



**Manchester
Metropolitan
University**

Sawczuk, T and Jones, B and Scantlebury, S and Read, D and Costello, N and Darrall-Jones, J and Stokes, K and Till, K and Weakley, J (2017) Between-Day Reliability and Usefulness of a Fitness Testing Battery in Youth Sport Athletes: Reference Data for Practitioners. *Measurement in Physical Education and Exercise Science*, 22. pp. 11-18. ISSN 1091-367X

Downloaded from: <https://e-space.mmu.ac.uk/625780/>

Version: Accepted Version

Publisher: Taylor & Francis

DOI: <https://doi.org/10.1080/1091367X.2017.1360304>

Please cite the published version

<https://e-space.mmu.ac.uk>

1 **Between-day reliability and usefulness of a fitness testing battery in youth sport**
2 **athletes: Reference data for practitioners**

3 Running Head: Reliability of a fitness testing battery in youth sport athletes

4
5 Thomas Sawczuk^{1,2}, Ben Jones^{1,2,3,4,5}, Sean Scantlebury^{1,2}, Jonathon Weakley^{1,2,3}, Dale
6 Read^{1,3}, Nessian Costello^{1,2,4}, Joshua David Darrall-Jones^{1,3}, Keith Stokes⁶ and Kevin Till^{1,3,4}

7
8 ¹ Institute for Sport, Physical Activity and Leisure, Leeds Beckett University, Leeds, United
9 Kingdom

10 ² Queen Ethelburga's Collegiate, Thorpe Underwood, York, United Kingdom

11 ³ Yorkshire Carnegie Rugby Club, Headingley Carnegie Stadium, Leeds, United Kingdom

12 ⁴ Leeds Rhinos Rugby Club, Headingley Carnegie Stadium, Leeds, United Kingdom

13 ⁵ The Rugby Football League, Red Hall, Leeds, United Kingdom

14 ⁶ University of Bath, Claverton Down, Bath, United Kingdom

15
16 Corresponding Author:

17 Thomas Sawczuk

18 Room G03, Macaulay Hall

19 Institute for Sport, Physical Activity and Leisure

20 Centre for Sports Performance

21 Leeds Beckett University, Headingley Campus

22 West Yorkshire

23 LS6 3QS

24 Phone: (0044) 7530945555

25 Email: t.sawczuk@leedsbeckett.ac.uk

26 Abstract

27 This study aimed to evaluate the between-day reliability and usefulness of a fitness testing
28 battery in a group of youth sport athletes. Fifty-nine youth sport athletes (age = 17.3 ± 0.7
29 years) undertook a fitness testing battery including the isometric mid-thigh pull (IMTP),
30 countermovement jump (CMJ), 5-40 m sprint splits, and the 5-0-5 change of direction test on
31 two occasions separated by 7 days. Usefulness was assessed by comparing the reliability
32 (typical error; TE) to the smallest worthwhile change (SWC). The TE was 5.5% for IMTP
33 and 3.8% for CMJ. The TE values were 2.7%, 2.5%, 2.2%, 2.2% and 1.8% for the 5, 10, 20,
34 30 and 40 m sprint splits, and 4.1% (left) and 5.4% (right) for the 5-0-5 tests. SWC ranged
35 from 1.1% to 6.1%. All tests were identified as having "good" or "acceptable" reliability. The
36 IMTP and CMJ had "good" usefulness, all other tests had "marginal" usefulness.

37

38 **Key words:** Reliability, usefulness, fitness testing, strength, power, speed

39 **Introduction**

40 The importance of developing strength, power, speed and change of direction
41 qualities to improve athletic performance and reduce injury risk in adolescent athletes has
42 been highlighted regularly in recent years (Lloyd et al., 2016; Reilly, Williams, Nevill, &
43 Franks, 2000; Young, 2006). This, in addition to the recent National Strength and
44 Conditioning Association (NSCA) position statements indicating that resistance training is
45 beneficial for the youth athlete, has resulted in an increase in the number of adolescents
46 participating in structured strength and conditioning programmes (Faigenbaum et al., 2009;
47 Lloyd et al., 2016). A number of these programmes take place at schools or colleges where
48 adolescents are given scholarships based on their sporting prowess. A recent report by Ofsted
49 has suggested that as many as 15% of current international athletes across a range of sports
50 received a sports scholarship at some point during their school life, highlighting their
51 importance (Ofsted, 2014). As part of the scholarships, it is common for coaches to use
52 fitness testing batteries to regularly measure and monitor the physical characteristics of their
53 athletes (Pyne, Spencer, & Mujika, 2014). However, little information is available regarding
54 the between-day reliability of these tests, particularly in a school sport environment where
55 athletes of different indoor and outdoor sports regularly train and test their physical
56 capabilities together.

57

58 The between-day reliability of a test refers to its ability to produce consistent results
59 from day to day (Hopkins, 2000). In order for coaches to be confident that changes in
60 performance from a specific test are "real" and not due to the daily variation in the test, it is
61 important that the test has good between-day reliability. Although good reliability of a test is
62 necessary, in order for its results to be interpretable it is also important that it is sensitive
63 enough to detect the smallest worthwhile change (SWC) in performance. This has been

64 termed its "usefulness" (Hopkins, 2004) and is assessed by comparing a test's between-day
65 reliability, or typical error (TE), to the SWC. To do this, the TE is usually converted into a
66 factor of the SWC, which can be termed the "TE:SWC ratio". If the SWC is greater than a
67 test's between-day reliability (i.e. TE:SWC ratio < 1) it is considered to have good usefulness.
68 Conversely, if the SWC is smaller than its between-day reliability, (i.e. TE:SWC ratio > 1),
69 its usefulness is said to be "marginal" (Hopkins, 2004). This information can be used to
70 assess the length of time which may be required between tests in order for a clear change in
71 performance to be seen. A number of tests of strength, power, speed and change of direction
72 ability have had their between-day reliability and usefulness considered in recent times
73 (Cormack, Newton, McGulgan, & Doyle, 2008; Darrall-Jones, Jones, Roe, & Till, 2015; De
74 Witt et al., 2016; Gabbett, Kelly, & Sheppard, 2008; Haff, Ruben, Lider, Twine, & Cormie,
75 2015; Roe et al., 2016; Stewart, Turner, & Miller, 2014), however the reliability has been
76 shown to vary between sports and cohorts, so a study considering the between-day reliability
77 and usefulness of these tests in a group of school based adolescent athletes across multiple
78 sports is warranted.

79

80 A recently conceptualised and novel single measure of strength is the isometric mid-
81 thigh pull (IMTP). The IMTP is designed to mimic the second pull phase of the snatch and
82 clean (Haff et al., 2005), and has shown to be strongly correlated with weightlifting
83 performance (Beckham et al., 2013). It requires little technical expertise indicating it is
84 suitable for athletes of all training ages with little familiarisation (Beckham, 2015). To date,
85 the majority of IMTP reliability studies have shown force plates to be reliable measures, with
86 intraclass correlation coefficients ranging from .92-.99 (De Witt et al., 2016; Haff et al.,
87 2015), however given the high cost of the equipment, force plates are likely only available
88 within universities or professional sports teams and not within a school environment. A lower

89 cost alternative is to use a back dynamometer, the reliability of which (test-retest correlation,
90 $r = .91$) has only been shown in older population groups (Coldwells, Atkinson, & Reilly,
91 1994). However, this study did not provide the typical error (TE) as a coefficient of variation
92 (CV), limiting its practical use (Hopkins, 2000), so a study considering the between-day
93 reliability and usefulness of the IMTP using a back dynamometer in youth sport athletes is
94 warranted.

95
96 The countermovement jump (CMJ) has received considerable attention in recent years
97 as a measure of neuromuscular power. Although the majority of this attention has come with
98 regards to its use as a daily monitoring tool for neuromuscular power (McLean, Coutts,
99 Kelly, McGuigan, & Cormack, 2010; Roe et al., 2016), CMJ height can still be used as a
100 surrogate measure of lower body power similar to the vertical jump used in the NFL
101 Combine (McGee & Burkett, 2003). As with the IMTP, the reliability of the CMJ has been
102 shown extensively when using a force plate, with TE's ranging from 5.2-6.8% for jump
103 height and 2.9-3.6% for flight time (Cormack et al., 2008; McLean et al., 2010; Roe et al.,
104 2016). However, the less expensive Optojump system for measuring CMJ height has only
105 once had its reliability confirmed (jump height TE = 2.2%) and this took place in a group of
106 older, less well trained individuals (Glatthorn et al., 2011). It is therefore important to assess
107 the between-day reliability and usefulness of the Optojump system as a measure of CMJ
108 height in youth sport athletes.

109
110 Timing gates are frequently used to measure the linear sprint ability of athletes
111 (Darrall-Jones et al., 2015; Duthie, Pyne, Ross, Livingstone, & Hooper, 2006; Young,
112 McLean, & Ardagna, 1995). Within a cohort of rugby players the reliability of timing gates
113 as a measure of linear sprint ability (TE = 1.3-3.1%) has been proven (Darrall-Jones et al.,

114 2015), however the movement demands of rugby are different to those of netball (Chandler,
115 Pinder, Curran, & Gabbett, 2014; Read et al., 2017), for example, thus differences in
116 reliability may exist between sports. Furthermore, Darrall-Jones and colleagues (2015) found
117 timing gates to have "marginal" usefulness, limiting the ability of the test to detect small
118 changes in performance. In addition to possessing linear speed, it is important that athletes
119 are able to move in a multi-directional manner (Sheppard & Young, 2006). As such, a test to
120 determine the change of direction ability of youth athletes, for example the 5-0-5 test (Draper
121 & Lancaster, 1985), should exist within a fitness testing battery. Previous studies have shown
122 the 5-0-5 test measured using timing gates to be reliable in adult rugby league players (TE =
123 1.9%; Gabbett et al., 2008) and school PE students measured indoors (TE = 2.8%; Stewart et
124 al., 2012). However, the 5-0-5 test is regularly assessed outdoors (Darrall-Jones, Jones, &
125 Till, 2015; Gabbett et al., 2008) and differences in the between-day reliability of a test may
126 be present in the same cohort in different conditions (i.e. indoor vs outdoor testing). Given
127 the common use of timing gates to measure sprinting ability, a study involving numerous
128 different sports, including both males and females, is required to establish the reliability and
129 usefulness of timing gates for measuring speed in a multi-sport setting. Furthermore,
130 establishing the reliability of the 5-0-5 test on an outdoor surface will provide useful
131 information for coaches working with multi-sport youth athletes and will allow the usefulness
132 of the test to be determined.

133

134 Despite a large number of sports scholarships across the globe, where general fitness
135 testing protocols may be put in place, the majority of research considering the between day
136 reliability of fitness testing protocols has taken place in individual sports (Cormack et al.,
137 2008; Darrall-Jones et al., 2015; Gabbett et al., 2008). Consequently, the aim of this study
138 was to assess the between-day reliability and usefulness of a fitness testing battery

139 incorporating measures of strength, power, speed and change of direction ability in
140 scholarship youth sport athletes.

141

142 **Methods**

143 *Participants*

144 Fifty-nine youth sport athletes (39 males, 20 females, age 17.3 ± 0.7 years, height $175.0 \pm$
145 17.4 cm, body mass 75.5 ± 14.0 kg) were recruited for this study from a local independent
146 school in the United Kingdom. The athletes were part of the school's sport scholarship
147 programme and had all previously competed at professional academy level or above, but
148 were now club/school ($n = 34$), professional academy ($n = 7$), county/regional ($n = 14$) or
149 international ($n = 4$) standard in their respective sports. The sports represented were
150 basketball ($n = 3$), cricket ($n = 5$), football ($n = 10$), hockey ($n = 9$), netball ($n = 10$) and
151 rugby ($n = 22$). Ethics approval was granted by the University Ethics Committee and written
152 informed consent was provided by all participants and their parents prior to the study.

153

154 *Research Design*

155 In order to assess the between day reliability of this fitness testing battery, participants
156 completed the tests on two separate occasions over a two-week period. Participants refrained
157 from strenuous exercise in the 24 hours prior to each testing day, and training volume was
158 standardised for the duration of the study, so that participants completed the same number of
159 sessions in both weeks, in line with previous studies (Duthie et al., 2006; Stewart et al.,
160 2012). These sessions consisted of strength and conditioning and technical training sessions,
161 both of which were controlled for intensity between weeks. On days one and seven, subjects
162 performed measures of strength via the IMTP and power via the CMJ. On days four and ten,
163 field based measures of 40 m sprints to measure speed and the 5-0-5 test to measure change

164 of direction ability were performed. On all testing days, the test inducing the greatest strain
165 on the neuromuscular system was performed first in order to enhance the reliability of all
166 maximal testing procedures (Harman, 2008). On days one and seven, this meant the CMJ was
167 performed first; on days four and ten, the 5-0-5 change of direction test was performed first.
168 Participants completed field based measures on either a 4G outdoor artificial grass playing
169 surface (cricket, football and rugby) or an outdoor running track (basketball, hockey and
170 netball), dependent on their sport. Ambient conditions were measured using a weather station
171 (Davis Instruments Corporation, Hayward, USA). These are shown in Table 1. A
172 standardised dynamic warm up including leg swings, lunges, squats was performed prior to
173 each testing session. Participants had been familiarised to all tests earlier in the academic
174 year.

175

176 ***INSERT TABLE 1 HERE***

177

178 *Protocols*

179 The IMTP was performed using a modified back dynamometer (Takei Scientific
180 Instruments Co., Niigata City, Japan). The modification increased the size of the base so that
181 participants could stand with their feet shoulder width apart rather than hip width as
182 necessitated by the original model. Participants were instructed to stand with their feet
183 shoulder width apart and knees bent at 120-135° in line with previous studies (Beckham et
184 al., 2013; Darrall-Jones et al., 2015). The bar was adjusted so that when they held it taut with
185 a straight back and arms, it reached the middle of their thigh. Participants were instructed to
186 pull directly upwards, keeping their feet flat on the platform and without leaning back. Two
187 warm up pulls were performed at 50% and 75% of maximum, before three all out efforts
188 were executed, each separated by a 3-minute rest.

189

190 The CMJ was performed using the Optojump system (Microgate, Bolzano, Italy).
191 Jump height was reported in centimetres. Participants began with their legs fully extended
192 and their feet at a self-selected width, with their hands on their hips. They were then
193 instructed to squat down and jump as high as they could in a fluid, countermovement motion.
194 The depth of the countermovement was self-selected. Participants were instructed to keep
195 their legs extended in flight and to land with their legs straight. Two warm up jumps were
196 completed, before three maximal efforts were executed with a 3-minute rest provided in
197 between each repetition.

198

199 Linear sprint speed was assessed via 40 m sprint with split times taken at 5, 10, 20,
200 30 and 40 m using single beam timing gates (Brower Timing Systems, IR Emit, USA). The
201 height of the timing gates was standardised at 1m in line with previous guidelines (Cronin &
202 Templeton, 2008). The sprinting direction was standardised as north-west for all sessions.
203 Participants were instructed to start 0.5 m behind the first timing gate and to start their sprints
204 at a self selected time. Three maximal efforts were performed, each separated by 3-minutes
205 rest. As part of their warm up, participants completed one practice sprint.

206

207 The 5-0-5 change of direction test was also measured using single beam timing gates
208 (Brower Timing Systems, IR Emit, USA) and was performed after the 40 m sprints,
209 following 5 minutes of rest. Participants began the test at a self selected time, sprinting 10 m
210 in a south-easterly direction (i.e. opposite to the sprints) before planting their foot beyond a
211 white line, turning 180° and sprinting back 5 m. The timing gates measured the time from 5
212 m before the line until they sprinted back through that point. Three maximal repetitions were

213 completed with 3-minutes rest in between each effort. One practice test was completed as
214 part of the warm up procedure.

215

216 *Data Analysis*

217 For all tests, the best of three efforts was taken for the between day reliability
218 analysis, in line with previous studies (Darrall-Jones et al., 2015; Gabbett et al., 2008). Data
219 were log transformed to allow the TE to be calculated as a CV (%), along with the SWC
220 using a premade Microsoft Excel spreadsheet (Hopkins, 2015). The TE is calculated as
221 follows:

$$222 \quad \text{TE} = S_{\text{diff}}/\sqrt{2}$$

223 where S_{diff} is the standard deviation of the difference score between two trials (Hopkins,
224 2000). Back transformation of the log transformed data provided the TE as a percentage
225 relative to the mean. Similar to previous studies, a CV of 5% or less was used to categorise a
226 reliable variable (Darrall-Jones et al., 2015; Roe et al., 2016). The SWC was calculated as 0.2
227 x between-subject standard deviation, in line with the Cohen's d effect size principle, and
228 expressed as a percentage of the mean in order to compare with the CV. The usefulness of the
229 test was classified according to the Hopkins (2000) criteria: Good (CV < SWC; TE:SWC
230 ratio < 1), OK (CV = SWC; TE:SWC ratio = 1) or Marginal (CV > SWC; TE:SWC ratio >
231 1).

232

233 **Results**

234 Table 2 shows the TE as a CV, SWC, TE:SWC ratio and usefulness of the tests. All
235 sprint splits had less than 5% CV and so were considered reliable tests. The CMJ also showed
236 good between day reliability. The IMTP and the 5-0-5 tests showed acceptable between day

237 reliability. All timing gate measured tests' usefulness was considered "marginal", whereas the
238 IMTP and CMJ tests were considered to have "good" usefulness.

239

240 ***INSERT TABLE 2 HERE***

241

242 **Discussion**

243 The aim of this study was to assess the between-day reliability and usefulness of a
244 fitness testing battery in a group of youth sport athletes of varying standards. The main
245 finding is that all fitness testing protocols were shown to have good (< 5% TE) or acceptable
246 (~5% TE) between-day reliability. A further finding of the study is that the strength and
247 power tests showed "good" usefulness, whereas speed and change of direction tests using the
248 timing gates showed "marginal" usefulness. As a consequence, the IMTP and CMJ tests are
249 able to detect smaller changes in performance with greater certainty than the timing gate
250 protocols.

251

252 A number of studies have considered the reliability of the IMTP using a force plate
253 (De Witt et al., 2016; Haff et al., 2015), however to these authors' knowledge the only study
254 to have considered a back dynamometer did not provide the TE of the test, and therefore did
255 not provide practically useful results (Coldwells et al., 1994). Our results confirm the
256 between-day reliability of the use of a back dynamometer to measure the IMTP strength of a
257 youth sport athlete. The TE of 5.4% is slightly greater than, but still comparable with, the 1.7
258 and 3.1% values previously reported using force plates (Haff et al., 2015; James, Roberts,
259 Haff, Kelly, & Beckman, in press). The difference in these figures is likely due to previously
260 reported values using more expensive devices with greater sampling frequencies (e.g. force
261 plates) and the different cohorts used. Furthermore, although it has previously been indicated

262 that little familiarisation is required to perform the test (Beckham, 2015), the ability to
263 consistently produce strength and power from this position is likely to require familiarisation
264 to and training in Olympic weightlifting techniques. The 1.4% difference in the previously
265 reported values (Haff et al., 2015; James et al., in press) provides support for this theory as
266 the study by Haff and colleagues, which used participants who regularly performed
267 "resistance training, including [Olympic] weightlifting movements" (Haff et al., 2015),
268 showed greater reliability than the "recreationally active with ≥ 6 months resistance training
269 exercise" participants used in James and colleagues' study (James et al., in press).

270

271 Our results demonstrate the reliability of the Optojump system for measuring power
272 via CMJ height in adolescent youth sport athletes. The 2.8% TE here is similar to the 2.2%
273 TE shown in an older, less trained age group by Glatthorn and colleagues (2011). Unlike the
274 back dynamometer, this is much lower than those values reported in force plate studies
275 (Cormack et al., 2008; Roe et al., 2016). It is suggested that this is because force plates
276 calculate jump height differently to the Optojump system. Force plates calculate jump height
277 using the velocity of the centre of mass at take off, whereas the Optojump system calculates
278 flight time using the breaking of an infrared beam. No equation was provided by either
279 previous study for their calculation of jump height from a force plate (Cormack et al., 2008;
280 Roe et al., 2016), however both provided between-day reliability values for the flight time of
281 the jump and our 2.8% between-day reliability value is comparable with the 3.3% (Cormack
282 et al., 2008) and 2.6% (Roe et al., 2016) measurements previously quoted.

283

284 The results of the sprints show all splits to have good between-day reliability.
285 Previous studies have either shown all sprint splits to be reliable (Darrall-Jones et al., 2015)
286 or indicated that the shorter distances of 5 or 10 m are more unreliable than longer splits

287 (Earp & Newton, 2012). Our results are similar to those of Darrall-Jones and colleagues
288 (2015) in that all splits were reliable, but also follow their trend of better reliability with
289 increased sprint split distance. The difference in reliability between shorter (e.g. 5 m) and
290 longer (e.g. 30 m) splits, however, was not sufficient for the shorter distances to provide a
291 value greater than the 5% threshold set for good between-day reliability.

292

293 The results of the 5-0-5 change of direction test showed acceptable levels of
294 reliability. Our values of 4.1% and 5.4% for left and right foot respectively are greater than
295 the previously reported values of 2.8% in adult rugby league players (Stewart et al., 2012)
296 and 1.9% in school physical education students, measured indoors (Gabbett et al., 2008). This
297 variation in the studies' results may be due to differences in the techniques used to measure
298 speed (e.g. single beam electronic timing gates vs the dual beam electronic timing gates used
299 by Gabbett and colleagues (2008)) and the differences between testing outdoors or on an
300 indoor running surface, as used by Stewart and colleagues (2012). Although there is little
301 difference in the ambient conditions shown in Table 1, the slight reduction in reliability is
302 likely due to the inherent variability in weather and ground conditions for outdoor field
303 testing, and thus should be a consideration for those involved in testing young athletes. This
304 is also shown by the difference in reliability between the two previous studies (Gabbett et al.,
305 2008; Stewart et al., 2012). The between-day difference in sprint time was lowest in the study
306 of Stewart et al. (2014), where testing was undertaken indoors, despite the more precise
307 timing gates used by Gabbett and colleagues (2008), where testing was undertaken outdoors.

308

309 With regards to the usefulness of the tests, the IMTP and CMJ tests were identified as
310 useful tests (TE:SWC ratio < 1). However, despite showing good reliability, the sprint splits
311 and 5-0-5 change of direction test were identified as having "marginal" usefulness (TE:SWC

312 ratio > 1). In order for a practitioner to assess whether a "real" change in performance has
313 occurred, it is important to plot the change in performance \pm TE against the SWC. If the TE
314 remains outside the SWC, it can be postulated that a change has occurred with 75%
315 probability, however if the TE crosses the SWC, the changes are deemed unclear (Hopkins,
316 2000). Figure 1 shows a practical example of this, plotting a sample change in performance
317 for the 10 and 20 m sprint splits of a player, using the SWC and TE values reported in this
318 paper. It is due to this method of interpreting results that the usefulness of a test is important.
319 In order for the change in performance to be considered "real" with 75% confidence, it must
320 be greater than or equal to the SWC + TE. In the case of the IMTP and CMJ tests, this means
321 that a change of 1.9 and 1.7 times the SWC would be required respectively; however for the
322 5-0-5 test this can rise to as much as four times the SWC. As a consequence, a much greater
323 improvement in performance is needed for practitioners to be confident that a "real" change
324 has occurred. Given an average 5-0-5 performance of 2.5 s and an SWC of 1.8% (0.05 s), a
325 change in performance of 0.20 s would be required to be sure with 75% probability that a
326 "real" change has occurred. This may be an unrealistic expectation if the test is to be used on
327 a regular (e.g. monthly/quarterly) basis, rather than a longer-term basis (e.g. bi-
328 annually/annually). In this situation, practitioners are advised to use the test less regularly
329 (e.g. bi-annually/annually) so that performance changes are greater and therefore more
330 certain, or accept that there will be a large element of uncertainty in results if testing occurs
331 regularly and performance changes are smaller.

332

333 *****INSERT FIGURE 1 HERE*****

334

335 Although this study has shown the reliability of a fitness testing battery in male and
336 female youth sport athletes with uneven representations from a number of sports, it is

337 important to understand the limitations inherent in the testing protocols. It could be argued
338 that not splitting the tests by sex is a limitation of the study, however in the only study to date
339 to compare the reliability of a fitness test between sex, no significant differences in reliability
340 were found (Augustsson et al., 2009). The multi-sport nature of the study could be seen as a
341 limitation. It has previously been noted that the movement demands of the sports are different
342 (Chandler et al., 2014; Read et al., 2017), however this study shows that within a multi-sport
343 environment, this fitness testing battery remains reliable. Future research may wish to assess
344 whether there are differences in reliability between sports, however given previous research
345 that shows no differences in reliability between sex (Augustsson et al., 2009), it is possible
346 that the differences between sports will be negligible. The use of two different surfaces for
347 the sprint based tests (outdoor athletics track and 4G artificial grass) could also be seen as a
348 limitation, however this could happen in practice if teams were to perform sprint testing on
349 their normal playing surface. Furthermore, as the surface remained constant between days, it
350 should not have had an impact on the results seen.

351

352 Finally, in order for this testing battery to be considered complete, a marker of aerobic fitness
353 should also be included. Along with its performance related benefits, aerobic fitness has been
354 associated with reduced injury risk in adolescents (Brenner, 2007; Difiori et al., 2014). It is a
355 limitation of this study that a test of aerobic fitness was not included, however both the Yo-
356 Yo Intermittent Recovery Test (Bangsbo, 1994; Bangsbo, Iaia, & Krstrup, 2008) and the 30-
357 15 Intermittent Fitness Test (Buchheit & Rabbani, 2014) have had their reliability and
358 sensitivity to training confirmed in similar populations and conditions to this study (Buchheit
359 & Rabbani, 2014; Deprez et al., 2014). Consequently, the authors decided not to include
360 either of these tests in this study.

361

362 In conclusion, this study has shown the reliability and usefulness of a fitness testing
363 battery aimed at monitoring strength, power and speed qualities in youth sport athletes. The
364 IMTP and CMJ were shown to be both reliable and have good usefulness. The sprint splits
365 and 5-0-5 test were shown to be reliable but had marginal usefulness. To this end, the IMTP
366 and CMJ are able to detect the smaller changes in performance with greater certainty than the
367 sprint splits and 5-0-5 test. It is recommended that either the Yo-Yo Intermittent Recovery
368 Test Level 1 or 30-15 Intermittent Fitness Test is added to ensure the battery provides a
369 complete understanding of the athlete's physical capabilities.
370

371 **References**

- 372 Augustsson, S. R., Bersas, E., Thomas, E. M., Sahlberg, M., Augustsson, J., & Svantesson,
373 U. (2009). Gender differences and reliability of selected physical performance tests in
374 young women and men. *Advances in Physiotherapy, 11*, 64–70.
- 375 Bangsbo, J. (1994). *Fitness training in football: A scientific approach*. Copenhagen
376 University: August Krogh Institute.
- 377 Bangsbo, J., Iaia, F. M., & Krstrup, P. (2008). The Yo-Yo Intermittent Recovery Test: A
378 useful tool for evaluation of physical performance in intermittent sports. *Sports*
379 *Medicine, 38*, 37–51. <http://doi.org/10.2165/00007256-200838010-00004>
- 380 Beckham, G. K. (2015). *The effect of various body positions on performance of the isometric*
381 *mid-thigh pull*.(PhD Thesis). East Tennessee State University. Retrieved from
382 <http://dc.etsu.edu/etd/2544>
- 383 Beckham, G., Mizuguchi, S., Carter, C., Sato, K., Ramsey, M., Lamont, H. S., ... Stone, M.
384 H. (2013). Relationships of isometric mid-thigh pull variables to weightlifting
385 performance. *Journal of Sports Medicine and Physical Fitness, 53*, 573–581.
- 386 Brenner, J. S. (2007). Overuse injuries, overtraining, and burnout in child and adolescent
387 athletes. *Pediatrics, 119*, 1242–1245. <http://doi.org/10.1542/peds.2007-0887>
- 388 Buchheit, M., & Rabbani, A. (2014). The 30-15 Intermittent Fitness Test versus the Yo-Yo
389 Intermittent Recovery Test Level 1: Relationship and sensitivity to training.
390 *International Journal of Sports Physiology and Performance, 9*, 522–524.
391 <http://doi.org/10.1111/j.1399-0012.2012.01641.x>
- 392 Chandler, P. T., Pinder, S. J., Curran, J. D., & Gabbett, T. J. (2014). Physical demands of
393 training and competition in collegiate netball players. *Journal of Strength and*
394 *Conditioning Research, 28*, 2732–2737. [http://doi.org/10.1519/1533-4287\(2005\)19<202](http://doi.org/10.1519/1533-4287(2005)19<202)
- 395 Coldwells, A., Atkinson, G., & Reilly, T. (1994). Sources of variation in back and leg

- 396 dynamometry. *Ergonomics*, 37, 79–86.
- 397 Cormack, S. J., Newton, R. U., McGulgan, M. R., & Doyle, T. L. A. (2008). Reliability of
398 measures obtained during single and repeated countermovement jumps. *International*
399 *Journal of Sports Physiology and Performance*, 3, 131–144.
- 400 Cronin, J. B., & Templeton, R. L. (2008). Timing light height affects sprint times. *Journal of*
401 *Strength and Conditioning Research*, 22, 318–20.
402 <http://doi.org/10.1519/JSC.0b013e31815fa3d3>
- 403 Darrall-Jones, J. D., Jones, B., Roe, G., & Till, K. (2015). Reliability and usefulness of linear
404 sprint testing in adolescent rugby union and league players. *Journal of Strength and*
405 *Conditioning Research*, 30, 1359–1364. <http://doi.org/10.1519/JSC.0000000000001233>
- 406 Darrall-Jones, J. D., Jones, B., & Till, K. (2015). Anthropometric and physical profiles of
407 English academy rugby union players. *Journal of Strength and Conditioning Research*,
408 29, 2086–2096.
- 409 De Witt, J. K., English, K. L., Crowell, J. B., Kalogera, K. L., Guilliams, M. E., Nieschwitz,
410 B. E., ... Ploutz-Snyder, L. L. (2016). Isometric mid-thigh pull reliability and
411 relationship to deadlift 1RM. *Journal of Strength and Conditioning Research*, epub
412 ahead of print. <http://doi.org/10.1519/JSC.0000000000001605>
- 413 Deprez, D., Coutts, A. J., Lenoir, M., Fransen, J., Pion, J., Philippaerts, R., & Vaeyens, R.
414 (2014). Reliability and validity of the Yo-Yo intermittent recovery test level 1 in young
415 soccer players. *Journal of Sports Sciences*, 32, 903–910.
416 <http://doi.org/10.1080/02640414.2013.876088>
- 417 Difiori, J. P., Benjamin, H. J., Brenner, J., Gregory, A., Jayanthi, N., Landry, G. L., & Luke,
418 A. (2014). Overuse injuries and burnout in youth sports: A position statement from the
419 American Medical Society for Sports Medicine. *Clinical Journal of Sport Medicine*, 24,
420 3–20. <http://doi.org/10.1097/JSM.0000000000000060>

- 421 Draper, J. A., & Lancaster, M. G. (1985). The 505 test: A test for agility in the horizontal
422 plane. *Australian Journal of Science and Medicine in Sport*, *17*, 15–18.
- 423 Duthie, G. M., Pyne, D. B., Ross, A. A., Livingstone, S. G., & Hooper, S. L. (2006). The
424 reliability of ten-meter sprint time using different starting techniques. *Journal of*
425 *Strength and Conditioning Research*, *20*, 246–251. <http://doi.org/10.1519/R-17084.1>
- 426 Earp, J. E., & Newton, R. U. (2012). Advances in electronic timing systems: Considerations
427 for selecting an appropriate timing system. *Journal of Strength and Conditioning*
428 *Research*, *26*, 1245–1248.
- 429 Faigenbaum, A. D., Kraemer, W. J., Blimkie, C. J. R., Jeffreys, I., Micheli, L. J., Nitka, M.,
430 & Rowland, T. W. (2009). Youth resistance training: Updated position statement paper
431 from the National Strength and Conditioning Association. *Journal of Strength and*
432 *Conditioning Research*, *23*, S60–S79. <http://doi.org/10.1519/JSC.0b013e31819df407>
- 433 Gabbett, T. J., Kelly, J. N., & Sheppard, J. M. (2008). Speed, change of direction speed, and
434 reactive agility of rugby league players. *Journal of Strength and Conditioning Research*,
435 *22*, 174–181. <http://doi.org/10.1519/JSC.0b013e31815ef700>
- 436 Glatthorn, J. F., Gouge, S., Nussbaumer, S., Stauffacher, S., Impellizzeri, F. M., &
437 Maffiuletti, N. A. (2011). Validity and reliability of Optojump photoelectric cells for
438 estimating vertical jump height. *Journal of Strength and Conditioning Research*, *25*,
439 556–560.
- 440 Haff, G. G., Carlock, J. M., Hartman, M. J., Kilgore, J. L., Kawamori, N., Jackson, J. R., ...
441 Stone, M. H. (2005). Force-time curve characteristics of dynamic and isometric muscle
442 actions of elite women Olympic weightlifters. *Journal of Strength and Conditioning*
443 *Research*, *19*, 741–748.
- 444 Haff, G. G., Ruben, R. P., Lider, J., Twine, C., & Cormie, P. (2015). A comparison of
445 methods for determining the rate of force development during isometric midthigh clean

- 446 pull. *Journal of Strength and Conditioning Research*, 29, 386–395.
- 447 Harman, E. (2008). Principles of test selection and administration. In T. R. Baechle & R. W.
448 Earle (Eds.), *Essentials of Strength and Conditioning* (3rd ed., pp. 238–246).
449 Champaign, IL: Human Kinetics.
- 450 Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports*
451 *Medicine*, 30, 1–15. <http://doi.org/10.2165/00007256-200030050-00006>
- 452 Hopkins, W. G. (2004). How to interpret changes in an athletic performance test.
453 *Sportscience*, 8, 1–7.
- 454 Hopkins, W. G. (2015). Spreadsheets for analysis of validity and reliability. *Sportscience*, 19,
455 36–42.
- 456 James, L. P., Roberts, L. A., Haff, G. G., Kelly, V. G., & Beckman, E. M. (2017). The
457 validity and reliability of a portable isometric mid-thigh clean pull. *Journal of Strength*
458 *and Conditioning Research*, 31, 1378–1386.
459 <http://doi.org/10.1519/JSC.0000000000001201>
- 460 Lloyd, R. S., Cronin, J. B., Faigenbaum, A. D., Haff, G. G., Howard, R., Kraemer, W. J., ...
461 Oliver, J. L. (2016). National Strength and Conditioning Association Position Statement
462 on Long Term Athletic Development. *Journal of Strength and Conditioning Research*,
463 30, 1491–1509.
- 464 McGee, K. J., & Burkett, L. N. (2003). The National Football League Combine: A reliable
465 predictor of draft status? *Journal of Strength and Conditioning Research*, 17(1), 6–11.
466 [http://doi.org/10.1519/1533-4287\(2003\)017<0006:TNFLCA>2.0.CO;2](http://doi.org/10.1519/1533-4287(2003)017<0006:TNFLCA>2.0.CO;2)
- 467 McLean, B. D., Coutts, A. J., Kelly, V., McGuigan, M. R., & Cormack, S. J. (2010).
468 Neuromuscular, endocrine, and perceptual fatigue responses during different length
469 between-match microcycles in professional rugby league players. *International Journal*
470 *of Sports Physiology and Performance*, 5, 367–383.

- 471 Ofsted. (2014). *Competitive school sport - Summary report*. Loughborough, United
472 Kingdom. Retrieved from
473 [https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/379986/N](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/379986/National_Governing_Bodies_of_Sport_survey_2014/Competitive_school_sport.pdf)
474 [ational_20Governing_20Bodies_20of_20Sport_20survey_2C_20Competitive_20school](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/379986/National_Governing_Bodies_of_Sport_survey_2014/Competitive_school_sport.pdf)
475 [_20sport.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/379986/National_Governing_Bodies_of_Sport_survey_2014/Competitive_school_sport.pdf)
- 476 Pyne, D. B., Spencer, M., & Mujika, I. (2014). Improving the value of fitness testing for
477 football. *International Journal of Sports Physiology and Performance*, *9*, 511–514.
- 478 Read, D., Jones, B., Phibbs, P., Roe, G., Darrall-Jones, J. D., Weakley, J., & Till, K. (2017).
479 Physical demands of representative match play in adolescent rugby union. *Journal of*
480 *Strength and Conditioning Research*, *31*, 1290–1296.
481 <http://doi.org/10.1519/JSC.0000000000001600>
- 482 Reilly, T., Williams, A. M., Nevill, A., & Franks, A. (2000). A multidisciplinary approach to
483 talent identification in soccer. *Journal of Sports Sciences*, *18*, 695–702.
484 <http://doi.org/10.1080/02640410050120078>
- 485 Roe, G., Darrall-Jones, J. D., Till, K., Phibbs, P., Read, D., Weakley, J., & Jones, B. (2016).
486 Between-day reliability and sensitivity of common fatigue measures in rugby players.
487 *International Journal of Sports Physiology and Performance*, *11*, 581–586.
488 <http://doi.org/10.1123/ijsp.2015-0413>
- 489 Sheppard, J. M., & Young, W. B. (2006). Agility literature review: Classifications, training
490 and testing. *Journal of Sports Sciences*, *24*, 919–932.
491 <http://doi.org/10.1080/02640410500457109>
- 492 Stewart, P. F., Turner, A. N., & Miller, S. C. (2012). Reliability, factorial validity and
493 interrelationships of five commonly used change of direction speed tests. *Scandinavian*
494 *Journal of Medicine & Science in Sports*, *24*, 500–506.
495 <http://doi.org/10.1111/sms.12019>

- 496 Young, W. B. (2006). Transfer of strength and power training to sports performance.
497 *International Journal of Sports Physiology and Performance*, 1, 74–83.
498 <http://doi.org/10.1123/ijsp.1.2.74>
- 499 Young, W., McLean, B., & Ardagna, J. (1995). Relationship between strength qualities and
500 sprinting performance. *Journal of Sports Medicine and Physical Fitness*, 35, 13–19.
501
502

Table 1: Ambient conditions for field testing for each sport.

Sport	Temperature (°C)	Humidity (%)	Air Pressure (hPa)	Wind (m/s)	Description
Basketball Day 1	18	54	1008	North (4.6)	Scattered Clouds
Basketball Day 2	12	82	1010	South (3.6)	Haze
Cricket Day 1	17	58	1008	North (4.6)	Scattered Clouds
Cricket Day 2	12	82	1010	South (3.6)	Haze
Football Day 1	19	46	1013	ENE (4.6)	Scattered Clouds
Football Day 2	13	58	1020	West (1.5)	Partly Cloudy
Hockey Day 1	18	54	1008	North (4.6)	Scattered Clouds
Hockey Day 2	12	82	1010	South (3.6)	Haze

Netball					Scattered
Day 1	18	54	1008	North (4.6)	Clouds
Netball					Overcast
Day 2	14	69	1010	SSE (3.1)	
Rugby					Scattered
Day 1	17	53	1013	ENE (4.6)	Clouds
Rugby					Partly
Day 2	12	67	1020	West (2.1)	Cloudy

Note: Running direction was standardised as North West for sprints and South East
(before the change of direction) for the 5-0-5 test.

503

504

Table 2: Summary of day 1 and day 2 raw values, TE as a CV (90% confidence intervals in brackets) and SWC as percentages, TE:SWC ratio and usefulness (Hopkins, 2000) for each test.

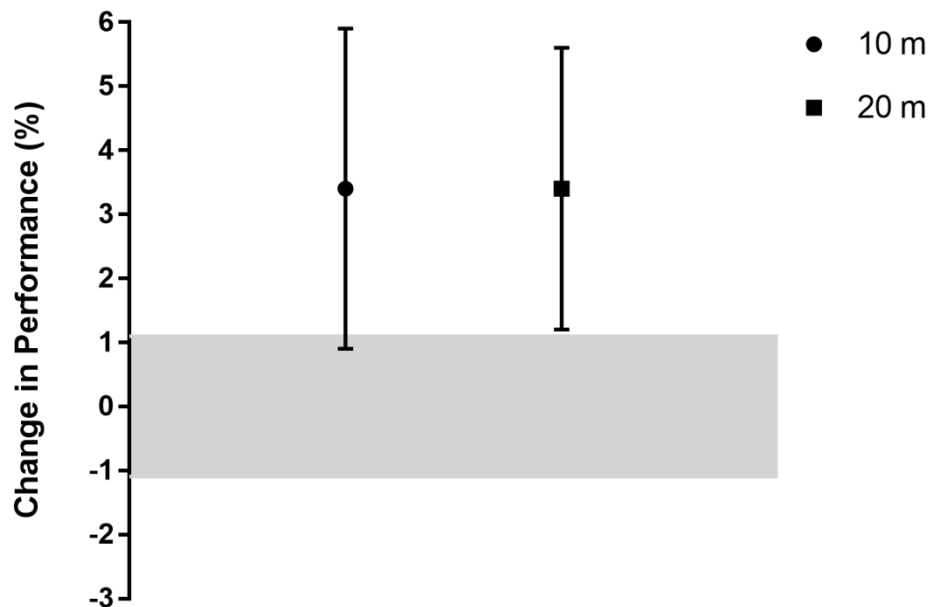
Test	Day 1	Day 2	TE as a CV (%)	SWC (%)	TE:SW C ratio	Usefulness
IMTP (kg)	170.6 ± 45.5	170.9 ± 46.4	5.5 (4.5 - 6.9)	6.1	0.9	Good
CMJ (cm)	34.4 ± 5.9	34.4 ± 6.4	2.8 (2.4 - 3.3)	3.9	0.7	Good
5 m (s)	1.08 ± 0.06	1.06 ± 0.06	2.7 (2.0 - 4.0)	1.1	2.5	Marginal
10 m (s)	1.82 ± 0.09	1.78 ± 0.11	2.5 (2.1 - 3.2)	1.1	2.3	Marginal
20 m (s)	3.19 ± 0.17	3.10 ± 0.19	2.2 (1.9 - 2.8)	1.1	2.0	Marginal
30 m (s)	4.45 ± 0.28	4.37 ± 0.28	2.2 (1.8 - 2.7)	1.3	1.7	Marginal
40 m (s)	5.75 ± 0.38	5.68 ± 0.42	1.8 (1.5 - 2.3)	1.4	1.3	Marginal
5-0-5 L (s)	2.54 ± 0.21	2.50 ± 0.22	4.1 (3.4 - 5.4)	1.7	2.4	Marginal
5-0-5 R (s)	2.49 ± 0.20	2.52 ± 0.25	5.4 (4.4 - 7.0)	1.8	3.0	Marginal

IMTP = Isometric Mid Thigh Pull; CMJ = Countermovement jump; L = Left; R = Right

505

506

507 Figure 1: An example of the change in performance of an athlete over two tests using our
508 reliability data. Data are percentage change in performance (\pm TE as a CV as error bars). The
509 shaded grey area represents the SWC, which is the same for both tests. Although the
510 magnitude of improvement in performance is the same, the difference in the TE results in the
511 error bar overlapping the SWC for the 10 m split, leading to an inability to describe the
512 changes as "real" with 75% confidence. The error bar for the 20 m split does not overlap the
513 SWC however, resulting in a clear improvement in performance, with 75% certainty
514 (Hopkins, 2000).



515

516