The effect of trial familiarisation on the validity and reproducibility of a field-based self-paced VO$_2$max test

AUTHORS: Lim W$^1$, Lambrick D$^2$, Mauger AR$^3$, Woolley B$^4$, Faulkner J$^4$

$^1$ Massey University, New Zealand  
$^2$ University of Southampton, United Kingdom  
$^3$ University of Kent, United Kingdom  
$^4$ University of Winchester, United Kingdom

ABSTRACT: The self-paced maximal oxygen uptake (VO$_2$max) test (SPV), which is based on the Borg 6-20 Ratings of Perceived Exertion (RPE) scale, allows participants to self-regulate their exercise intensity during a closed-loop incremental maximal exercise test. As previous research has assessed the utility of the SPV test within laboratory conditions, the purpose to this study was to assess the effect of trial familiarisation on the validity and reproducibility of a field-based, SPV test. In a cross-sectional study, fifteen men completed one laboratory-based graded exercise test (GXT) and three field-based SPV tests. The GXT was continuous and incremental until the attainment of VO$_2$max. The SPV, which was completed on an outdoor 400m athletic track, consisted of five x 2 min perceptually-regulated (RPE11, 13, 15, 17 and 20) stages of incremental exercise.

There were no differences in the VO$_2$max values from the SPV, the application of a single SPV test is an appropriate stand-alone protocol for gauging VO$_2$max. Reliability implies greater precision of one-off measures and better tracking of changes in measurements in research or practical settings [7].

The self-paced maximal oxygen uptake (VO$_2$ max) test (SPV), which is based upon the Borg 6-20 Ratings of Perceived Exertion (RPE) scale, allows participants to self-regulate their exercise intensity during a closed-loop incremental maximal exercise test. Laboratory-based SPV protocols have been shown to elicit higher[1-3] or comparable[4-6] VO$_2$max values to those reported from a conventional open-loop laboratory-based graded exercise test (GXT). Despite these findings, caution should be exercised with their interpretation as the reliability of the SPV protocol has yet to be examined. Enhanced reliability implies greater precision of one-off measures and better tracking of changes in measurements in research or practical settings [7].

The self-paced nature of exercising in an outdoor environment, where an individual is free to vary their pace, cannot be easily replicated in the laboratory environment [8]. The ecological validity of laboratory-based running protocols are reduced as the conscious decision to manually control the treadmill belt speed does not occur as quickly or as frequently as during self-paced running outside [9,10]. As there are small but significant differences between running on a track and running on a treadmill due to variation in airstream, ground surface and movement patterns [8], the transferability of laboratory measurements to training and competitive situations may be limited.

Accordingly, it is of interest to investigate the efficacy and reproducibility of the SPV in measuring VO$_2$max and the pace response during field-based exercise. Should the SPV prove to be reliable, this test may be of practical value to coaches when monitoring and prescribing exercise for athletes.

The purpose of this study was to assess the concurrent validity and reproducibility of a field-based SPV test in comparison to a conventional GXT. It was hypothesised that the SPV protocol would validly measure VO$_2$max and elicit reproducible findings in relation to VO$_2$max and the pacing response across three field-based SPV tests.
MATERIALS AND METHODS

Subjects. This was a cross-sectional experimental study wherein fifteen men (23.9 ± 4.5 y, 1.74 ± 0.05 m, 73.9 ± 7.5 kg) who were recreationally-trained (>3 h·wk⁻¹ of vigorous athletic training) participated in the study. Based on effect sizes (δ = 0.88) and mean (± SD) statistics for VO₂max [1], a minimum sample size of n = 12 was calculated to achieve a statistical power of 80% at an alpha level of .05. Participants were injury free, healthy and asymptomatic of any illness as confirmed through health screening procedures [11]. All participants had previous experience of undertaking maximal exercise testing in a laboratory environment, but none had completed the SPV test. Institutional ethical approval was obtained prior to the study in accordance with the spirit of the Helsinki Declaration, and participants provided written informed consent.

Procedures

Following basic anthropometric measurements, participants completed a laboratory-based GXT (19.6 ± 1.2 °C, 40.8 ± 5.1% [humidity], 1001.2 ± 8.3 hPa) on a motorised treadmill (True 825, Fitness Technologies, St Louis, USA). Subsequently, participants completed three field-based SPV exercise tests (SPV1, SPV2, SPV3), on an outdoor 400m athletics track at the same time of day (±1h) (19.8 ± 3.5 °C, 51.7 ± 13.9% [humidity], 1012.3 ± 6.3 hPa, 2.9 ± 2.1 km·h⁻¹ [maximum wind speed]). All tests were separated by a 72h recovery period in which participants refrained from any normal training activities. Respiratory gases (oxygen uptake [VO₂], ventilation [VE], respiratory exchange ratio [RER]) were continuously measured using a portable breath-by-breath sampling system (K4-b-TX Module, Cosmed, Roe, Italy), and a global positioning system (GPS) watch (Suunto Ambit, Vaarimote, Finland) was used to assess heart rate and the pacing response (using 5 s data averages) throughout each exercise test. Blood lactate (Lactate Pro, Kyoto, Japan) was assessed prior to and 1 min post each exercise test. Participants were perceptually anchored to the Borg 6-20 RPE scale [12] and to a 0-10 localized-leg-pain scale[13] by definition and recall procedures. By definition, participants were instructed that the differing numerical values equated to the feelings associated with the corresponding written definitions on each scale (e.g., RPE 20 equates to ‘maximal exertion’). Anchoring by recall refers to encouraging the participant to remember the range of feelings previously experienced during exercise of a similar nature (e.g., Pain 10 equates to an extremely intense pain). Physiological data (and treadmill speeds during the GXT) were masked from participants throughout both exercise tests.

During the GXT, VO₂max was confirmed by a visible plateau in oxygen consumption of ≤ 2 mL·kg⁻¹·min⁻¹ with a standard increment in exercise intensity, and / or any secondary criterion indicators (visible signs of exhaustion; HRmax = 10 b·min⁻¹; RER ≥ 1.15), at or around the point of volitional exhaustion [11]. The VO₂max value determined in the GXT was clarified using a verification stage [14].

The design and closed-loop nature of the SPV, whereby a non-linear change in running speed is expected in the final stage of the test, the highest measure of VO₂ was taken as the VO₂max, independent of changes in running speed.

Laboratory-based GXT

Participants initially completed a self-directed warm-up on the treadmill (2.5 min at a running speed equivalent to an RPE11; 2.5 min at an RPE13). The GXT commenced at the speed which was equivalent to RPE13 from the warm-up; a speed deemed sufficient to elicit VO₂max within 10 (±2) min [14]. The GXT was continuous and incremental, commencing at the chosen speed and increasing by 1 km·hr⁻¹ every 2 min thereafter until volitional exhaustion. The treadmill gradient was set at 1% to reflect the energy cost associated with outdoor running [15]. During the final 20 s of each increment of the GXT, participants reported their overall RPE and localised pain perception. Following a 15 min recovery period, participants completed a VO₂max verification stage whereby speed was gradually increased over a 30 s period to a speed which was 1 km·hr⁻¹ higher than the final stage of the GXT. Participants exercised at this elevated speed until volitional exhaustion. Respiratory markers were monitored throughout. Peak speed at VO₂max was the running speed at which VO₂max occurred, provided that the running time was greater than 1 min.

Field-based SPV

Following a similar warm-up to the GXT, participants commenced the SPV. This test was 10 min in duration and comprised five x 2 min stages of incremental exercise, which were adjusted according to prescribed RPE levels, equating to a light (RPE11), somewhat hard (RPE13), hard (RPE15), very hard (RPE17) and maximal RPE (RPE20). Participants were instructed to modify their running speed on a moment-to-moment basis in line with the prescribed RPE, rather than the end-point of the task, so that their RPE (not their speed) remained clamped for each given 2 min stage. The RPE scale was visible at regular intervals (every 100 m). During the final 20 s of each increment of the SPV, the researchers would show the exercise-induced pain scale to the participants whilst they were running to obtain their current perception of pain.

Following a 15 min recovery period, participants completed a VO₂max verification stage whereby participants repeated the final two, 2 min stages of the SPV test (RPE17 & RPE20). Respiratory markers were monitored throughout.

Statistical analysis

Breath-by-breath data (VO₂, VE, RER) from the GXT and SPV tests were averaged into 10 s bins and used in the following analyses. To assess the concurrent validity and reproducibility of the SPV, one-way ANOVAs compared the maximal physiological (VO₂max, HRmax, etc.), perceptual and physical (peak speed, average speed, distance, duration) data reported from the GXT to each SPV test (e.g., GXT vs. SPV1), and between each SPV test (e.g., SPV1 vs. SPV2). A Bland and Altman 95% limits-of-agreement (LoA) analysis quantified the
Validit y and reproducibility of a field-based SPV VO₂max test

agreement (bias ± random error [1.96 × SD]) between the VO₂max reported from the GXT and each of the SPV tests [16]. Intraclass correlation coefficients (ICCs) were calculated to quantify the reproducibility of the maximal physiological criteria between the GXT and SPV tests, and between each of the SPV tests. A reliability coefficient (the smallest detectable difference) was also used to determine the critical difference in a parameter that must be exceeded between two sequential results in order for a statistically significant change to occur in an individual. A one-sample t-test was used to compare the maximal RPE from the GXT and SPV (RPE20).

To assess the pacing response during the SPV, two factor repeated measures ANOVA’s; Test (SPV1, SPV2, SPV3) by RPE (RPE11, RPE13, RPE15, RPE17, RPE20) were used to compare the distance covered, and the peak, mean and end running speeds for each SPV test. Where significant differences were reported, Tukey’s HSD was used to identify the location of any statistical differences (between SPV1, SPV & SPV3). A similar analysis was conducted to compare the coefficient of variation (CV) in running speed between each of the five perceptual intensities. The CV was calculated using the following equation:

\[ c_v = \frac{\sigma}{\mu} \]

Whereby, \( \sigma \) is the standard deviation of the running speed for a given perceptual intensity, and \( \mu \) is the mean. Two-factor repeated measures ANOVA’s were also used to assess the physiological responses (VO₂, HR, V̇E, RER) during the SPV tests. Partial eta squared \( (\eta_p^2) \) was used to demonstrate the effect size, with .0099, .0588 and .1379 representing a small, medium and large effect, respectively [17]. Partial eta squared was calculated using the following formula:

\[ \eta_p^2 = \frac{SS_{Effect}}{SS_{Effect} + SS_{Error}} \]

Whereby, \( SS_{Effect} \) is the estimated variance for a given outcome measure, and \( SS_{Error} \) is the error variance that is attributable to the effect. Alpha was set at 0.05 throughout all analyses. All data were analysed using the statistical package SPSS for Windows, PC Software, version 22.

RESULTS

As there were no differences between VO₂max values reported from the GXT and SPV when compared to their corresponding verification test, the GXT and SPV data are used in the proceeding analyses.

### TABLE 1. Mean (± SD) physiological, perceptual and physical responses recorded at completion of the GXT and SPV tests.

<table>
<thead>
<tr>
<th></th>
<th>GXT</th>
<th>SPV1</th>
<th>SPV2</th>
<th>SPV3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO₂max (ml·kg⁻¹·min⁻¹)</td>
<td>63.5 ± 10.1</td>
<td>65.5 ± 8.7</td>
<td>65.4 ± 7.0</td>
<td>66.7 ± 7.7</td>
</tr>
<tr>
<td>V̇E max (L·min⁻¹)</td>
<td>155 ± 17.2</td>
<td>161 ± 16.9</td>
<td>161 ± 16.5</td>
<td>163 ± 16.0</td>
</tr>
<tr>
<td>HR max (b·min⁻¹)</td>
<td>192 ± 6.8</td>
<td>191 ± 7.1</td>
<td>190 ± 8.0</td>
<td>191 ± 7.8</td>
</tr>
<tr>
<td>Maximal RER</td>
<td>1.00 ± 0.9</td>
<td>1.05 ± 0.1</td>
<td>1.01 ± 0.1</td>
<td>1.00 ± 0.1</td>
</tr>
<tr>
<td>Maximal Blood Lactate (mmol·L⁻¹)</td>
<td>11.7 ± 3.1</td>
<td>11.0 ± 2.5</td>
<td>10.3 ± 2.4</td>
<td>10.9 ± 2.0</td>
</tr>
<tr>
<td>Maximal Pain</td>
<td>8.4 ± 2.3</td>
<td>8.0 ± 2.5</td>
<td>7.7 ± 2.6</td>
<td>7.6 ± 2.6</td>
</tr>
<tr>
<td>Maximal RPE</td>
<td>19.8 ± 0.4</td>
<td>20.0 ± 0.0</td>
<td>20.0 ± 0.0</td>
<td>20.0 ± 0.0</td>
</tr>
<tr>
<td>Peak Speed (km·h⁻¹)</td>
<td>17.7 ± 0.9</td>
<td>19.8 ± 2.5</td>
<td>18.9 ± 2.8</td>
<td>19.9 ± 2.9</td>
</tr>
<tr>
<td>Average Speed (km·h⁻¹)</td>
<td>13.8 ± 1.0</td>
<td>14.4 ± 1.1*</td>
<td>14.6 ± 1.1*</td>
<td>15.0 ±1.2*</td>
</tr>
<tr>
<td>Total Distance (m)</td>
<td>2620 ± 284.6</td>
<td>2389 ± 160*</td>
<td>2425 ± 177</td>
<td>2462 ± 184</td>
</tr>
<tr>
<td>Test Duration (sec)</td>
<td>621 ± 43</td>
<td>600 ± 0</td>
<td>600 ± 0</td>
<td>600 ± 0</td>
</tr>
</tbody>
</table>

Note: *Significant difference between SPV and GXT (P < 0.001).

### TABLE 2. Reliability of the physiological, perceptual and physical responses recorded from all SPV tests. Data is presented as the mean difference between tests, ICC, SEM and RC.

<table>
<thead>
<tr>
<th></th>
<th>SPV 1-SPV2</th>
<th>SPV2-SPV3</th>
<th>SPV1, SPV2 &amp; SPV3</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO₂max (ml·kg⁻¹·min⁻¹)</td>
<td>Mean Diff (±) SD</td>
<td>ICC</td>
<td>SEM</td>
</tr>
<tr>
<td></td>
<td>0.05 ± 12.0</td>
<td>0.81</td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>0.27 ± 3.6</td>
<td>0.95</td>
<td>3.62</td>
</tr>
<tr>
<td></td>
<td>0.7 ± 23.6</td>
<td>0.94</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>0.5 ± 10.6</td>
<td>0.35</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>0.04 ± 0.12</td>
<td>0.66</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>0.69 ± 3.6</td>
<td>0.94</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>0.9 ± 3.4</td>
<td>0.83</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>-0.2 ± 1.6</td>
<td>0.92</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>-26 ± 293</td>
<td>0.95</td>
<td>44.7</td>
</tr>
</tbody>
</table>

Note: ICC, Intraclass correlation coefficients; SEM, Standard error of the mean; RC, Reliability coefficient
There were no differences in the maximal physiological or perceptual values reported when comparing the GXT with each of the SPV tests (all \( P > .05 \); Table 1). Significant differences in average speed (\( F(3,56) = 20.54, \ P < .001, \eta^2 = .60 \)) and total distance (\( F(3,56) = 3.601, \ P < .05, \eta^2 = .16 \)) were however observed (Table 1). For the \( \dot{V}O_2 \max \) reported from the GXT and SPV1, SPV2 and SPV3, the corresponding ICC and 95% LoA were; 0.85 and 2.01 ± 13.26 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV1), 0.85 and 1.96 ± 12.28 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV2) and 0.87 and 3.27 ± 11.93 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV3, respectively).

**Concurrent validity: GXT vs SPV 1, SPV 2, & SPV 3**

There were no differences in the maximal physiological or perceptual values reported when comparing the GXT with each of the SPV tests (all \( P > .05 \); Table 1). Significant differences in average speed (\( F(3,56) = 20.54, \ P < .001, \eta^2 = .60 \)) and total distance (\( F(3,56) = 3.601, \ P < .05, \eta^2 = .16 \)) were however observed (Table 1). For the \( \dot{V}O_2 \max \) reported from the GXT and SPV1, SPV2 and SPV3, the corresponding ICC and 95% LoA were; 0.85 and 2.01 ± 13.26 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV1), 0.85 and 1.96 ± 12.28 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV2) and 0.87 and 3.27 ± 11.93 ml·kg\(^{-1}\)·min\(^{-1}\) (GXT vs. SPV3, respectively).

**Reproducibility: SPV1 vs SPV2 vs SPV3**

There were no differences in the maximal physiological, perceptual or physical values reported when comparing the three SPV tests (all \( P > .05 \); Tables 1 and 2). Although RER and BLa provided low to moderate ICCs, all other parameters provided strong ICCs between SPV1 and SPV2, SPV2 and SPV3, and between all SPV tests (all \( \geq 0.80 \)).

**Pacing response during SPV tests**

Figure 1 demonstrates the mean (± SD) pacing strategy at each perception of exertion for SPV1-SPV3. There were no Test by RPE interactions for the peak speed, mean speed, end speed and for the...
Validity and reproducibility of a field-based SPV VO₂max test

distance ran from the SPV tests (all $P > .05$; Figure 2). Similar findings were observed when considering the CV for running speed ($F(3.1,42.3) = 1.44, P > .05; \eta_p^2 = .09$; Table 3). An RPE main effect was observed for peak ($F(1,4,19.1) = 18.06, P < .001, \eta_p^2 = .56$), mean ($F(1,4,19.6) = 26.69, P < .001, \eta_p^2 = .66$) and end speed ($F(2.1,29.4) = 16.95, P < .001, \eta_p^2 = .55$), and for the distance ran ($F(2.3,32) = 23.47, P < .001, \eta_p^2 = .63$), with significant increases occurring at various increments in exercise intensity (all $P < .01$; Figure 2). For the CV in running speed, significant changes were only observed between RPE11 and RPE13 ($F(1,14) = 18.5, P < .01, \eta_p^2 = .57$) and RPE17 and RPE20 ($F(1,14) = 14.1, P < .01, \eta_p^2 = .50$; Table 3).

Significant Test main effects were observed for peak ($F(2,28) = 4.35, P < .05; \eta_p^2 = .24$), mean ($F(1,4,19.4) = 7.88, P < .01; \eta_p^2 = .36$) and end speed ($F(2,28) = 6.05, P < .01, \eta_p^2 = .30$), and distance ran ($F(2,28) = 6.35, P < .01, \eta_p^2 = .31$), with participants in SPV3 running, on average, faster and further than both SPV1 and SPV2.

**Physiological response during SPV tests**

A significant Test by RPE interaction was observed for HR ($F(4.3, 60.1) = 4.19, P < .05; \eta_p^2 = .23$), $V_{E}$ ($F(8,112) = 3.98, P < .001, \eta_p^2 = .22$) and RER ($F(2.8,39.1) = 3.19, P < .05, \eta_p^2 = .19$; Figure 3). During SPV1, participants experienced a greater change in HR, $V_{E}$ and RER between RPE13 and RPE15 than either SPV2 or SPV3. RPE main effects were observed for VO₂, HR, $V_{E}$, RER and pain (all $P < .001$; Figure 3).

**DISCUSSION**

This study assessed the validity and reproducibility of a field-based self-paced VO₂max test in recreationally-trained men. In this study, SPV1, SPV2 and SPV3 elicited similar VO₂max values to those obtained from a traditional laboratory-based GXT. Accordingly, and in support of the study hypotheses, a field-based SPV test may be con-
considered a valid test of VO$_2$max, in keeping with previous laboratory-based research [4-6]. Furthermore, the SPV test was considered reliable, as a period of trial familiarisation did not statistically moderate the VO$_2$max or pacing response elicited.

Despite no statistical differences and a high correlation (ICC: 0.85-0.87) between the GXT and the three SPV tests, slightly higher VO$_2$max values were observed with the SPV (3.1%, 3.0% & 4.8% for SPV1, SPV2 & SPV3, respectively), similar to past research [1, 2, 5]. This is pertinent from a coach’s perspective as positive changes in aerobic capacity in the region of 3-5% have been accepted as a meaningful performance improvement [18, 19]. Therefore, it is important to recognise: i) the practical implications of measuring an athlete’s maximal aerobic capacity using differing methods of assessment, and ii) that the SPV may consistently elicit higher practical values than a laboratory-based GXT. However, it should be noted that near optimal environmental conditions were experienced in the course of the field data collection (e.g., wind speed $\leq$ 3 m·s$^{-1}$) and thus, our findings should be considered in the context of this. Furthermore, as there appears to be a difference in VO$_2$max that is of practical significance, it is likely that the non-statistical difference in VO$_2$max between the GXT and SPV in the current study was due to the wide variance associated with the fitness level of the participants recruited (i.e., recreationally-trained). Future research should therefore ensure that participants are more closely matched for fitness when assessing the validity and reliability of the SPV protocol.

In this study, a mean difference in VO$_2$max of 0.05 mL·kg$^{-1}$·min$^{-1}$ (0.2%) and 1.3 mL·kg$^{-1}$·min$^{-1}$ (1.8%) was observed between SPV1 and SPV2, and SPV2 and SPV3, respectively. This is encouraging when considering that laboratory-based maximal exercise tests have shown a 3 to 6% variation in VO$_2$max values following three or more repeated trials [19, 20]. The reliability of the SPV protocol is further demonstrated using ICC (all >0.80), SEM and RC analyses for VO$_2$max and other variables of interest (Table 2). Accordingly, this study confirms that the application of a single SPV, which to date, has been the only way the test has been implemented previously [1-6], may be appropriate for gauging VO$_2$max. However, as the ICC and RC’s were shown to improve with trial familiarisation (Table 2), and as participants ran, on average, faster and further during SPV3 than either SPV1 (4.3 & 3.1%, respectively) or SPV2 (2.8 & 1.6%, respectively), there may be benefit in repeat assessments.

Despite the positive VO$_2$max findings, it is an athlete’s running speeds and pacing response at differing submaximal intensities that are perhaps more pertinent for training prescription. In this study, participants’ changes in speed (peak, mean and end speed) and distance between the five stages of the SPV protocol were similar between SPV1, SPV2 and SPV3 (see Figure 2). Although a more variable running speed was observed at the start (RPE11) and end (RPE20) of each SPV (Table 3), a similar overall pacing response was observed between the three trials. Each of the perceptual intensities corresponded to 66-69% (RPE11), 69-73% (RPE13), 75-79% (RPE15) and 78-82% (RPE17) of the peak running speed observed from the VO$_2$max test. Knowledge of the speed, distance and heart rate responses at submaximal and maximal SPV stages may be a useful reference for coaches when determining and prescribing appropriate training intensities, negating the need for expensive equipment, designated laboratory facilities and trained technicians associated with traditional, maximal GXT protocols.

Regardless of the SPV test, the pacing response varied within SPV protocols. Although continuous increments in peak running speed were observed after RPE13, there were no differences in the peak running speed between RPE11 and RPE13 (Figure 2a). The mean speed and end speed for RPE11 was however lower than RPE13 (Figure 2b & 2c), suggesting that participants adopted an inappropriate starting speed (e.g. ran too fast) at the start of the protocol. In this regard, it is plausible that afferent feedback involving physiological systems, environmental surroundings and psychological constructs (mood, self-efficacy, etc.) helped to adjust the pacing response after the initial peak speed was achieved to ensure that an appropriate running speed was elicited thereafter. This appears to be corroborated by the CV in running speed, which demonstrated a more variable pacing response at RPE11 compared to RPE13 (6.4 ± 3.1 vs. 2.9 ± 1.1%, respectively).

Knowledge of the exercise end-point has been suggested to be the single most important factor in influencing a pacing strategy [21]. In this study, a more conservative pacing strategy was observed during the latter portion (RPE15 & RPE17) of the SPV (Figure 2). This finding is complemented by the physiological data, as similar VO$_2$, HR, $V_{E}$ and RER values were reported between RPE15 and RPE17 (Figure 3). The anticipated 2 min maximal sprint at RPE20 may have encouraged participants to utilise a more conservative pacing strategy during the final moments of the penultimate stage of the exercise test. Similar findings have been shown elsewhere [3,22,23]. For example, Sperlich et al. [23] demonstrated that when participants could manipulate both treadmill speed and inclination throughout a self-selected maximal exercise test, participants often elicited a rapid change in one or both of these factors in the final few minutes of the exercise test. In the present study, it is also of interest to note that RPE20 elicited a greater CV in running speed than RPE17 (Table 3). The greater variation at RPE20 is likely due to the rapid increase and attempted maintenance of peak speed during this stage, stimulating a greater accumulation of metabolic by-products, and thus facilitating a greater perception of pain and discomfort than that associated with RPE17. Consequently, the SPV protocol may be better suited to experienced athletes as it may reflect the physiological and pacing demands encountered during competition.

The researchers do recognise certain limitations to the study. In the present study, participants completed their laboratory-based GXT prior to any of the field-based SPV tests. To minimise the risk of any potential confounding effects associated with test order, it would have been useful to implement a randomised and counterbalanced cross-over design. Our study also compared the physiological, physical and perceptual responses between treadmill and over-ground running.
Validity and reproducibility of a field-based SPV VO₂max test

As kinematic [24,25] and perceptual [26] differences may be evident between these ambulatory modalities, it may be speculated that these differences could have contributed to some of the statistical differences reported in Table 1. Similar to Hogg and colleagues [3], the lack of difference in peak running speed between the treadmill GXT and the field-based SPV may be due to limitations in achieving a ‘true’ peak running speed during treadmill exercise, which in-turn may confound the validity of the reported findings.

CONCLUSIONS

This is the first study to demonstrate the concurrent validity and reliability of a field-based SPV. Measures of VO₂max in the field-based SPV were not statistically different from those in the laboratory-based GXT, and trial familiarisation did not moderate the SPV were not statistically different from those in the laboratory-based GXT and the field-based SPV. Measures of VO₂max in the field-based SPV were not statistically different from those in the laboratory-based GXT, and trial familiarisation did not moderate the SPV. Measures of VO₂max and the field-based SPV may be due to limitations in achieving a ‘true’ peak running speed during treadmill exercise, which in-turn may confound the validity of the reported findings.

Acknowledgements

The authors would like to thank the Massey University Summer Scholarship scheme for supporting the costs of the study.

Conflict of interests: the authors declared no conflict of interests regarding the publication of this manuscript.

REFERENCES