



CREaTE

Canterbury Research and Theses Environment

Canterbury Christ Church University's repository of research outputs

<http://create.canterbury.ac.uk>

Please cite this publication as follows:

Hurst, P. and Board, L. (2016) Reliability of 5-km running performance in a competitive environment. *Measurement in Physical Education and Exercise Science*. pp. 1-5. ISSN 1091-367X.

Link to official URL (if available):

<http://dx.doi.org/10.1080/1091367X.2016.1233421>

This version is made available in accordance with publishers' policies. All material made available by CReaTE is protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.

Contact: create.library@canterbury.ac.uk



- 1 **Article title:** Reliability of 5km running performance in a competitive environment
- 2 **Running head:** Reliability of a competitive 5 km time-trial

3 **Abstract**

4 The aim of this study was to examine the reliability of a 5 km time-trial during a competitive
5 outdoor running event. Fifteen endurance runners (age = 29.5 ± 4.3 years, height = $1.75 \pm$
6 0.08 m, body mass = 71.0 ± 7.1 kg, 5 km lifetime personal best = $19:13 \pm 1:13$ minutes)
7 completed two competitive, 5 km time-trials over two weeks. No systematic differences in
8 run time between Trial 1 and Trial 2 were reported (Trial 1; 1217 ± 85 s, 95% CI [1170,
9 1264] and Trial 2; 1216 ± 79 s, [1172 to 1260], $p = .855$). Absolute reliability, expressed as
10 the typical error (TE; 14.7s, 95% CI = 11.3 to 21.4 s) and coefficient of variation (CV; $0.95 \pm$
11 0.65% , [0.59 to 1.31]) confirms the reliability of 5 km running performance in a competitive
12 time trial.

13 **Key words**

14 Ecological validity, reproducibility, time-trial, competition, performance

15 Researchers investigating the efficacy of an intervention must use a test that has high
16 reliability (Currell & Jeukendrup, 2008). For this reason, numerous studies have examined
17 the reliability of running time-trial performance in the laboratory. Laursen, Francis, Abbiss,
18 Newton and Nosaka (2007) reported coefficient of variations (CV) of 3.3 and 2.0%, for 1500
19 m and 5 km running trials on a motorised treadmill respectively and Stevens et al. (2015)
20 reported a similar CV of 1.2% during 5 km running on a non-motorised treadmill. The low
21 CV for these measurements provides the researcher with confidence that any observed
22 change in performance is attributed to the intervention and not to other extraneous variables
23 (e.g. measurement error and inter-individual variation). However, the ecological validity of
24 these performance measures are questionable, as performance tests conducted within the
25 controlled laboratory environment are artificial and may not provide a true reflection of real-
26 world outdoor events. If a performance measurement fails to adequately represent the target
27 environment, then scientific experimental outcomes may not translate into practice and may
28 lack true relevance and impact (Araújo & Davids, 2009).

29 The differences in performance between artificial (e.g. laboratory) and natural (e.g.
30 outdoor) environments has been extensively investigated. Higher running velocities have
31 been reported during field-based running at fixed blood lactate concentrations in comparison
32 to laboratory based trials (Kunduracioglu, Guner, Ulkar, & Erdogan, 2007), and higher blood
33 lactate concentrations have been reported during treadmill running compared to running on
34 synthetic surfaces (Di Michele, Di Renzo, Ammazalorso, & Merni, 2009). Others have
35 reported different energetic/metabolic costs (Jones & Doust, 1996) and biomechanical
36 differences (Ali, Caine, & Snow, 2007) between treadmill and outdoor environments. While
37 more recent laboratory investigations have attempted to replicate the outdoor environment
38 with the use of non-motorised treadmills (Stevens et al., 2015). Although authors have
39 attempted to stimulate outdoor running performance in the laboratory with specialised

40 equipment, the use of such protocols have shown differences in performance of 22% (Stevens
41 et al., 2015) and suggest that the use of actual outdoor time-trials may be a more pragmatic
42 and cheaper alternative. Furthermore, and from a psychological perspective, Terry,
43 Karageorghis, Saha, and D'Auria (2012) proposed that the lack of visual stimulation within
44 the laboratory environment may increase the tedium of the task compared to the outdoor
45 external environment, whereas McAuley, Mihalko, and Bane (1997) purported the
46 unfamiliarity and perceived threat of the laboratory environment, and/or testing equipment,
47 may negatively influence anxiety and arousal. It is possible that these factors may negatively
48 influence levels of athlete motivation, effort and perceived exertion, which may consequently
49 influence performance and the inferences that can be made from interventions using these
50 protocols.

51 Collectively, current research suggests that performance measured in the laboratory
52 may not be an adequate representation of actual performance. Some studies have therefore
53 investigated the reliability of time-trials outdoors. Hodges, Hancock, Currell, Hamilton, and
54 Bruce (2006) and O'Rourke, Obrien, Knez, and Paton (2007) measured the reliability of 1500
55 m and 5000 m running and reported CVs of 0.8 and 1.4%, respectively. These results are
56 similar, if not better, to the equivalent running time-trials performed in the laboratory (e.g.
57 Laursen et al., 2007) and are more representative of actual running performance. However,
58 the studies highlighted above did not investigate the effects of direct competition during
59 performance and, like the laboratory protocols, may lack ecological validity. The effect of
60 competition can have a significant impact on the physiology of the athlete and subsequently
61 the performance. Pierce, Kuppert, and Hardy (1976) reported that adrenaline is significantly
62 higher during competitions in comparisons to training for basketball and track and field
63 athletes, whereas Viru et al. (2010) reported differences in the peak oxygen consumption
64 ($\dot{V}O_{2peak}$) and performance between competitive and non-competitive situations in treadmill

65 running. These studies highlight the physiological differences competitive and non-
66 competitive environments can have, which can have a significant impact on the performance
67 of the athlete. This illustrates the argument that for researchers to truly elucidate the efficacy
68 of an intervention, the methods employed have to replicate the athletes' actual performance.
69 That is, moving away from laboratory based measures to assessing actual performance in the
70 field, and where possible, in a competitive environment.

71 Since 2004, weekly, open entry, free and timed 5 km road race events (parkrun®),
72 have become increasingly popular throughout the United Kingdom (UK) and offer the
73 opportunity for researchers to understand the efficacy of running interventions on a
74 heterogeneous sample in a competitive environment. However, the reliability of these events
75 has not been established. The aim of this study was therefore to assess the reliability of
76 running performance during an outdoor running event in a competitive environment in
77 trained athletes.

78 **Method**

79 Participants

80 Institutional ethical approval was obtained from the University of Sunderland. Fifteen,
81 competitive, male, endurance runners (mean \pm standard deviation [SD]; age = 29.5 ± 4.3
82 years, height = 1.75 ± 0.08 m, body mass = 71.0 ± 7.1 kg, 5 km personal best = $19:13 \pm 1:13$
83 minutes) were recruited following a call out for participants made through social media to
84 Newcastle-upon-Tyne parkrunners. Participants trained regularly (>5 d \cdot week $^{-1}$) during the 6
85 months prior to the study and regularly participated in 5 km competitive races. All
86 participants were habituated with the selected course and event, having each completed >10
87 parkruns® in Newcastle-upon-Tyne prior to the study. Written informed consent was
88 obtained from all participants prior to participation. Participants were informed that they
89 could withdraw from the study at any point in time should they wish to do so without reprisal.

90 Procedure

91 A within-participant study design was adopted. Participants completed two 5 km
92 time-trial runs (Trial 1 and Trial 2) in a competitive environment within a 7-21 day period.
93 The 5 km parkrun® trials took place in Newcastle upon-Tyne, UK. The Newcastle-upon-
94 Tyne parkrun® is run on tarmac and has been accurately measured using a professional
95 measuring wheel. The course is officially certified and is located 61-75 m above sea level.
96 Approximately 500 runners compete weekly. For this reason, participants were asked to
97 begin the trial at the front of the mass start to ensure times were not hindered by other
98 runners. Participants were asked to prepare for and treat each run as they would for a
99 competition. They were asked to maintain a similar diet for 48 hours, rest adequately (>8
100 hours of sleep) and maintain their pre-competition training routines before each trial.
101 Participants performed individual warm up routines and were asked to keep this the same for
102 subsequent trials. The 5 km runs started promptly at 09:00 and participants were instructed to

103 complete the distance as fast as possible. Environmental conditions, wind speed (m/s),
104 temperature (°C), relative humidity (%) and wind chill (°C) were recorded weekly using the
105 Pasco weather sensor (PS-2174, Pasco, Roseville CA, USA) attached to the Xplorer GLX
106 graphing data-logger (PS-2002, Pasco, Roseville CA, USA). Measures were taken at various
107 points around the course and a mean value recorded. Time trials were not recorded on days
108 when the wind speed exceeded ± 2 m/s. Weather conditions remained stable (cool and dry)
109 across all trials, with wind speed ranging from 0.9 to 1.8 m/s, temperature from 4 to 7 °C,
110 relative humidity from 82 to 92% and wind chill from 3 to 4 °C.

111 Statistical analysis

112 Data are presented as mean \pm standard deviation (SD) and 95% confidence intervals
113 (95% CI) in brackets. Normality was assessed using the Shapiro-Wilk test for normality.
114 Paired samples t-tests were conducted to determine any systematic difference in performance
115 time between the two runs (Trial 1 - Trial 2). Cohen's d was calculated to determine the
116 effect size (d) of the mean differences (Cohen, 1977) and interpreted using the modified scale
117 proposed by Hopkins (2002): trivial ≤ 0.2 ; small 0.2-0.6; moderate, 0.6-1.2; and large, >1.2 .
118 Absolute reliability of 5 km performance time was determined using the within-participant
119 coefficient of variation (CV) and typical error (TE) expressed in seconds. A CV $\leq 1.5\%$ was
120 set as a criterion for absolute reliability (Hopkins & Hewson, 2001). Within-participant CV's
121 were calculated for individual participants by dividing the standard deviation of their Trial 1
122 and Trial 2 performances by their mean performance and multiplying by 100 (SD [Trial 1 and
123 Trial 2] / mean [Trial 1 and Trial 2]*100). The mean CV is reported. Relative reliability was
124 established using the intra-class correlation coefficient (ICC). TE and ICC were calculated
125 using an online statistical spreadsheet (Hopkins, 2009). The precision of the ICC (95% CI)
126 was established using the McGraw and Wong (1996) formula. The ICC was interpreted as
127 follows: ICC <0.80 low reliability; ICC 0.80 to 0.90 moderate reliability; ICC >0.9 high

128 reliability (Vincent & Weir, 2005). Statistical analyses were conducted using Microsoft Excel
129 and SPSS for windows version 20.0 (SPSS Inc., Chicago, USA) software packages.
130 Significance was accepted at $p < .05$.

131 **Results**

132 Twenty participants completed the first initial trial. Five participants withdrew from
133 the study before completing the second trial (injury, n = 2; no reason provided, n = 2; non-
134 availability, n = 1). Data analysis is based on the 15 participants who successfully completed
135 both time-trials. No differences existed between participants who withdrew and the
136 participants who remained for the demographic variables (i.e. age, height, weight, PB; p >
137 .05).

138 Individual performances are presented in table 1. Performance times for the two 5 km
139 time-trials were highly reproducible (mean \pm standard deviation [SD]; 1217 ± 85 s, 95% CI
140 [1170, 1264] and 1216 ± 79 s, [1172 to 1260] for Trial 1 and Trial 2 respectively). The mean
141 difference in running performance between Trial 1 and Trial 2 was 1.0 ± 20.8 s [-10.5, 12.5].
142 A paired samples t-test revealed no differences between the two trials ($t_{(14)}$, p = .855, d <
143 0.01). The coefficient of variation (CV) was $0.95 \pm 0.65\%$ [0.59, 1.31], typical error (TE) =
144 14.7 s [11.3, 21.4] equating to approximately 1.2% of mean performance, and Intra-class
145 correlation = 0.97 [0.93, 0.99].

146 **Discussion**

147 The aim of this present study was to evaluate the reproducibility of an outdoor,
148 competitive time-trial. To our knowledge, this is the first study to assess the reproducibility of
149 this type of measure. Results suggest that an outdoor 5 km time-trial within a competitive
150 environment is highly reproducible in a population of trained athletes. Results have
151 implications for future research that seek to understand the effects of interventions on
152 endurance running performance. The use of a reliable competitive, outdoor time-trial could
153 provide researchers with greater confidence that results of intervention studies can be
154 extrapolated to real world environments.

155 The use of an indoor, laboratory based time-trial to distinguish the effects of an
156 intervention for endurance performance has been frequently used within sport and exercise
157 science (Stevens & Dascombe, 2015). This is commonly perceived to be a more reliable
158 method compared to outdoor time-trials (Reilly, Morris, & Whyte, 2009). However, results
159 from this study suggest that this may not be entirely accurate. The coefficient of variation
160 (CV) of $0.95 \pm 0.65\%$, 95% CI [0.59, 1.31] reported in this study is similar, if not better than
161 indoor laboratory based time-trials. Russell, Redmann, Ravussin, Hunter, and Larson-Meyer
162 (2004) reported CV of 1.0% for 10 km treadmill based time-trials and Laursen et al. (2007)
163 reported CV of 3.3 and 2.0%, for 1500 m and 5 km treadmill based time-trials, respectively.
164 In addition, the results of Stevens et al. (2015) who attempted to better simulate the outdoor
165 environment with a non-motorised treadmill, reported a similar CV of 1.2% for 5 km time-
166 trials. The use of an outdoor competitive time-trial is therefore comparable, if not more
167 reliable than an indoor, laboratory based time-trial. This holds important implications and
168 considerations for researchers aiming to establish the effectiveness of an intervention on
169 running performance. For inferences to be extrapolated to performance, the use of a protocol
170 that holds high reliability and validity should be used. If this isn't achieved, the inferences

171 reported may not be translated accurately to actual performance. We therefore encourage the
172 use of the 5 km outdoor, competitive time-trial (i.e. parkrun®) as a means of confidently
173 assessing the efficacy of running interventions.

174 The typical error (TE; 1.2% or 14.7 s) reported is also lower compared to previously
175 reported variability in non-elite distance runners and corroborates a similar TE (1.3%)
176 observed in trained endurance athletes over a 3000 m indoor time-trial (Durussel et al., 2013)
177 and a 1.4% TE for distances between 3000 m and 10000 m in elite athletes (Hopkins &
178 Hewson, 2001). The low values for TE and CV in the current study may be attributed to the
179 level of participant familiarisation with the competitive, parkrun® time-trial adopted for the
180 study. Stevens et al. (2015) emphasised that to minimise the test-retest variation, running
181 time-trials should include participants that are familiar with the testing procedures. For this
182 study, prior to their recruitment, participants had completed on average 51 ± 38 5 km
183 parkrun® time-trials (range = 14 to 144). This highlights a strength of the current study and
184 of parkrun®, as it ensures that the participants were well versed and familiar with the course
185 and distance. The parkrun® events can provide a unique advantage for future research, as it
186 allows the researcher the opportunity to access a population of athlete that are experienced
187 with the running protocol already, without having to include familiarisation trials prior to the
188 study. In short, this would save time and resources during repeated measure, cross-over
189 design studies.

190 The results of this study should take into consideration a number of potential
191 limitations. First, the use of pacing and drafting practices were not fully explored during this
192 study. We acknowledge that time-trials by design, permit athletes to alter their pace, which
193 may influence reproducibility of performance. However, we, like others (Hampson, Gibson,
194 Lambert, & Noakes, 2001; Laursen et al., 2007), argue pacing strategies are an integral
195 component of real-life, competitive performance. To maximise ecological validity, athletes in

196 this study were permitted to alter their race-pace to suit the interactions between their
197 perceptions of fatigue and external motivational cues (Hampson et al., 2001). Secondly, it
198 may be argued the ‘competitive’ environment in this study would be better described as semi-
199 competitive as it does not mimic the atmosphere, pressures and demands, arousal or anxiety
200 experienced by elite level athletes competing for podium status at international athletic
201 competitions.

202 These limitations notwithstanding, the results provide valuable information for
203 researchers wishing to ascertain the effect of interventions on outdoor endurance running
204 performance. Empirical investigations establishing ecologically valid research protocols, in a
205 competitive environment, have been difficult to design. The timing of competition events
206 often differs between venues and are rarely held at the same time of the day. The parkrun®
207 event used in this research eliminates such confounds as parkrun® events are scheduled
208 weekly at the same time of day.

209 To conclude, mean performance time was found to be highly reproducible over
210 repeated competitive, outdoor-based 5 km time-trials. Results are similar, if not better, than
211 indoor based, treadmill and outdoor, track based time-trials. Results provide a useful platform
212 from which to measure the magnitude of performance changes following future interventions.

213 **References**

- 214 Ali, A., Caine, M. P., & Snow, B. G. (2007). Graduated compression stockings: physiological
215 and perceptual responses during and after exercise. *Journal of Sports Sciences*, 25(4), 413-
216 419. doi:10.1080/02640410600718376
- 217 Araujo, D., & Davids, K. (2009). Ecological approaches to cognition and action in sport and
218 exercise: Ask not only what you do, but where you do it. *International Journal of Sport*
219 *Psychology*, 40(1), 5-37.
- 220 Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York:
221 Academic Press.
- 222 Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of
223 sporting performance. *Sports medicine*, 38(4), 297-316. doi:10.2165/00007256-
224 200838040-00003
- 225 Di Michele, R., Di Renzo, A. M., Ammazalorso, S., & Merni, F. (2009). Comparison of
226 physiological responses to an incremental running test on treadmill, natural grass, and
227 synthetic turf in young soccer players. *The Journal of Strength & Conditioning Research*,
228 23(3), 939-945. doi:10.1519/JSC.0b013e3181a07b6e
- 229 Durussel, J., Daskalaki, E., Anderson, M., Chatterji, T., Wondimu, D. H., Padmanabhan, N.,
230 ... & Pitsiladis, Y. P. (2013). Haemoglobin mass and running time trial performance after
231 recombinant human erythropoietin administration in trained men. *PloS one*, 8(2), 1-8.
232 doi:10.1371/journal.pone.0056151
- 233 Hampson, D. B., Gibson, A. S. C., Lambert, M. I., & Noakes, T. D. (2001). The influence of
234 sensory cues on the perception of exertion during exercise and central regulation of
235 exercise performance. *Sports Medicine*, 31(13), 935-952. doi:10.2165/00007256-
236 200131130-00004

237 Hodges, K., Hancock, S., Currell, K., Hamilton, B., & Jeukendrup, A. E. (2006).
238 Pseudoephedrine enhances performance in 1500-m runners. *Medicine and Science in*
239 *Sports and Exercise*, 38(2), 329-333. doi:10.1249/01.mss.0000183201.79330.9c

240 Hopkins, W. G. (2002). A scale of magnitudes for effect statistics. Retrieved from
241 <http://www.sportsci.org/resource/stats/effectmag.htm>

242 Hopkins, W. G. (2009). Calculating the reliability intraclass correlation coefficient and its
243 confidence limits (Excel spreadsheet). Retrieved from
244 <http://www.sportsci.org/resource/stats/relycalc.html>

245 Hopkins, W. G., & Hewson, D. J. (2001). Variability of competitive performance of distance
246 runners. *Medicine and Science in Sports and Exercise*, 33(9), 1588-1592.
247 doi:10.1097/00005768-200109000-00023

248 Jones, A. M., & Doust, J. H. (1996). A 1% treadmill grade most accurately reflects the
249 energetic cost of outdoor running. *Journal of Sports Sciences*, 14(4), 321-327.
250 doi:10.1080/02640419608727717

251 Kunduracioglu, B., Guner, R., Ulkar, B., & Erdogan, A. (2007). Can heart rate values
252 obtained from laboratory and field lactate tests be used interchangeably to prescribe
253 exercise intensity for soccer players? *Advances in Therapy*, 24(4), 890-902.
254 doi:10.1007/BF02849982

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

258 McAuley, E., Mihalko, S. L., & Bane, S. M. (1997). Exercise and self-esteem in middle-aged
259 adults: multidimensional relationships and physical fitness and self-efficacy
260 influences. *Journal of Behavioral Medicine*, 20(1), 67-83. doi:10.1023/A:1025591214100

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

263 O'Rourke, M. P., O'Brien, B. J., Knez, W. L., & Paton, C. D. (2008). Caffeine has a small
264 effect on 5-km running performance of well-trained and recreational runners. *Journal of*
265 *Science and Medicine in Sport*, 11(2), 231-233. doi:10.1016/j.jsams.2006.12.118

266 Pierce, D., Kupprat, I., & Harry, D. (1976). Urinary epinephrine and norepinephrine levels in
267 women athletes during training and competition. *European journal of applied physiology*
268 *and occupational physiology*, 36(1), 1-6. doi:10.1007/BF00421628

269 Reilly, T., Morris, T., & Whyte, G. (2009). The specificity of training prescription and
270 physiological assessment: A review. *Journal of Sports Sciences*, 27(6), 575-589.
271 doi:10.1080/02640410902729741

272 Russell, R. D., Redmann, S. M., Ravussin, E., Hunter, G. R., & Larson-Meyer, D. E. (2004).
273 Reproducibility of endurance performance on a treadmill using a preloaded time
274 trial. *Medicine and Science in Sports and Exercise*, 36(4), 717-724.
275 doi:10.1249/01.MSS.0000121954.95892.C8

276 Stevens, C. J., & Dascombe, B. J. (2015). The Reliability and Validity of Protocols for the
277 Assessment of Endurance Sports Performance: An updated review. *Measurement in*
278 *Physical Education and Exercise Science*, 19(4), 177-185.
279 doi:10.1080/1091367X.2015.1062381

280 Stevens, C. J., Hacene, J., Sculley, D. V., Taylor, L., Callister, R., & Dascombe, B. (2015).
281 The reliability of running performance in a 5 km time trial on a non-motorized
282 treadmill. *International Journal of Sports Medicine*, 36(9), 705-709. doi:10.1055/s-0034-
283 1398680

284 Terry, P. C., Karageorghis, C. I., Saha, A. M., & D'Auria, S. (2012). Effects of synchronous
285 music on treadmill running among elite triathletes. *Journal of Science and Medicine in*
286 *Sport*, 15(1), 52-57. doi:10.1016/j.jsams.2011.06.003

287 Vincent, W.P. & Weir, J.P. (2005). *Statistics in Kinesiology*. Leeds: Human Kinetics.

288 Viru, M., Hackney, A., Karelson, K., Janson, T., Kuus, M., & Viru, A. (2010). Competition
289 effects on physiological responses to exercise: performance, cardiorespiratory and hormonal
290 factors. *Acta Physiologica Hungarica*, 97(1), 22-30. doi:10.1556/APhysiol.97.2010.1.3

291 **Table captions**

292 Note: CV = coefficient of variation

Table 1. Individual differences for outdoor competitive 5km time-trial

Participant	Trial 1 (s)	Trial 2 (s)	Differences (s)	Percentage Change in Performance	CV (%)
1	1173	1170	3	0.26%	0.18
2	1157	1150	7	0.61%	0.43
3	1328	1289	39	3.03%	2.11
4	1363	1343	20	1.49%	1.05
5	1121	1154	-33	-2.86%	2.05
6	1180	1176	4	0.34%	0.24
7	1070	1078	-8	-0.74%	0.53
8	1127	1111	16	1.44%	1.01
9	1152	1156	-4	-0.35%	0.25
10	1212	1227	-15	-1.22%	0.87
11	1286	1254	32	2.55%	1.78
12	1261	1250	11	0.88%	0.62
13	1282	1296	-14	-1.08%	0.77
14	1271	1301	-30	-2.31%	1.65
15	1270	1283	-13	-1.01%	0.72
Mean ± SD	1216.9 ± 84.7	1215.9 ± 79.2	-1.0 ± 20.8	0.08 ± 0.02	0.95 ± 0.65