J Sci Cycling. Vol. 3(3), 29-33

RESEARCH ARTICLE **Open Access**

The reliability of performance during computer-simulated varying gradient cycling time trials

Bradley Clark[1]✉, Carl D Paton [2], Brendan J O'Brien [1]

Abstract

Ergometer based time trials are commonly used to assess performance changes due to training or other interventions. This investigation establishes the reliability of a novel computer simulated cycling time trial. Nineteen cyclists (age: 32 ± 12 years, mass 73 ± 11 kg, height 178 ± 5 cm) completed four time trials over a 20-km course which included numerous changes in gradient. The time trials were completed over a 4-week period in order to establish both short and long-term reliability. Performance time (mean ± SD) for trials one to four was 2265 ± 149 s, 2252 ± 153 s, 2236 ± 146 s and 2240 ± 154 s respectively; the corresponding power output for consecutive trials was 293 ± 35 W, 297 ± 36 W, 299 ± 35 W and 299 ± 35 W. The coefficient of variation (± 90% confidence limits) of performance for trials separated by 7, 14, 21 and 28 days was 1.1% (0.8% – 1.5%), 1.3% (1.1% – 1.9%), 1.3% (1.1% – 1.9%) and 1.5% (1.1% – 2.1%) respectively for time; the corresponding values for power output were 2.0% (1.5% – 2.7%), 2.3% (1.8% – 3.2%), 2.6% (2.0% – 3.6%) and 3.2% (2.5% – 4.5%). Further analysis based on rider ability indicated slower riders were less reliable than faster riders by a factor of ~1.1. Reliability of time trial performance diminishes with increasing time between trials. Additionally, faster riders show better reliability than slower riders over time. Researchers should consider the effect of time between trials and athlete ability when making conclusions about intervention effectiveness.

Keywords: athlete; variability; enhancements; competition; power output; exercise test

✉**Contact email:** *bclark@federation.edu.au* (B. Clark)

[1] Federation University Australia

[2] Eastern Institute of Technology, New Zealand

Received: 8 April 2014. Accepted: 2 October 2014.

Introduction

Laboratory based assessments of physiology and performance form an integral part of athlete monitoring and preparation for competition. Establishing the physiological capacities and performance standards of athletes, allows sports scientists and coaches to assess the effectiveness of training programmes and other experimental interventions. The performance capabilities of competitive cyclists are often assessed using simulated time trials completed under controlled conditions in a laboratory. Laboratory based cycling trials can take several forms (Hopkins et al. 2001), and there has been considerable debate on the advantages and disadvantages of the different types of test (Currell and Jeukendrup 2008). However from an ecologically valid perspective, fixed distance self-paced time trials most closely represent a true competitive situation and are often the preferred option when investigating athlete performance enhancement strategies.

Irrespective of the test design, any test must have good reliability to monitor the small changes in performance that matter to competitive athletes (Hopkins et al. 1999; Paton and Hopkins 2006). Several previous studies have investigated the reliability of different types of time trial protocols. The re-test reliability (reported as a coefficient of variation) for simulated cycling time-trials of ~30-60 minutes duration, completed on a flat course and bereft of changes in gradient or prescribed changes in intensity is reportedly between 0.7%-1.5% and 1.9%-3.6% for time and power respectively (Smith et al. 2001; Sporer and McKenzie 2007; Zavorsky et al. 2007). Similar reliability measures have also been reported for time (1.4%-2.9%) and power (1.7-3.5%) during a simulated up-hill time trial completed on a constant gradient 8-mile course (Noreen et al. 2010). In a more recent study Driller et al. (2013) reported excellent reliability (~1.3% for power) for a short duration 15- minute self-paced time trial following a 15-minute pre-load activity at a fixed intensity. However, whilst these previous studies have reported the reliability of performance measures between consecutive trials over short intervening periods (typically 1-10 days between trials), none have reported the effects of increasing time between trials on test reliability. Further, a common issue with these previous studies is they lack the variations in the external environment that are typically seen during real competitions.

Unlike traditional laboratory based time trials, competitive cycling events typically take place on public roadways and as such consist of constant changes in road gradient. Perception of these changes in combination with internal physiological feedback mechanisms combine to determine how an individual cyclists adjusts pace and effort (Atkinson et al. 2007). Pacing strategy is therefore adjusted according to perception of the internal and external environment by important brain centres (Atkinson et al. 2007). Currell and Jeukendrup (2008) suggest that any laboratory measure of sporting performance should allow participants to adopt a pacing strategy similar to that which is required by competitive situations. By providing a constant external environment, most laboratory test protocols do not challenge the perceptive skills of the cyclists and present a testing stimulus that is unlike competitive situations. In one of the few studies to examine reliability of performance when the test required substantial changes in intensity, Schabort et al. (1998) reported short-term (>7days between trials) reliability for both total time and repeated high intensity efforts (1-km and 4-km) time of ~2% during a simulated 100-km time trial. Conversion of this reliability in time to an equivalent mean power yields relatively poor reliability of ~3.7% (Hopkins et al. 1999). In a more recent modification of the Schabort et al. (1998) study using a shorter duration 30-km time trial, Abbiss et al. (2008) reported reliability in mean power of 2.4% after subjects had completed a familiarisation session. Interestingly in their study, Abbiss et al. (2008) reported a large decrease in test reliability (~11%|) when trials were separated by large intervening periods.

While these two previous studies address some of the issues associated with variations in pace during laboratory based time trials, they do not fully simulate a competition situation requiring almost constant changes in pacing strategy in response to variation in the external environment. However the development of new computer technology and bicycle ergometers which allow accurate simulation of real race course profiles provides an opportunity to study the effects of scientific interventions in a more realistic environment. Therefore the aim of this study was to establish the short and long term re-test reliability of a novel computer simulated cycling time trial completed on a course of varying gradients.

Materials and methods

Participants

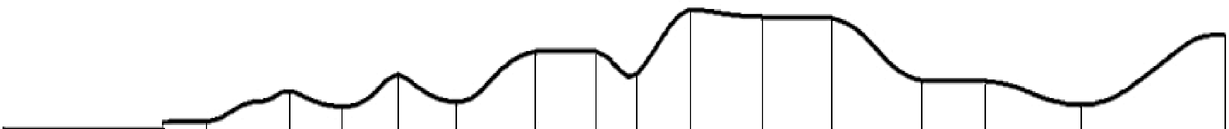
Nineteen competitive cyclists (17 males, 2 females) volunteered to participate in this study (Age: 32 ± 12 years, mass 73 ± 11 kg, height 178 ± 5 cm). All cyclists were well-trained with a minimum of two years racing experience at an A or B grade standard. All testing was performed in the athlete's competition phase of the season. Participants were free from illness or injury and gave their written informed consent to participate in the study. The study was carried out in accordance with the ethical and procedural requirements of the journal (Harriss and Atkinson 2013) and approved by the institutional human research ethics committees.

Design and procedure

The study was a repeated measures design requiring cyclists to complete four simulated 20-km cycling time trials at set time intervals. Trials one to two and two to three were separated by 7 days and trials three to four by 14 days. Each trial was completed at a similar time of day (±2 hours) and was preceded by a standardized 20 minute warm up. Participants were instructed to treat each trial as it was an important competition and refrain from vigorous exercise and maintain a consistent diet in the 24 hours before each trial. Cyclists were requested not to consume any alcohol, caffeine or other substances that may affect performance in the 12 hours immediately preceding each trial.

Methology

All test sessions were completed on a Velotron Dynafit Pro cycle ergometer (RacerMate Inc, WA, USA) using the company's associated 3D race course software. Prior to the first trial, the Velotron factory calibration was confirmed according to manufacturer instructions using the "Accuwatt" function. During the first session each participant was fitted to the ergometer in a manor to replicate their own racing bicycle. The fit measurements were recorded and repeated for each subsequent testing session. Cyclists initially completed a 20 minute standardised warm up consisting of three repeated increasing intensity bouts. The first two minutes were completed at 2-2.5 $W.kg^{-1}$, followed by two minutes at 3-3.5 $W.kg^{-1}$ and finally one minute at 4-4.5 $W.kg^{-1}$ repeated consecutively. For the final five minutes cyclist pedalled at a fixed intensity of 100W. The time trial was completed on an experimenter designed course which replicated a typical racing circuit and contained numerous changes in gradient represented by both ascents and descents as shown in Figure 1. The total elevation gain over the 20km was 300 meters leading to an average gradient of ~1.5%.



Distance (km)	.5	1.6	1.05	1.0	1.1	1.4	1.0	.6	1.0	1.3	1.3	1.7	1.1	1.9	2.2
Grade (%)	0	2.5	-2.0	4.0	-3.2	4.6	0	-8	10	-1.5	0	-4.7	0	-1.5	4.1

Figure 1. The computer simulated course showing the profile of the varying gradient time trial used in the study.

Participants were able to view their progress over the course on a computer monitor and were provided with information on distance completed and gear selected; all other information was blinded to remove any potential pacing feedback. Participants were requested to complete each time trial as quickly as possible with no restriction on gear selection, cadence or cycling posture (seated or standing). Participants were not restricted to a set pacing strategy, were not coached on how to best ride the course and in order to control for extrinsic motivation, no encouragement was given to cyclists during the trials. Throughout the trial participants were cooled by two 30 cm pedestal fans and were able to consume water *ad libitum*.

Table 1. Performance characteristics for each trial for all cyclists, and sub-groups of fastest and slowest cyclists, Mean ± SD

	$T_{all\ (s)}$	$T_{fast\ (s)}$	$T_{slow\ (s)}$	W_{all} (W)	W_{fast} (W)	W_{slow} (W)
Test 1	2265 ± 149	2153 ± 87	2390 ± 90	293 ± 35	314 ± 28	269 ± 26
Test 2	2252 ± 153	2137 ± 85	2379 ± 98	297 ± 36	320 ± 28	271 ± 24
Test 3	2236 ± 146	2122 ± 75	2363 ± 83	299 ± 35	323 ± 23	273 ± 26
Test 4	2240 ± 154	2115 ± 68	2379 ± 85	299 ± 35	324 ± 20	271 ± 25
Mean	2248 ± 151	2132 ± 79	2378 ± 89	297 ± 35	320 ± 25	271 ± 26

Abbreviations: T_{all} = performance time all cyclists; T_{fast} = performance time fastest cyclists; T_{slow} = performance time slowest cyclists; W_{all} = mean power all cyclists; W_{fast} = mean power fastest cyclists; W_{slow} = mean power slowest cyclists.

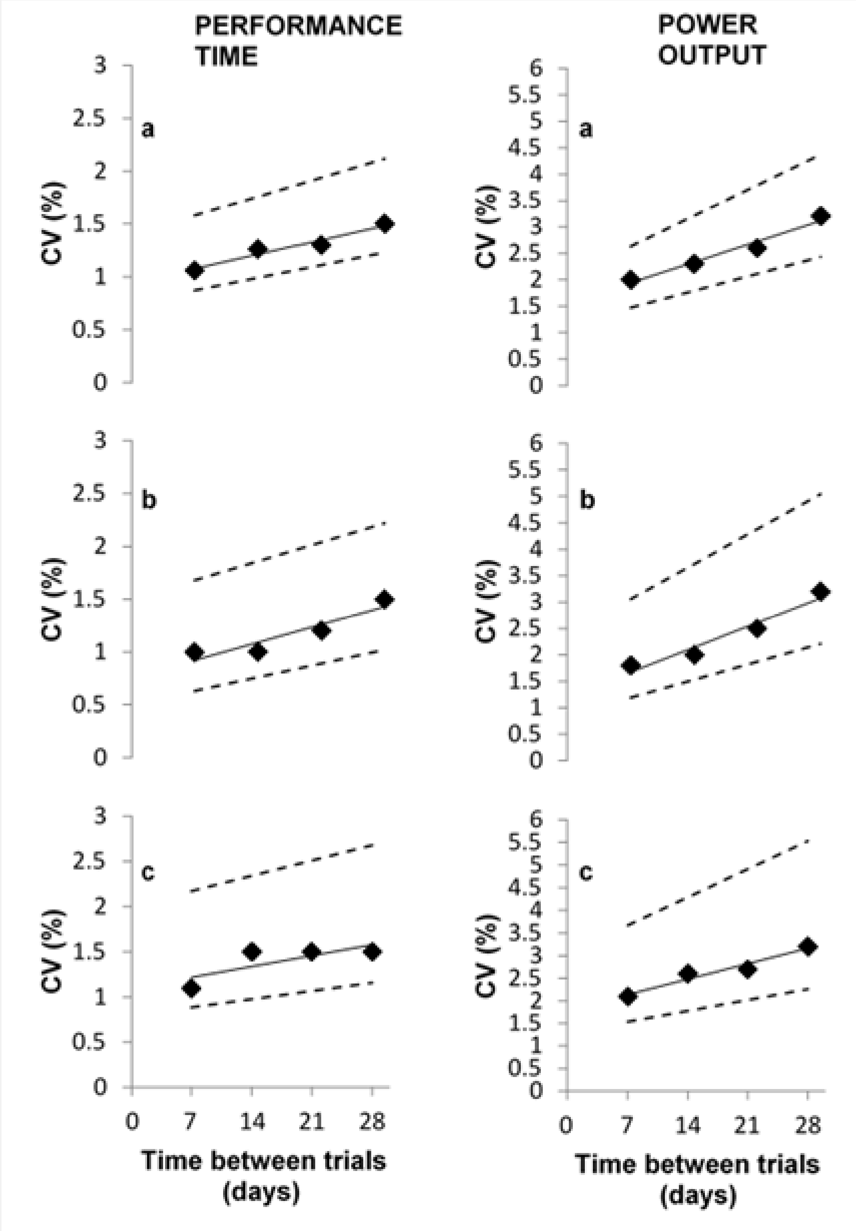


Figure 2. Coefficient of variation (CV) for time and power output (±90% CI) as time increases between trials for all cyclists (a), fastest cyclists (b) and slowest cyclists (c).

Statistical analysis

Simple descriptive statistics are shown as means ± between-subject standard deviations. All measures were log transformed to reduce bias arising from non-uniformity of error and analysed using a made for purpose Excel spread sheet for reliability analysis (Hopkins 2006). Typical error was determined as coefficients of variation (CV%) along with their 90% confidence intervals (CI). The spreadsheet also provided the intra-class correlation (± 90% CI) between trials. Analysis was performed for all subjects together and as separate analysis for the fastest (n=10) and slowest (n=9) sub-groups in the time trial.

Results

Table 1. shows the time and power output (mean ± SD) for all cyclists, and the sub-groups of fastest and slowest cyclists across all four trials. The change in mean of the performance variable represents the size of any learning effect between trials. For all cyclists there was a change of -0.6%, -0.7% and 0.2% in mean performance time between consecutive trials; the corresponding change in mean power between consecutive trials was 1.3%, 0.9% and -0.1% respectively. The magnitude of the mean change between trials was largest from trial 1-2 and reduced with subsequent trials, however all changes were deemed trivial (ES<0.2) The fastest subgroup of cyclists was ~10% faster and produced ~18% more power across all four trials than the slower sub-group.

Figure 2. shows the coefficient of variation of performance for trials separated by 7, 14, 21 and 28 days. The CV for seven days was calculated by taking the average CV from tests 1-2 and 2-3, 14 days by taking the average CV from tests 1-3 and 3-4, 21 days by taking the CV from tests 2-4 and 28 days the CV from tests 1-4. The variation in performance time for all cyclists' increased linearly from 1.1% to 1.5% with increasing time between trials. Similarly the variation for mean power increases from 2.0% to 3.2% with increasing time between trials. The faster cyclists were marginally more reliable than the slower cyclists over the short term (7-14 days between trials) but there were no substantial differences in reliability between sub-groups over the longer term.

Table 2. The changes in intra-class correlation coefficient (ICC ± 90% CI) for all cyclists, fastest cyclists and slowest cyclists with increasing time between trials.

	T_{all}	T_{fast}	T_{slow}	W_{all}	W_{fast}	W_{slow}
ICC_{7days}	0.98 (0.960.99)	0.95 (0.81-0.99)	0.93 (0.79-0.98)	0.98 (0.95-0.99)	0.97 (0.87-0.99)	0.97 (0.89-0.99)
ICC_{14days}	0.97 (0.940.99)	0.95 (0.84-0.98)	0.86 (0.58-0.96)	0.97 (0.94-0.99)	0.95 (0.84-0.98)	0.95 (0.84-0.99)
ICC_{21days}	0.97 (0.930.98)	0.92 (0.78-0.98)	0.88 (0.65-0.97)	0.96 (0.91-0.98)	0.92 (0.76-0.97)	0.94 (0.81-0.98)
ICC_{28days}	0.96 (0.910.98)	0.87 (0.65-0.96)	0.88 (0.64-0.96)	0.94 (0.87-0.97)	0.87 (0.63-0.96)	0.92 (0.76-0.98)

Abbreviations: T_{all} = performance time all cyclists; T_{fast} = performance time fastest cyclists; T_{slow} = performance time slowest cyclists; W_{all} = mean power all cyclists; W_{fast} = mean power fastest cyclists; W_{slow} = mean power slowest cyclists; 7days = seven days between trials; 14days = 14 days between trials; 21days = 21 days between trials; 28days = 28 days between trials.

Table 2, shows the intra-class correlations (± 90% CI) for performance time and power output for all cyclists, and sub-groups of fastest and slowest cyclists as time increases between trials. A gradual decline in reliability is evident for time and power with increasing time between trials.

Discussion

The major findings of the present study is that a novel laboratory based simulated cycling time trial performed on a course of varying gradients is a reliable test in terms of time (~1.2%) and power output (~2%) with competitive cyclists when trials are separated by less than 14 days. However reliability of performance declines substantially as time between trials increases beyond this period. In addition it was evident that faster cyclists were more reliable in the short term in comparison to their slower counterparts, though this finding was not apparent when trials were separated by longer intervening trial periods. We also found evidence of a learning effect between particularly between trials 1-2; though this was deemed statistically trivial. Evidence of a learning effect, all be it small, is a finding consistent with previous studies (Abbiss et al. 2008; Noreen et al. 2010; Zavorsky et al. 2007) and adds support to the requirement of at least one familiarisation trial for subjects prior to performing any experimental study trials.

The observed short term (7-14 days between trials) reliability for performance in our study was similar to, and in some cases better, than the short term reliability reported in previous studies using constant grade time trials (Noreen et al. 2010; Smith et al. 2001; Sporer and McKenzie 2007; Zavorsky et al. 2007). However, a unique aspect of our study is the inclusion of frequent variations in terrain which we may have expected to increase performance variation compared to a constant gradient time trial. Importantly, the similarity in short term reliability between this study and others indicates the presence of changes in gradient does not appear to adversely affect the tests reliability.

The variation in performance we report here is also substantially smaller than that reported in previous studies using dynamic changes in effort over both 100-km and 30-km distances (Abbiss et al. 2008; Schabort et al. 1998). The reasons for the better reliability in the current study are unclear, since both the previous studies used cyclists of similar ability. However a possible explanation relates to the differing nature of the dynamic tests. In both previous dynamic studies cyclist were required to perform set periods (0.25-4-km) of high-intensity activity during the trial when instructed by the researchers, whereas in the current study the cyclists were free to modify their intensity in response to their perceived feelings at the time. The ability to make smaller but continuous modifications to exercise intensity may have allowed the athletes in our study to adopt a more even pacing strategy and this therefore may lead to better reliability. It is also possible the shorter distance in the current study influenced reliability, as longer distances would allow for greater errors in a cyclists self-pacing strategy to manifest. Clearly changes in feeding, for example, during a 100-km trial would have a much bigger effect on pacing than during a 20-km trial.

We also observed a substantial decrease (Fig 2.) in reliability of cycling performance with increasing time between trials. The decrease in reliability over time is consistent with the findings of Abbiss et al. (2008) who reported a very large decline in reliability (CV of ~11%) when time-trials were separated by six-weeks. A likely explanation for the increased variation in performance within our study (and that of previous studies) is during long intervening periods subjects simply lose their perception of the appropriate pacing strategy. It is also likely individual variations in fitness over longer time-periods contribute to greater variations in performance within a study group.

Separate analysis of reliability based on cyclists ability in our study also indicated the faster cyclists were more reliable in performance than slower cyclists (CV~1.9% & 2.4% respectively) at least in the short term; this finding is in agreement with previous investigations (Zavorsky et al. 2007). However, reliability declined linearly in both groups with increasing time between trials and was similar after a 28-day period. Irrespective of athlete ability, the decrease in trial reliability over time has important implications for studies examining training and other interventions where time between experimental trials exceeds 14-days. In these situations we would recommend that researchers perform regular re-habituation trials so that subjects might remain familiar with testing conditions. Theoretically this could improve the ability to detect meaningful and important changes to performance in experimental studies with a large intervening time period between pre and post testing.

Conclusions

A novel computer simulated cycling time trial completed over a course of varying gradient is a reliable measure of performance, when trials are separated by short intervening periods. However a substantial decline in performance reliability was evident when more than 14 days elapsed between trials. Furthermore, faster cyclists were generally more reliable in performance than slower cyclists over the short term though any differences were insubstantial over the longer term. Future studies are needed to confirm the reliability of variable gradient time trials and determine the effects of individual variations in fitness on test reliability.

Practical applications

The novel protocol investigated in the present study may detect meaningful changes in performance that matter to athletes and can therefore be used by coaches and sports scientists to examine the efficacy of training and other scientific interventions. However continued habituation is necessary in all cyclists when a larger period of time elapses between trials. Habituation could be achieved by including the performance trial as part of any training intervention in long duration experimental trials or as a prescribed training session if monitoring performance throughout a competitive season. There was also evidence of a small learning effect between trials 1-2 and we therefore recommend that all athletes undertake a familiarisation session prior to any experimental study.

Acknowledgment

The authors would like to thank the cyclists for their commitment to this study. We would also like to thank Professor Will Hopkins from the Auckland University of technology for his valuable advice on statistical analysis. The authors did not receive any external funding for this research.

Conflict of interest

There are no conflicts of interest for any of the contributing authors of this manuscript.

References

1. Abbiss CR, Levin G, McGuigan MR, Laursen PB (2008) Reliability of power output during dynamic cycling. International Journal of Sports Medicine 29 (7): 574-578
2. Atkinson G, Peacock O, St Clair Gibson A, Tucker R (2007) Distribution of power output during cycling: impact and mechanisms. Sports Medicine 37 (8): 647-667
3. Currell K, Jeukendrup AE (2008) Validity, reliability and sensitivity of measures of sporting performance. Sports Medicine 38 (4): 297-316
4. Driller MW, Argus CK, Bartram JC, Boniventura J, Martin DT, West NP, Halson SL (2013) The reliability of a two-bout exercise test on a Wattbike cycle ergometer. International Journal of Sports Physiology and Performance 8 (4): 379-383
5. Harriss DJ, Atkinson G (2013) Ethical standards in sport and exercise science research: 2014 update. International Journal of Sports Medicine 34 (12): 1025-1028
6. Hopkins WG (2006) Spreadsheets for analysis of controlled trials with adjustment for a predictor. Sportscience 10: 46-50
7. Hopkins WG, Hawley JA, Burke LM (1999) Design and analysis of research on sport performance enhancement. Medicine and Science in Sports and Exercise 31 (3): 472-485
8. Hopkins WG, Schabort EJ, Hawley JA (2001) Reliability of power in physical performance tests. Sports Medicine 31 (3): 211-234
9. Noreen E, Yamamoto K, Clair K (2010) The reliability of a simulated uphill time trial using the Velotron electronic bicycle ergometer. European Journal of Applied Physiology 110 (3): 499-506
10. Paton CD, Hopkins WG (2006) Variation in performance of elite cyclists from race to race. European Journal of Sport Science 6 (1): 25-31
11. Schabort EJ, Hawley JA, Hopkins WG, Mujika I, Noakes TD (1998) A new reliable laboratory test of endurance performance for road cyclists. Medicine and Science in Sports and Exercise 30 (12): 1744-1750
12. Smith MF, Davison RCR, Balmer J, Bird SR (2001) Reliability of mean power recorded during indoor and outdoor self-paced 40 km cycling time-trials. International Journal of Sports Medicine 22 (04): 270,274
13. Sporer BC, McKenzie DC (2007) Reproducibility of a laboratory based 20-km time trial evaluation in competitive cyclists using the Velotron Pro ergometer. International Journal of Sports Medicine 28 (11): 940-944
14. Zavorsky GS, Murias JM, Gow J, Kim DJ, Poulin-Harnois C, Kubow S, Lands LC (2007) Laboratory 20-km cycle time trial reproducibility. International Journal of Sports Medicine 28 (9): 743-748