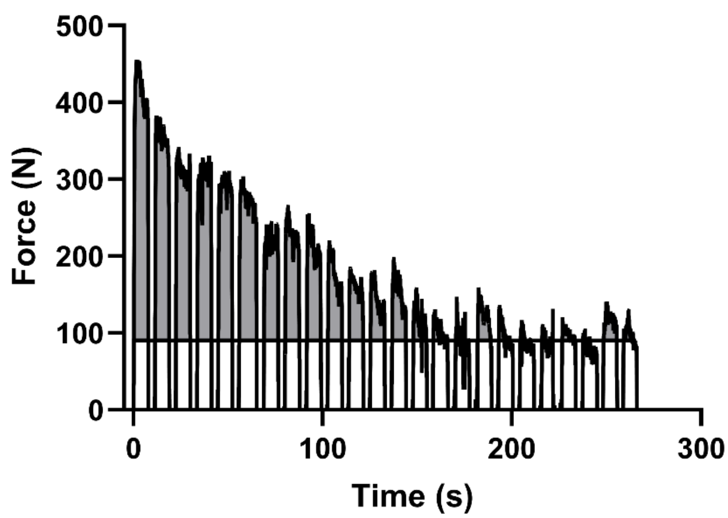

Figure 1. Representative force profile for a 24-repetition all-out critical force test on a 20 mm hangboard. Participants (n=13) conducted an all-out critical force test. End-test force was calculated as the average force across

the final six repetitions, shown as the horizontal thick black line. Impulse above end-test force was calculated as the total force- time integral above the end-test force, represented as the sum of the areas shaded in grey.

Statistical Analysis

Data were exported to Microsoft Excel (Microsoft Office 2017, Microsoft, Redmond, USA), and the end-test force (ET), used as an estimate of critical force, was calculated as the average force during the final six contractions of the all-out trial. The impulse above EF (IEF), used as an estimate of W′, was calculated as the total area under the force-time curve above the estimated critical force (see Figure 1). Peak force was calculated as the highest instantaneous force reached during the all-out trial. As the Entralpi system outputs force in kg, these values were converted to force (N): Force (N) = Entralpi force (in kg) × *g* (where *g* is the acceleration due to gravity, 9.81 $m \cdot s^{-2}$) before statistical analysis. Descriptive analysis is presented as mean ± standard deviation. Statistical analyses were conducted using GraphPad Prism v 9.0.0 (GraphPad Software, San Diego, California, USA) and SPSS version 25 (IBM Corp, Armonk, New York, USA). Normality was confirmed by Shapiro-Wilk tests. The presence of systematic changes between tests was assessed using a paired t-test. Relative reliability was assessed using ICC (2-way mixed effects, single rater, absolute agreement) (22). ICC values were interpreted according to Koo and Li (22), whereby ICC less than 0.5 is poor, between 0.5 and 0.75 is moderate, between 0.75 and 0.9 is good, and greater than 0.9 is excellent. Absolute reliability was assessed using within subject CV using the root mean square method (20), as well as Bland-Altman analysis with mean bias and 95% LoA. To check for proportional bias, linear regression of the differences between tests as a function of the mean value was performed to determine if the slope was significantly different from zero, as described by Bland and Altman (3). Statistical significance was set to an alpha level of $p < 0.05$.

RESULTS

Group mean and individual values between test 1 and test 2 for EF, IEF, and peak force are shown in Figure 2A, 2B, and 2C, respectively. There were no significant differences between test 1 and test 2 for any of the reported measures (P = 0.455, 0.753, and 0.727 for EF, IEF, and peak force, respectively). The ICC for EF, IEF, and peak force between test 1 and test 2 were 0.848, 0.820, and 0.938, respectively. The within-subject CV for these same values were 21.0%, 13.2%, and 5.6%. Bland-Altman analyses showed a mean relative bias of -2.3%, -2.8%, and -1.3%, with 95% limits of agreement of -62.6% to 58.1%, -40.5% to 30.9%, and -17.2% to 14.6% for EF, IEF, and peak force, respectively (Figure 3). Linear regression of the differences between the two tests against the mean, revealed no proportional bias for peak force or IEF, as evident from the estimate of the slope being not significantly different than zero ($p = 0.248$ and $p = 0.487$, respectively). However, there was a proportional bias for EF ($p = 0.026$, $R^2 = 0.377$), whereby the difference between tests became more positive as the EF value became larger (see Figure 3A).

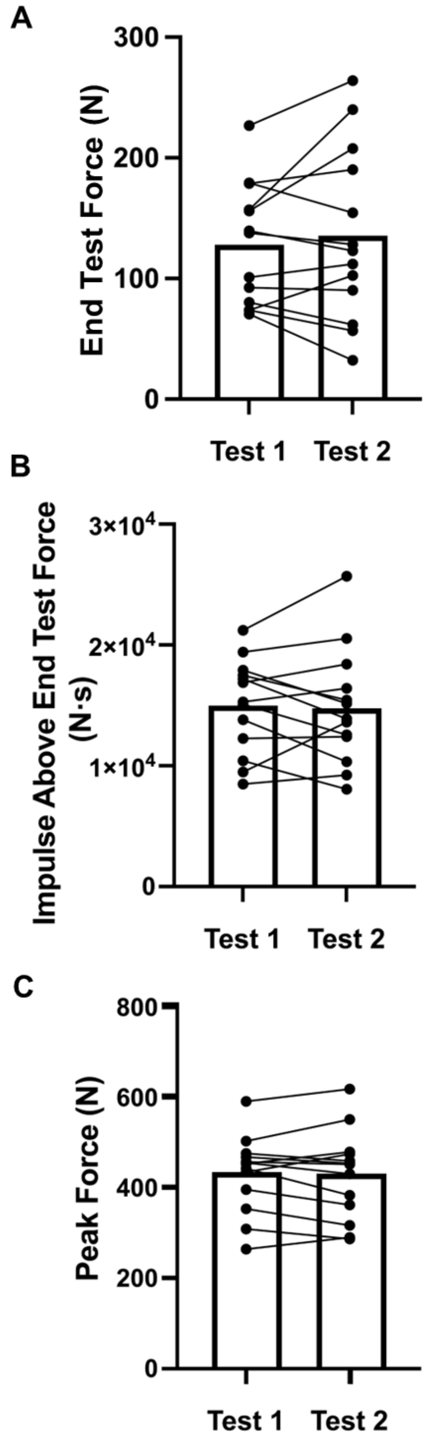

Figure 2. Group mean and individual (n = 13) differences between Test 1 and Test 2 for end-test force, impulse above end-test force, and peak force.

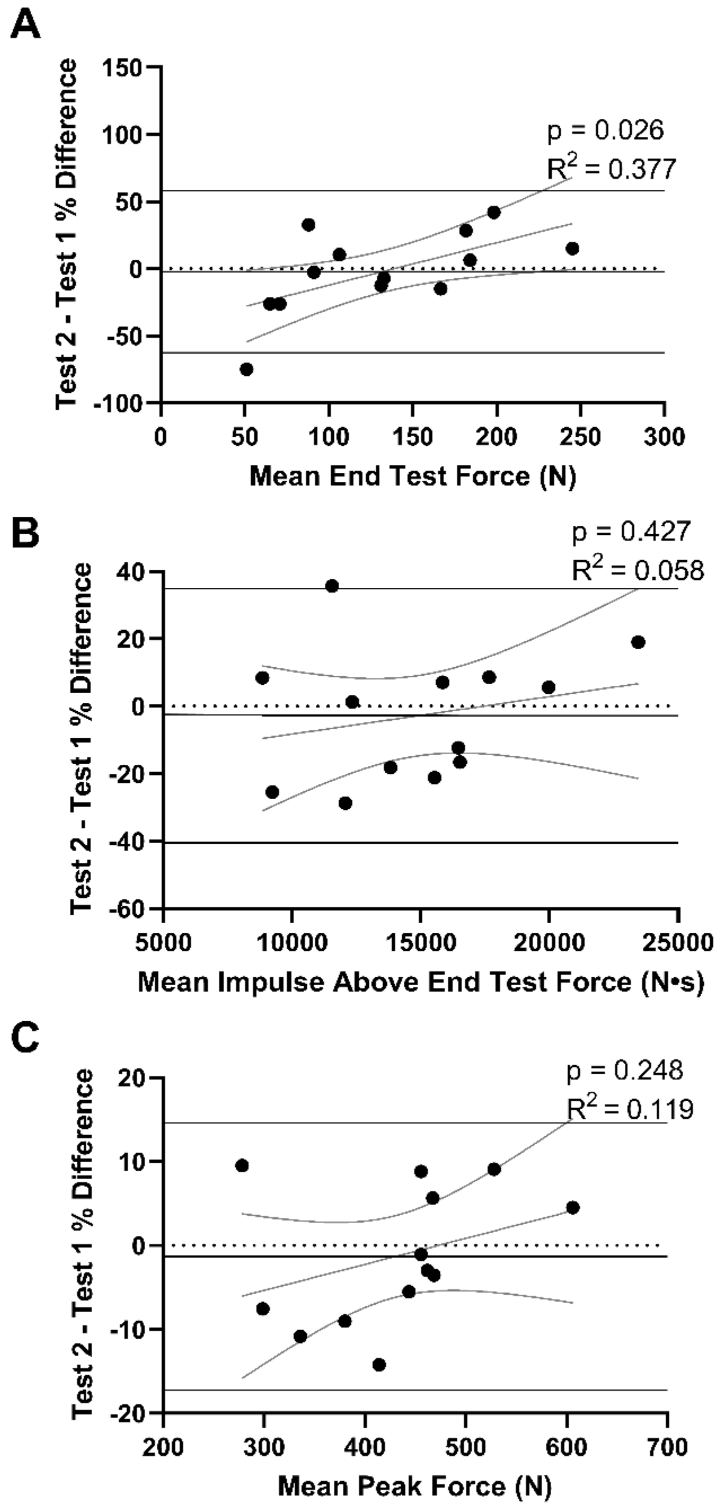

Figure 3. Bland-Altman analyses (n = 13) showing the mean relative bias and 95% limits of agreement between Test 1 and Test 2 for end-test force, impulse above end-test force, and peak force. From the regression of the differences, a proportional bias was observed for end test force ($p = 0.026$, $R^2 = 0.0377$).

DISCUSSION

This study aimed to assess the test-retest reliability of several parameters estimated from a 4-min all-out critical force test in trained rock climbers. The relative reliability of the EF and IEF, which corresponds to the degree which individuals maintain their position in a sample of with repeated measurements (4), was good, while that of peak force was excellent. Similarly, the absolute reliability, i.e., the degree to which repeated measures varies for individuals (4), was highest for peak force, followed by IEF and EF, respectively.

In comparison to the pilot data of Giles *et al.* (17), we reported a larger bias and LoA for EF. One potential explanation for the discrepancy in results here could be that the level of experience in climbing and/or hangboard training was different between the two studies. Although the average IRCRA grade in our study (18.4) was like that of Giles and colleagues' full study sample (18.1), the reliability analysis conducted by these researchers was limited to a sub-sample of participants for whom experience or climbing ability was not reported. While Giles *et al.* (17) did not report the CV of their pilot testing, our results showed CV for EF of 21%, which is higher than previous studies reporting the reliability of 10-min all out handgrip test (CV = 5.5%) (21), but similar to that reported for time to exhaustion trials in running and cycling (8). It is important to note, however, that the individual variability we observed for EF was substantially affected by a single participant, who showed ~ 54% decrease between test 1 and test 2. When this participant was removed from the analysis, the CV improved to a much more acceptable 15.6 %. Moreover, when this data point was removed from the regression of the differences, no proportional bias was present ($p = 0.094$, $R^2 = 0.256$). Although removing this outlier from the statistical analysis may have been acceptable in some contexts, considering the aim of this study was to determine individual variability of the test performance, it was decided to keep the outlier in the analysis.

While not exactly comparable to the critical force test used here, Michailov *et al.* (23) assessed the test-retest reliability of the percentage decrease in force from the beginning to the end of a 30 s maximal continuous hang, using a comparable experimental setup and similarly trained climbers to our group. For this measure they reported an ICC of 0.701 and a large LoA, similarly indicating that the end-test force during single trial all-out tests may have relatively high session-to-session variability. Interestingly, compared to EF, our mean absolute bias (-241.4 N·s) and 95% LoA for IEF (-5548 to 5065 N·s) were smaller than that reported in the pilot data of Giles *et al.* (17) (mean absolute bias = -1363.6 N·s, LoA = -9574 to 6847). Some of this difference may be a result of how IEF was calculated, whereby while we summed the integrals above EF for each individual contraction, Giles *et al.*,(17) appeared to calculate IEF as one continuous integral across the entire 4-min test. Nevertheless, as the IEF is reflective of the overall force profile during the test, despite observing consistent values for end-test force, the large variability in IEF suggests that there may have been substantial individual variation in the overall performance of the test.

The comparatively lower absolute reliability of the EF estimated here compared to parameters such as critical power derived from all-out cycling tests (27,29) can likely be ascribed to the difference in potential variability in more continuous versus intermittent exercise protocols. During continuous cycling, power output is averaged across many revolutions, resulting in a more stable output across a given period. Alternatively, EF here was estimated as the average force across only six contractions at the end of the test. Thus, if any one of these contractions was markedly different than the others, for example due to poor finger positioning on the hangboard, or subconscious pacing strategies nearing the end of the test, this may have a substantial effect on the estimated critical force. Furthermore, in contrast to activities in which limbs are essentially fixed to the measurement instrument, for example with clipless pedals in cycling, or many dynamometer setups, the downward force applied by the fingers on a small wooden rung is greatly dependent on the coefficient of friction between the skin and the rung (13). While testing was completed in an indoor, temperature controlled environment, and participants were able to re-chalk their fingers when needed, factors such as varying skin moisture content (e.g., see Gerhardt *et al.* (15)), or perhaps variation in freely chosen chalking strategies (see Amca *et al.* (1)), could have impacted the capacity to apply force to the rung. If factors such as these were different day to day, this could have impacted the consistency of measurements.

Conventionally, the critical intensity is estimated by plotting the time to exhaustion from 3-5 constant intensity trials performed in separate sessions. Employing this method in a two-handed hangboarding protocol, Giles *et al.* (16) determined the agreement between critical force and W′ estimated from 3 constant intensity trials to exhaustion performed on multiple days compared to that done in a single session with a 20 min recovery between trials. They reported excellent relative reliability for critical force (ICC = 0.900), but not for W′ (ICC = 0.768), although the 95% LoA for critical force still demonstrated considerable individual variation. Presumably, a large part of the variability in these results was due to the accumulated muscle fatigue when performing multiple trials in the same session. Additionally, compared to multiple submaximal trials during conventional testing, it is plausible that all-out tests might require greater mental engagement, which could be another source of intra-individual variability. Thus, compared to a single trial/session, it may be assumed that spreading the variability across multiple sessions may provide a more accurate reflection of the true value. Currently, there has been no published work examining the agreement between the critical force estimated from an all-out trial and that estimated from multiple constant-intensity trials. However, a valid protocol must be reliable as well, and so in any case, researchers, coaches, and athletes must weigh the pros and cons between single versus multi-trial tests in regard to their overall time efficiency, as well as their degree of precision.

One limitation of our study is the relatively small sample size employed. Practically speaking, however, any conclusions made about the importance of recognizing the potential for considerable individual variation in the performance of the 4-min all-out critical force test would still stand. In this regard, despite including a familiarization trial, we reported a proportional bias for EF, in which there tended to be greater improvement from test 1 to test 2 in those participants with a higher EF. This improvement could not be due to a training effect of the testing, because if this was the case, those participants with a *lower* baseline would be expected to improve to a relatively greater extent. Alternatively, this performance improvement may be explained by changes in motivation between sessions, where the stronger individuals were perhaps more motivated to try to improve their previous performances compared to the other participants. This notion once again relates to the potential impact that mental engagement may have on performance during all-out tests, and should be considered in their implementation. One other consideration in this regard is the mixed-sex sample used in our study, which may have impacted our results. Although most recent studies indicate that menstrual cycle phase has a negligible effect on performance (7), others have reported lower reliability measures for muscle endurance tasks (but not strength tasks) for females compared to males (14). This may be an additional consideration by coaches or researchers interested in utilizing this test for evaluation or prescriptive purposes.

Overall, we found that critical force estimated from a 4 min all-out trial had good relative reliability, but larger within subject variation. Comparatively, IEF estimated from the total impulse above the end test force had a similar relative reliability as the end test force, but a higher absolute reliability. Peak force had the highest relative and absolute reliability of all measured variables. Considering the significant correlations previously reported between these measures and climbing ability (17,23), this time-efficient test may be useful as part of a battery of tests designed to differentiate climbers of different abilities or climbing potential, and monitor athlete performance over time. However, compared to some other activities, rock climbing or hangboarding tasks may involve more variables that are difficult to control, such as changes in friction, that may impact the consistency of some performance measures. Thus, some caution should be taken in the interpretation of performance changes from single trials, and it is recommended that coaches and athletes weigh the pros and cons of time efficiency of single trial tests compared to a potentially more robust assessment of monitoring performance trends over a longer time period and a larger number of trials.

REFERENCES

1. Amca AM, Vigouroux L, Aritan S, Berton E. The effect of chalk on the finger-hold friction coefficient in rock climbing. Sports Biomech 11(4): 473–9, 2012.

2. Baláš J, Michailov M, Giles D, Kodejška J, Panáčková M, Fryer S. Active recovery of the finger flexors enhances intermittent handgrip performance in rock climbers. Eur J Sport Sci 16(7): 764–72, 2016.

3. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 8(2): 135–60, 1999.

4. Bruton A, Conway JH, Holgate ST. Reliability: What is it, and how is it measured? Physiotherapy 86(2): 94–9, 2000.

5. Burnley M. Estimation of critical torque using intermittent isometric maximal voluntary contractions of the quadriceps in humans. J Appl Physiol 106(3): 975–83, 2009.

6. Burnley M, Doust JH, Vanhatalo A. A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. Med Sci Sports Exerc 38(11): 1995–2003, 2006.

7. Colenso-Semple LM, D'Souza AC, Elliott-Sale KJ, Phillips SM. Current evidence shows no influence of women's menstrual cycle phase on acute strength performance or adaptations to resistance exercise training. Front Sports Act Living 5: 1054542, 2023.

8. Currell K, Jeukendrup AE. Validity, reliability, and sensitivity of measures of sporting performance. Sports Med 38(4): 297–316, 2008.

9. Dekerle J, Vanhatalo A, Burnley M. Determination of critical power from a single test. Sci Sports 23(5): 231–8, 2008.

10. Draper N, Dickson T, Blackwell G, Fryer S, Priestley S, Winter D, et al. Self-reported ability assessment in rock climbing. J Sports Sci 29(8): 851–8, 2011.

11. Draper N, Giles D, Schöffl V, Konstantin Fuss F, Watts P, Wolf P, et al. Comparative grading scales, statistical analyses, climber descriptors, and ability grouping: International Rock Climbing Research Association position statement. Sports Technol 8(3–4): 88–94, 2015.

12. Florence S, Weir JP. Relationship of critical velocity to marathon running performance. Eur J Appl Physiol 75(3): 274–8, 1997.

13. Fuss FK, Niegl G. The importance of friction between hand and hold in rock climbing. Sports Technol 5(3–4): 90–9, 2012.

14. García-Vaquero MP, Barbado D, Juan-Recio C, Lopez-Valenciano A, Vera-Garcia FJ. Isokinetic trunk flexion–extension protocol to assess trunk muscle strength and endurance: Reliability, learning effect, and sex differences. J Sport Health Sci 9(6): 692–701, 2020.

15. Gerhardt LC, Strässle V, Lenz A, Spencer ND, Derler S. Influence of epidermal hydration on the friction of human skin against textiles. J R Soc Interface 5(28): 1317–28, 2008.

16. Giles D, Chidley JB, Taylor N, Torr O, Hadley J, Randall T, et al. The determination of finger-flexor critical force in rock climbers. Int J Sports Physiol Perform 14(7): 972–9, 2019.

17. Giles D, Hartley C, Maslen H, Hadley J, Taylor N, Torr O, et al. An all-out test to determine finger flexor critical force in rock climbers. Int J Sports Physiol Perform 16(7): 942–9, 2021.

18. Giles LV, Rhodes EC, Taunton JE. The physiology of rock climbing. Sports Med 36(6): 529–45, 2006.

19. Hill DW. The critical power concept: A review. Sports Med 16: 237–54, 1993.

20. Hyslop NP, White WH. Estimating precision using duplicate measurements. J Air Waste Manag Assoc 59(9): 1032–9, 2009.

21. Kellawan JM, Tschakovsky ME. The single-bout forearm critical force test: A new method to establish forearm aerobic metabolic exercise intensity and capacity. PLoS ONE 9(4): e93481, 2014.

22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15(2): 155–63, 2016.

23. Michailov ML, Baláš J, Tanev SK, Andonov HS, Kodejška J, Brown L. Reliability and validity of finger strength and endurance measurements in rock climbing. Res Q Exerc Sport 89(2): 246–54, 2018.

24. Navalta JW, Stone WJ, Lyons TS. Ethical issues relating to scientific discovery in exercise science. Int J Exerc Sci 12(1): 1-8, 2019.

25. Poole DC, Burnley M, Vanhatalo A, Rossiter HB, Jones AM. Critical power: An important fatigue threshold in exercise physiology. Med Sci Sports Exerc 48(11): 2320-34, 2016.

26. Smith JC, Dangelmaier BS, Hill DW. Critical power is related to cycling time trial performance. Int J Sports Med 20(06): 374–8, 1999.

27. Vanhatalo A, Doust JH, Burnley M. Determination of critical power using a 3-min all-out cycling test. Med Sci Sports Exerc 39(3): 548–55, 2007.

28. Wakayoshi K, Ikuta K, Yoshida T, Udo M, Moritani T, Mutoh Y, et al. Determination and validity of critical velocity as an index of swimming performance in the competitive swimmer. Eur J Appl Physiol 64(2): 153–7, 1992.

29. Wright J, Bruce-Low S, Jobson SA. The reliability and validity of the 3-min all-out cycling critical power test. Int J Sports Med 38(06): 462–7, 2017.