Narrative Review

Paradigm shift in critical speed application and modification for shuttle running: a narrative review

Mark Kramer [1*], and Eva Piatrikova [2]

[1] Physical Activity, Sport, and Recreation (PhASRec) Unit, North-West University, Potchefstroom, South Africa; mark.kramer@nwu.ac.za

[2] Department of Health, University of Bath, Bath, United Kingdom; ep398@bath.ac.uk

* Correspondence: mark.kramer@nwu.ac.za

Abstract: The overarching purpose of this review was to highlight the utility of different aerobic field tests in terms of the parameters they provide, with a specific focus on shuttle running and all-out testing. Various field tests are discussed in detail and are categorised according to *linear continuous running tests* (e.g. 12-minute Cooper Test, University of Montreal Track Test [UMTT], 1200/1600 m time trials, 3-minute all-out test for running [3MT]), *intermittent shuttle running tests* (e.g. yo-yo intermittent recovery test level 1 [YYIR1], 30-15 intermittent fitness test [IFT], and the intermittent all-out shuttle test [IAOST]), and *continuous shuttle running tests* (e.g. 1.2 km shuttle run test [1.2$_{SRT}$], maximal multi-stage 20-m shuttle test [MSR], 25-m, 30 m and 50-m 3-minute all-out shuttle test [AOST]). Readers will be guided through the theoretical and practical underpinnings of the 3MT methodology, where the all-out testing methodology is stationed within the testing paradigm, and how to practically implement and interpret the results thereof.

Keywords: Critical speed, exercise prescription, team sports, thresholds, shuttle running

1. Introduction

The regular testing and monitoring of an athletes' performance status is a well-established component of a successful training programme. However, the associated testing procedures are especially challenging to conduct in team sports. Hindrances to testing at scale are primarily due to the large number of players involved as well as the limited time, resources, and physiological testing expertise often available to coaches in such multifaceted sports.

Gradations of maximal oxygen uptake (VO$_2$max), which is considered a valid proxy for cardiorespiratory fitness (CRF), tends to discretize different levels of performance in field sports such as rugby, soccer and hockey [1–4]. As match-play requirements escalate from junior to senior levels, or from regional to international levels, so too do the CRF requirements in order to meet the physiological demands imposed by commensurately higher levels of match-play. Consequently, many coaches require regular physiological testing of athletes in order to (i) monitor and predict physical performance, (ii) assess talent, and (iii) possibly design training programmes [5].

Historically, the gold standard for the assessment of CRF was, and still is, the laboratory-based graded exercise test (GXT) [6,7]. Although first developed in the early 20th century, and despite nuanced variations in stage gradations and durations, the essence of the test has remained intact for almost 100 years [8]. While the GXT is considered the gold standard, it exhibits several limitations in that it (i) fails to scale, (ii)

shows methodological variability, (iii) is costly, (iv) requires substantial expertise, and (v) possibly lacks ecological validity for field-based sports [9]. Based on these limitations the need for valid and reliable field-based testing was borne, leading to the subsequent development of several methodologically grouped tests such as:

(i)   *linear continuous running tests* (e.g. 12-minute Cooper Test, University of Montreal Track Test [UMTT], 1200/1600 m time trials, 3-minute all-out test for running [3MT]) [10–12]

(ii)   *intermittent shuttle running tests* (e.g. yo-yo intermittent recovery test level 1 [YYIR1], 30-15 intermittent fitness test [IFT], and the intermittent all-out shuttle test [IAOST]) [13–15].

(iii)   *continuous shuttle running tests* (e.g. 1.2 km shuttle run test [1.2_{SRT}], maximal multi-stage 20-m shuttle test [MSR], 25-m, 30 m and 50-m 3-minute all-out shuttle test [AOST]) [16–20]

In most instances the purpose of a given test is to produce a measure that relates to a specific physiological state (e.g. VO_2max or sVO_2max) [10,21]. For this reason, many of the tests are validated against the 'gold-standard' whereby field-test performance outcomes, such as end-stage speed, are used to estimate the laboratory equivalent VO_2max (or sVO_2max) to varying degrees of success [22,23]. Although numerous research studies have used shuttle running to predict VO_2max through various regression equations [14,17,24], there appears to be a dearth of research *directly* relating the maximal VO_2 achieved during a graded exercise test (GXT; lab or field) to that achieved during a variant of the shuttle test. Some evidence suggests that the VO_2 response during intermittent activity is significantly different when measured in the laboratory compared to on the field, with the VO_2 response being approximately 3.3 $ml.kg^{-1}.min^{-1}$ higher during intermittent field running compared to treadmill running [25]. Furthermore, research indicates that the VO_2 response is dependent on both the turn count (i.e. for the same relative velocity, end exercise VO_2 increases linearly with the number of shuttle turns) as well as the velocity of running (i.e. for the same number of turns, VO_2 increases linearly with speed) [26,27].

Interestingly, some research has shown that the upper threshold requirement of VO_2max for elite soccer and rugby may be around 60 $ml.kg^{-1}.min^{-1}$ [28–30]. It is therefore likely that other skills are more important than a reliance on higher VO_2max values at, or beyond, a given level of play. A possible implication here is that the evaluation of VO_2max in field-sports is important only up to a certain point. Thereafter, other parameters, such as submaximal thresholds (e.g. lactate threshold [LT], maximal lactate steady state [MLSS], critical speed [CS] etc), maximal sprint speed (MSS), skills and tactics should be more readily evaluated as these may be more distinguishing and provide more valuable performance information than VO_2max alone.

While numerous field-based tests exist to evaluate the CRF of athletes, these are often limited to the indirect assessment of VO_2max through various regression equations [14,31–34]. Although useful, most research is this line fails to unpack the practical utility

of the field test beyond just providing a surrogate estimate of CRF, nor how test parameters may be used to guide performance assessment and/or training prescription.

As such, the purpose of this narrative review is to highlight the utility of different field tests in terms of the parameters they provide, with a specific focus on all-out testing. Readers will be guided through the theoretical and practical underpinnings of the 3MT methodology, where the 3MT is stationed within the testing spectrum, and how to practically implement and interpret the results thereof. Information will be highlighted in terms of the physiological state that each test is meant to represent and how these parameters are typically utilised. The following sections will expand on the various field tests which have been categorised according to linear continuous running tests, intermittent shuttle running tests, and continuous shuttle running tests.

2. Linear continuous running tests

The 12-minute Cooper test is arguably one of the most widely used linear continuous field tests to predict VO_2max based on the distance achieved during the test ($r = 0.90$-0.92) [24,35]. Although such an approach has the benefit of enabling testing on a larger scale, it is reliant on high levels of effort at a relatively constant pace for the entirety of the test. Furthermore, the only outcome of the test is an approximation of the VO_2max, based on the distance achieved, without taking cognisance of the average speed maintained during the test, nor the heart rate response throughout the test. Therefore, the utility of the test for programme prescription is limited based on the fact that prescribing intensities as a percentage of VO_2max is not considered a valid method for exercise prescription on the basis that this leads to substantial heterogenous homeostatic responses both in and across individuals [9]. The utility of the test for evaluating field-sport athletes may also questionable on the basis that no changes of direction are incorporated into the test.

Another popular field test used to predict VO_2max is the University of Montreal Track Test (UMTT) [10]. The UMTT is a graded test in that it starts at a speed of 8 $km.hr^{-1}$ and increases by a speed of 1 $km.hr^{-1}$ every 2-minutes [14]. Such progressions in workload may be deemed more tolerable by participants when compared to the Cooper test. The advantages of the UMTT are that it (i) is scalable, (ii) provides an estimated maximal aerobic speed (MAS), and (iii) provides an estimate of the VO_2max based on the stage achieved ($r = 0.98$) [10]. If the goal of testing is to provide an estimate of MAS and VO_2max, then the utility of the test is clear. However, using these parameters for exercise prescription suffers the same critiques as for the Cooper test, coupled with the fact that the test lacks the specificity needed for field sports.

The final linear continuous test within this grouping is the 3-minute all-out test for running (3MT) [11,12]. The 3MT deviates substantially from the other tests in that (i) it requires *all-out* effort from the beginning of the test, (ii) provides both a submaximal threshold and maximal sprinting speed (iii) has had VO_2 measured directly during testing allowing for both VO_2max and VO_2 kinetics assessment, and (iv) can be accurately used for programme prescription, training and monitoring [19,36–39].

The concept of the 3MT originated from cycling research originally based on the hyperbolic relationship between power output and the sustainable duration of that workload (see Figure 1) [40–42]. The asymptote of this hyperbolic relationship was referred to as 'critical power' (CP) whereby work above CP was limited by the fatigability constant (termed W-prime [W']) which represented a finite energy reserve available to an athlete above CP. Work below CP on the other hand could be sustained for substantially longer periods of time [43,44]. The successful implementation of this relationship however required multiple testing bouts (i.e. n = 4-7 bouts), each lasting between 2-15 minutes in order to accurately parametrise the mathematical relationship [45]. Although attractive, the feasibility was questionable. Later it was discovered that the same parameters, CP and W', could be derived from an all-out test lasting 3-minutes (i.e. 3MT) whereby the end-phase power (i.e. final 30-seconds) was equivalent to CP, and the area under the curve but above CP, was equivalent to W' [42]. Furthermore, the VO2max values attained towards the end of the test were consistent with the VO2max achieved in a ramp test [42]. Using these principles, the running version of the test was first implemented in 2012 [11] whereby the running equivalents could be derived in the same all-out manner (i.e. CP is replaced by critical speed [CS] which is the end-stage speed, and W' is replaced by D' as the fatigability constant).
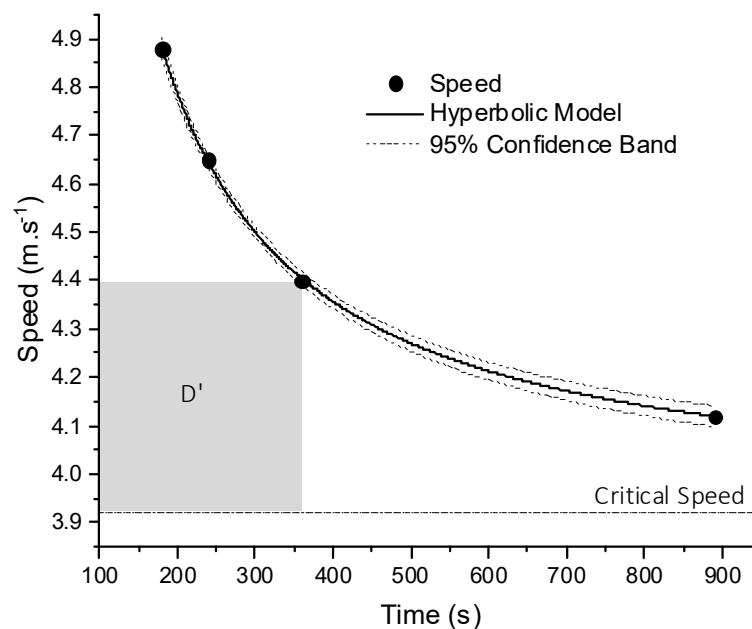

Figure 1: Hyperbolic representations of the speed-time relationship for running. CS is the asymptote of the speed-time relationship, and D' is the curvature constant.

From a physiological perspective, the CS is considered as a 'gold standard method' for the assessment of the critical metabolic threshold that separates heavy- from severe-intensity exercise, and typically occurs at a relative intensity of ~70-90% VO2max depending on training status [9,46–48]. The ability of CS to delineate these intensity

domains is one of its primary strengths on the basis that the metabolic responses, and therefore the tolerable durations (t_{lim}) across domains, is vastly different. Exercise below CS is usually tolerable for up to approximately 3 hours (depending on proximity of work rate to CS), whereas exercise above CS is highly predictable when CS and D′ are combined   [49]. Although other thresholds exist that can delineate the heavy- from the severe-intensity domain (e.g. respiratory compensation point, maximal lactate steady state, etc.), these are dependent on the completion of a GXT with respiratory gas analysis or lactate sampling (Jamnick et al., 2018), whereas the 3MT offers a short-duration, non-invasive alternative. The characteristic speed-time curve of the running 3MT was later mathematised using a bi-exponential model which provided unique parameters, such as maximal 3MT sprint speed (MS_{3MT}), fatigue index (FI%), and the speed decay time constant (τ_d), that offered additional insights into the test performance of participants [50]. Furthermore, parameters from the 3MT can be used to predict the t_{lim} within the severe-intensity domain, and thereby provide customisable training intervals, via the following two equations:

$$S = D'\%/t_{lim} + CS \tag{1}$$

$$t_{lim} = D'\%/(S - CS) \tag{2}$$

$$t_{lim} = (D - D'\%)/CS \tag{3}$$

where S is the running speed (m.s^{-1}), D′% is the proportion of D′ depleted within a given interval, t_{lim} is the interval duration (sec) and CS is the critical speed (m.s^{-1}). For example, if an athlete with a CS of 3.5 m.s^{-1} and a D′ of 130 m, would like to complete 4 bouts of high-intensity exercise, with each bout lasting 2-minutes, then as per Equation 1, the athlete would need to run each bout at a speed of 4.26 m.s^{-1} [S = 0.7 · 130 m/120 s + 3.5 m.s^{-1} = 4.26 m.s^{-1}]. Such interval prescription is based on experimental evidence that has shown that a D′ depletion of 70% per bout, coupled with a 1:2 work-to-rest ratio, would allow for the successful completion of 4 bouts [37]. Towards the end of the 4th bout the athlete would likely achieve their maximum heart rate (HR_{max}) coupled with an RPE of at least 18. The same principles can be applied to Equation 2 where the coach/athlete is free to choose a desired running speed above CS, with the exception that an interval duration (sec) is provided. If all other parameters from the previous example are kept constant (e.g. 70% depletion of D′ and CS of 3.5 m.s^{-1}), and if a desired interval speed of 5 m.s^{-1} is chosen, then the calculated interval duration is 61 sec per bout). Again, based on previous evidence, the athlete would be able to complete 4 bouts before reaching HR_{max} and exercise cessation. Other variations such as 5 x 60% or 3x 80% depletion schemes intervals with 1:1 and 1:1.5 work to rest ratios have also been successfully utilised (Pettitt,

2016, Clark et al., 2013). Presently, few other tests, if any, can provide such intervals with the same fidelity.

To support coaches/practitioners with the interpretation of parameters derived from the linear 3MT (e.g. CS, D', MS_{3MT}, FI% etc), normative data tables have been created for *rugby* [51], *athletic*, and *non-athletic* individuals [52]. For example, if a rugby player successfully completed a linear 3MT and obtained a CS of 3.70 m.s^{-1} and a D' of 272m, these parameters would be classified as 'above average' and 'average' respectively. If these same values were obtained by a non-athletic (i.e. moderately trained) individual, the same parameters would be classified as 'extremely high' and 'very high' respectively. Although it is important to be mindful of the participant characteristics when interpreting normative data, such data tables may help guide coaches/practitioners obtain more meaningful insights when translating sport-specific performances.

Recent research has even linked 3MT performance with VO$_2$max by providing a regression equation with CS and gender as input parameters [53]. The typical error, coefficient of variation, and intraclass correlation coefficient between predicted and observed VO$_2$max values were 2.26 ml.kg^{-1}.min^{-1}, 4.90% and 0.864 respectively. Although the regression equation shows promise, the equation was not validated against an independent dataset which may limit the transferability/generalisability of the equation to a broader population. As per previous critiques, knowledge of VO$_2$max may be useful for approximating and interpreting CRF but should not be used for standardising exercise prescription.

Although the 3MT has many strengths, some of the limitations must also be mentioned. Firstly, the 3MT is an all-out test and its use is therefore likely limited to those that are at least moderately fit. Secondly, the motivation to complete the test at full effort is a factor that would affect test reliability as pacing would be a risk for test invalidation. Thirdly, as discussed here, the 3MT is a linear all-out test and therefore the transferability to field sports could be questioned. In light of this however, it is pertinent to mention that parameters from the linear 3MT have been successfully implemented in training interventions focused on soccer players [36] and moderately trained individuals [37], showing significant improvements in CS (+ 0.46-0.62 m.s^{-1}) and VO$_2$max (+ 3.81-5.19 ml.kg^{-1}.min^{-1}).

Nonetheless, all continuous linear tests ultimately lack the 'specificity' often sought by coaches for evaluating CRF in a more sport-specific manner (i.e. incorporating directional change). The assimilation of directional changes for field tests is therefore an important component in the testing regimen. Whether or not the linear 3MT could be modified to incorporate directional changes, and whether the speed profile of all-out shuttle running was similar to that of linear running was previously unknown. Both aspects were however recently answered in a series of papers (Kramer *et al.*, 2018; Kramer, Du Randt, *et al.*, 2019b, 2019a). The evidence suggests that all-out linear and shuttle running (i.e. 3MT concept) tend to be energetically similar (total energetic cost: p>0.05; total metabolic power: p>0.05), but sufficiently distinctive to provide test-specific CS, D' and S$_{max}$ values (p < 0.01) [54]. Coaches should however be cognisant of differences in shuttle testing methodologies such as *intermittent* vs. *continuous* protocols as these

would provide meaningful differences in parameter estimates. The following sections will therefore unpack both continuous and intermittent shuttle-based tests typically used for the assessment of CRF in field-sports.

3. Intermittent shuttle running tests

The intermittent nature of field-based ball sports (e.g. hockey, soccer, rugby etc.) necessitated the development of an intermittent test that evaluated both the ability of players to repeatedly perform intense exercise, and the ability to recover from intensive exercise [15,55]. This led to the development of the Yo-Yo intermittent recovery test level 1 (YYIR1) [55].The end-stage speed attained during a YYIR1 (sYYIR1) test usually eclipses sVO_2max [21], although this is dependent on the stage gradation and duration of the GXT. The intermittent nature of the test allows for substantial metabolic recovery kinetics to take effect which in turn allows for the attainment of higher shuttle stages than would not have been possible during a continuous test (e.g. MSR). Additionally, considering that recovery of W'/D' appears to be $CP/CS/VO_2max$ dependant, this perhaps introduces further variability in the results of intermittent tests between individuals (Chorley et al., 2020).

The relationship between the calculated VO_2max from the YYIR1 and the VO_2max determined via a GXT is usually moderate at best (r = 0.47-0.57) [56]. A possible limitation is that the YYIR1 is often used to evaluate changes in performance outcomes (e.g. distance, level, or predicted VO_2max) as a function of some intervention, rather than using actual metrics from the test to prescribe training. Completing a test without using parameters from that test to guide exercise prescription seems somewhat counterintuitive on the basis that players are evaluated for changes in performance (e.g. shuttle level, total distance, VO_2max), without necessarily knowing the mitigating circumstances under which changes in performance may have accrued. Stated differently, it is entirely possible that changes in bioenergetics may have occurred without commensurate changes in VO_2max (or shuttle level by extension); although more research would be needed.

A further limitation relates to the fact that exercise prescription, using sYYIR1, fails to standardise the metabolic response across athletes. The lack of intensity standardisation is especially pertinent on the basis that sYYIR1 is usually substantially above sVO_2max. Exercise prescription using sYYIR1 is therefore problematic on two fronts: (i) using purely sYYIR1 would imply exercise within the extreme intensity domain which would be sustainable for less than 2-minutes depending on the proximity of sYYIR1 to sVO_2max, or (ii) adjusting sYYIR1 to be below sVO_2max, where exercise would likely fall within the severe intensity domain, would require knowledge of sVO_2max. The derivation of sVO_2max would necessitate the completion of either a laboratory- or field-based assessment to derive this additional parameter, thereby negating the purpose of the YYIR1 test. Without knowledge of submaximal intensity anchor points, exercise standardisation is impossible, and it is therefore highly unlikely that parameters from the YYIR1 could be reliably used for exercise prescription.

The evolution of intermittent shuttle-based testing led to the development of the 30-15 intermittent fitness test (IFT) which first appeared in print in 2008 [14]. The IFT may be useful for tracking changes in athlete performances with respect to parameters associated with the IFT. As with other field-based tests of this nature, there are strong correlations between IFT achievement and VO_2max (r = 0.58) as well as certain match-play parameters (e.g. high-intensity running) [14].

In many ways the IFT suffers from the same predilections as the YYIR1 in that the end-stage velocity associated with the IFT (i.e. vIFT) falls above sVO_2max [21]. Such supra-threshold running speeds would again make it difficult to standardise training adaptations based on athlete-specific metabolic responses to the proposed intensity. Without knowing the proximity of vIFT to any maximal or submaximal thresholds would preclude any comprehension of the intensity domain within which the athlete may be training, and therefore the metabolic flux would be more difficult to determine and control for. Although VO_2max can be indirectly estimated from the IFT, it should already be clear that doing so is problematic based on the over- or under-estimations associated with the regression equations. Shuttle running is clearly different to linear running, and intermittent protocols elicit different responses compared to continuous protocols, thereby bringing into question the interchangeability of parameters. To date, no study has directly evaluated the end-exercise VO_2 associated with the IFT.

To overcome some of the limitations of the intermittent shuttle tests mentioned, some researchers aimed to utilise the principles of the hyperbolic speed-time relationship (akin to the hyperbolic power-time relationship mentioned previously), such that an intermittent critical speed (iCS) and intermittent D' (iD') could be calculated (see Figure 1 for a representative example) [13,57,58]. Although the method is theoretically sound, the limitation here is the number of trials that need to be completed since this affects the robustness of the mathematical fit. Based on the original research it is recommended that between 4-to-7 trials be used to maximise the accuracy of the fit between the data and the model and for standard error to fall within 5% for CP/CS and 10% for W'/D' [49]. Much of the research on the intermittent version of the test however only used 3 trials (Berthoin *et al.*, 2006; Buchheit *et al.*, 2008; Fukuda *et al.*, 2011, 2012).

There is inevitably a convenience-accuracy trade-off using the aforementioned approach. On the one hand, to have a more accurate model fit, and more robust model parameters, a greater number of trials are required. However, the implication is that the participant would need to complete a greater number of trials over several days, which in most cases is not practically viable. Therefore, although the approach is theoretically attractive, the apparent lack of accuracy coupled with the inconvenience of repeated trials, has invariably reduced the popularity of the approach on a larger scale.

Clearly, coaches must select a test to match the requirements of the evaluation, such as (i) the determination of a physiological threshold (e.g. VO_2max, CS etc.), (ii) determining a maximal running speed (e.g. MSS, MS_{3MT}), and/or (iii) to guide training. Thus far, arguably only the linear 3MT manages to tick *all* the requisite boxes. The

following section will explore the utility of various continuous shuttle running tests in the context of the abovementioned requirements.

4. **Continuous shuttle running tests**

One of the first shuttle running tests to evaluate VO_2max was devised by Léger and Lamberts (1982) which utilised stage gradations of 0.5 $km.hr^{-1}$ every 2-minutes. This same test was later modified to incorporate the same stage gradations (i.e. 0.5 $km.hr^{-1}$) but using 1-minute stages [60]. Researchers have found significant correlations between shuttle level achieved during the test and VO_2max obtained in the laboratory depending on the group being analysed (r = 0.51-0.95) [17,22,61,62]. Consequently, regression equations have been developed to indirectly estimate VO_2max from the MSR which allows for a higher testing density (i.e. more athletes can be evaluated simultaneously), low learning curve (i.e. easy to follow instructions across numerous age cohorts), lower cost implications (i.e. minimal equipment, higher participant : expert ratios) and faster turn-around times (i.e. more tests can be conducted over a given season). These reasons may account for the broad appeal and successful implementation of the test on a global scale, even now, more than 30 years after its founding.

A potential draw-back of the MSR is that the actual VO_2 cost associated with the test has not been directly measured. Early research compared the VO_2 from the end-stage of the MSR to that obtained from a GXT using a backwards extrapolation method from

the O_2 recovery curve and found high correlations (n=11; r=0.90-0.92) [16]. Despite this,

it is important to mention that more recent research on shuttle running has shown that higher VO_2 requirements are exhibited compared to linear running at the same relative velocities [26]. For example, treadmill running at just 8 $km.hr^{-1}$ exerts a VO_2 requirement of 27.3 ± 1.5 $ml.kg^{-1}.min^{-1}$, whereas incorporating just 13 $turns.min^{-1}$ at the same velocity

($\approx$ 1 turn every 10-m), substantially increases the VO_2 requirement to 38.4 ± 1.9

$ml.kg^{-1}.min^{-1}$ (i.e. an increase of 11.1 $ml.kg^{-1}.min^{-1}$ or 41%).

One of the primary outcomes of the MSR is an estimated VO_2max based on the end-stage speed. Knowledge of the VO_2max is useful for predicting and monitoring physical performance, assessing talent, and possibly designing training programmes [5]; although evidence for the latter is severely limited. A recent review by Jamnick et al. (2020), highlighted the fact that prescribing intensities as a percentage of VO_2max is not a valid method for exercise prescription on the basis that this leads to substantial heterogenous   homeostatic responses in and across individuals.

To address the limitations imposed by the MSR, additional continuous shuttle tests were developed such as the 1.2-km shuttle run test (1.2SRT) and the shuttle 3MT. The 1.2SRT was developed in 2012 and consists of running two 20-m shuttles, followed by two 40-m shuttles and finally two 60-m shuttles in series, which are then repeated in five continuous sets for a total of 1200-m [63]. The primary outcomes measure of the 1.2SRT is a maximal running speed (MRS) which has shown high correlation to the MRS obtained

during the IFT ($r = 0.86$); although the MRS from the 1.2$_{SRT}$ is lower than vIFT [63,64]. The 1.2$_{SRT}$ has also been linked to on-field performances for rugby-league players [65], although more research is required. The benefits of the 1.2$_{SRT}$ however are the (i) relatively short duration of 4-7 minutes and (ii) scaling for mass testing. It is nevertheless pertinent to mention that drawbacks do exist in that (i) the test has not been validated against direct/indirect VO$_2$max measures, (ii) MRS is calculated as the average speed of the entire test, and that (iii) the utility of MRS data is limited for training purposes. It is therefore not known whether MRS occurs above or below sVO$_2$max.

To encompass a more 'sport specific' approach, the linear 3MT was modified to incorporate directional changes by using 25-m, 30-m and 50-m shuttle distances (Kramer *et al.*, 2018, 2019; Saari et al., 2019). It is clear from Figure 2 that the speed-time profile of the modified AOST is similar to that of the linear 3MT. Consequently, the modified AOST provides similar parameters in that a shuttle specific CS, D', τ$_d$ and maximal all-out shuttle speed (MS$_{AOST}$), which when combined with CS can provide a potentially more ecologically alternative parameter to the popular anaerobic speed reserve concept as well [66].
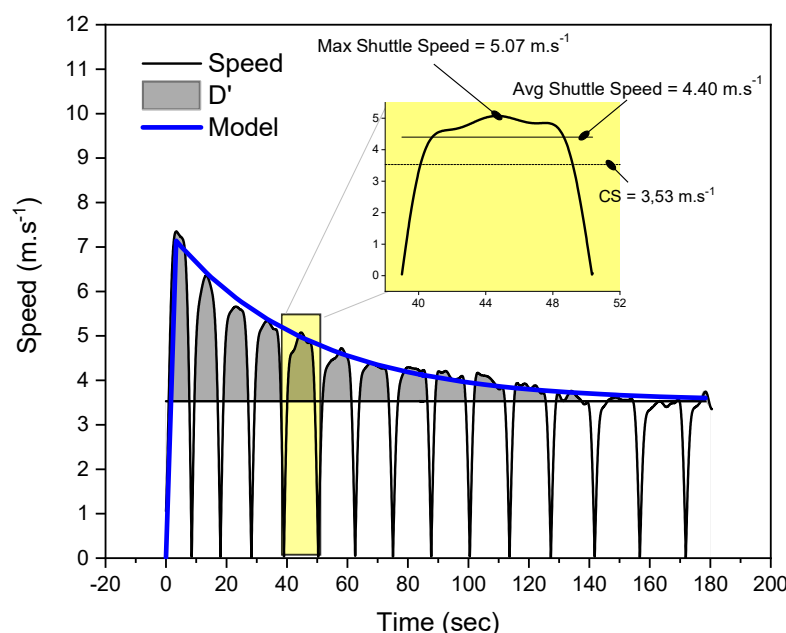

Figure 2: Representation of a 50-m shuttle-based 3MT. Black horizontal line shows the critical speed (CS); blue line indicates the bi-exponential method used to model changes in peak shuttle speed throughout the test; grey area indicates the D'. Inset shows the speed profile of a single shuttle coupled with key parameters such as the maximal shuttle speed, average shuttle speed (solid horizontal line) and CS (dotted horizontal line). Adapted from [20]

One of the only studies to investigate the VO$_2$ response during the 3MT (both linear and shuttle versions), was conducted by Kramer *et al.* (2018) which showed clear similarities across 3MT testing modalities. The data from the original manuscript was

used to evaluate whether the $\dot{V}O_2$max values derived from each version of the test were statistically equivalent to each other and to that derived from a laboratory-based GXT (see Figure 3).
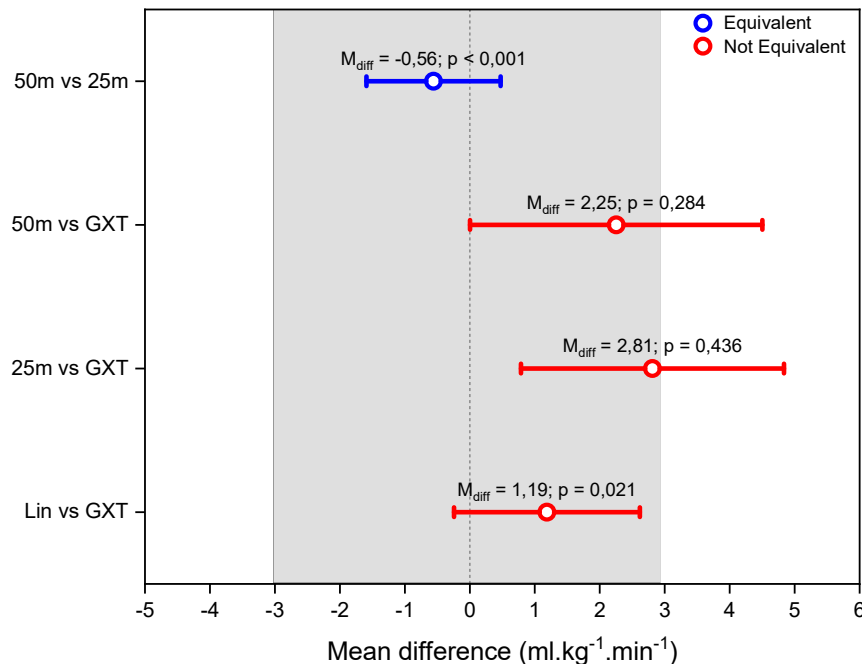

Figure 3: Assessing $\dot{V}O_2$max equivalence across laboratory and field tests. Equivalence bounds ($\pm$ 3 ml.kg^{-1}.min^{-1} from zero) were established from available literature and evaluated using the

TOST procedure [67]. The alpha (α) is adjusted for multiple comparisons using a Bonferroni cor-

rection (α = 0.05/4 = 0.0125).

In terms of the mean difference in $\dot{V}O_2$ between the linear 3MT and GXT as well as between shuttle 3MTs and GXT, after correcting for multiple comparisons, the equivalence results were largely inconclusive (p>α). We can therefore neither reject the null hypothesis, nor the smallest effect size of interest of $\pm$ 3.0 ml.kg^{-1}.min^{-1}. The effect size is relatively small, but more data are needed to draw a conclusion (given our desired error rate of 5%). The mean $\dot{V}O_2$ differences between 25-m and 50-m 3MT are however statistically equivalent.

In terms of using the parameters from the 3MT and AOST, Kramer *et al*. (2019), investigated whether the t_{lim} of an 800-m shuttle could be predicted. The t_{lim} accurately predicted the 25-m AOST (r = 0.93, SEE = 6.60 sec) when athletes were instructed to complete the interval at their fastest self-selected pace [20]. It must however be stated that it is presently unknown whether the use of customised speed or time intervals, without implementing a speed correction, would work as effectively for shuttle running

as for linear running. The energetics and neuromuscular loading characteristics of shuttle running are substantially different from linear running, and it is therefore theorised that the fatigue mechanisms inherent to shuttle running would therefore play a role in shuttle running prescription using the methods previously mentioned.

One of the strengths of the 3MT in general is that the parameters from the test can not only be used to assess and monitor robust and ecologically valid parameters specific to team sports, but can also be used to prescribe training intervals that are either based on time (see Equation 1) or speed (see equation 2), or distance (see Equation 3) requirements [37]. In addition to this, given the problematic nature of using absolute/fixed speed thresholds for match/training load monitoring across players (i.e., 5.5 m.s^{-1} for high speed running; 7 m.s^{-1} for sprinting), the results of 3MT can arguably provide more specific data to more specifically individualise these processes.

The overall results of the speeds derived from the various testing methodologies discussed in the text can be contextualised in Figure 4.
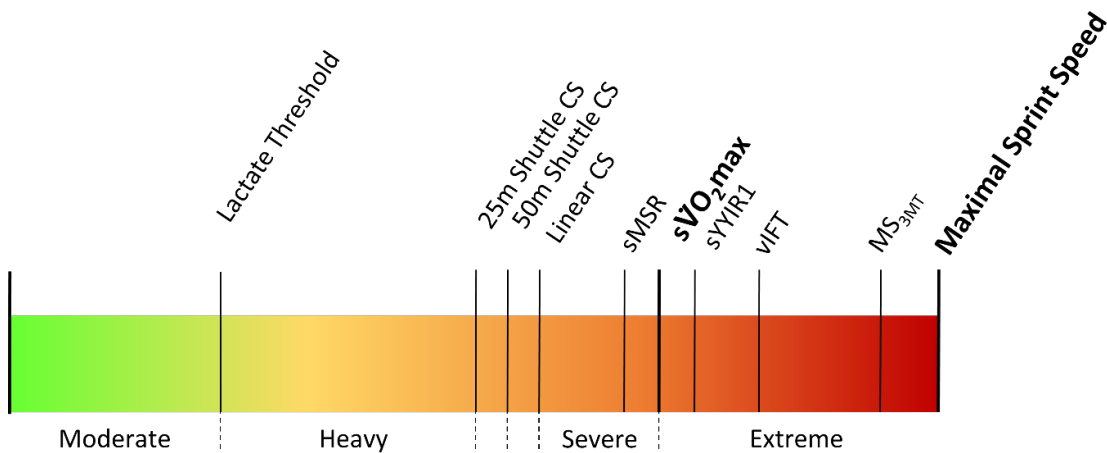

Figure 4: Continuum of running speeds and intensity domains. sLT = speed associated with the lactate threshold; CS = critical speed; sMSR = end-stage speed of the multistage shuttle run; sV̇O$_2$max = speed evoking maximal oxygen uptake; sYYIR1 = speed associated with the yo-yo intermittent recovery test level 1; vIFT = velocity associated with the intermittent fitness test; MS$_{3MT}$ = maximal speed attained during the 3MT. Intensity domain classification ranges: moderate ($\leq$ LT); heavy (> LT, but $\leq$ CS); severe (> CS, but $\leq$ sV̇O$_2$max); extreme (> sV̇O$_2$max). Adapted from [21].

As practitioners it is important to be wary of the benefits and costs of each testing method. Shuttle-based tests (e.g. MSR, YYIR and IFT), which are appropriate for team-based sports, should not necessarily be used to estimate VO$_2$max, but rather focus on monitoring performance improvement and evaluating intervention effectiveness. Prescribing exercise sessions as a function of VO$_2$max (or sVO$_2$max) is not appropriate due to the lack of physiological standardisation which may be more desirable in team

sports. As this review has shown, the AOST represents a viable alternative. The CS derived from the shuttle 3MT would be synonymous with a sub-maximal threshold intensity that other tests cannot provide. Moreover, there are additional parameters such as MS_{3MT}, FI%, D', and τ_d that can provide insights into not only the performance characteristics of players, but also exercise prescription and monitoring that can be individualised, yet standardised. Such data can be easily obtained via GPS units or stopwatches which can be found in most fitness facilities. Furthermore, multiple players can be tested at the same time allowing effective and efficient testing processes to be utilised in the time-constrained environments of applied practice.

5. Conclusions

Field-tests are multifactorial and can provide a multitude of performance-related parameters. Coaches should be aware of the advantages and disadvantages of each test and should be mindful of the purpose, outcomes, and utility of each test and how this may both inform and guide training prescription for performance enhancement

References

1.   Paraskevas G, Hadjicharalambous M. Aerobic Fitness of Starter and Non-Starter Soccer Players in the Champion's League. J Hum Kinet. 2018;61: 99–108. doi:10.1515/hukin-2017-0135

2.   Bishop DJ, Girard O. Determinants of team-Sport performance: Implications for altitude training by team-Sport athletes. Br J Sports Med. 2013;47. doi:10.1136/bjsports-2013-092950

3.   Smith VG, Jayaraman RC. Physiological Fitness Profile of NCAA Division I Female Field Hockey Players. ARC J Res Sport Med. 2019;4: 17–24.

4.   Glassbrook DJ, Doyle TLA, Alderson JA, Fuller JT. The Demands of Professional Rugby League Match-Play: a Meta-analysis. Sport Med - Open. 2019;5. doi:10.1186/s40798-019-0197-9

5.   Slimani M, Znazen H, Miarka B, Bragazzi NL. Maximum Oxygen Uptake of Male Soccer Players According to their Competitive Level , Playing Position and Age Group : Implication from a Network Meta-Analysis by. J Hum Kinet. 2019;66: 233–245. doi:10.2478/hukin-2018-0060

6.   Hale T. History of developments in sport and exercise physiology: A. V. Hill, maximal oxygen uptake, and oxygen debt. Journal of Sports Sciences. 2008. doi:10.1080/02640410701701016

7.   Beltz NM, Gibson AL, Janot JM, Kravitz L, Mermier CM, Dalleck LC. Graded Exercise Testing Protocols for the Determination of VO2max: Historical Perspectives, Progress, and Future Considerations. J Sports Med. 2016;2016: 1–12.

doi:10.1155/2016/3968393

8.   Hill A V., Lupton H. Muscular exercise, lactic acid, and the supply and utilization of oxygen. Q J Medica. 1923;16: 135–171. doi:10.1093/qjmed/os-16.62.135

9.   Jamnick NA, Pettitt RW, Granata C, Pyne DB, Bishop DJ. An Examination and Critique of Current Methods to Determine Exercise Intensity. Sport Med. 2020. doi:10.1007/s40279-020-01322-8

10.   Léger L, Boucher R. An indirect continuous running multistage field test: the Université de Montréal track test. Can J Appl Sport Sci. 1980;5: 77–84.

11.   Pettitt RW, Jamnick N, Clark IE. 3-min All-out Exercise Test for Running. Int J Sports Med. 2012;33: 426–431.

12.   Broxterman RM, Ade CJ, Poole DC, Harms CA, Barstow TJ. A single test for the determination of parameters of the speed-time relationship for running. Respir Physiol Neurobiol. 2013;185: 380–385. doi:10.1016/j.resp.2012.08.024

13.   Fukuda DH, Smith AE, Kendall KL, Cramer JT, Stout JR. The determination of critical rest interval from the intermittent critical velocity test in club-level collegiate hockey and rugby players. J Strength Cond Res. 2011;25: 889–895.

14.   Buchheit M. The 30-15 Intermittent Fitness Test: Accuracy for Individualizing Interval Training of Young Intermittent Sport Players. J Strength Cond Res. 2008;22: 365–374.

15.   Krustrup P, Mohr M, Amstrup T, Rysgaard T, Johansen J, Steensberg A, et al. The Yo-Yo Intermittent Recovery Test: Physiological Response, Reliability, and Validity. Med Sci Sports Exerc. 2003;35: 697–705. doi:10.1249/01.MSS.0000058441.94520.32

16.   Léger L, Seliger V, Brassard L. Backward extrapolation of Vo2max values from the O2 recovery curve. Med Sci Sports Exerc. 1980;12: 24–27.

17.   Léger L, Gadoury C. Validity of the 20 m shuttle run test with 1 min stages to predict VO2max in adults. Can J Sport Sci. 1989;14: 21–26.

18.   Léger LA, Lambert J. A Maximal Multistage 20-m Shuttle Run Test to Predict VO2max*. Eur J Appl Physiol. 1982;49: 1–12.

19.   Kramer M, Du Randt R, Watson M, Pettitt RW. Oxygen uptake kinetics and speed-time correlates of modified 3-minute all-out shuttle running in soccer players. PLoS One. 2018;13: e0201389.

20.   Kramer M, Watson M, Du Randt R, Pettitt RW. Critical speed as a measure of aerobic fitness for male rugby union players. Int J Sports Physiol Perform. 2019;8: 518–524.

21.   Clarke R, Dobson A, Hughes J, Kingdom U, Sciences S, Kingdom U. Metabolic Conditioning: Field Tests to Determine a Training Velocity. Strength Cond J. 2016;38: 38–47.

22.   Mayorga-Vega D, Aguilar-Soto P, Viciana J. Criterion-related validity of the 20-m shuttle run test for estimating cardiorespiratory fitness: A meta-analysis. J Sport Sci Med. 2015;14: 536–547.

23.   Martínez-Lagunas V, Hartmann U. Validity of the Yo-Yo Intermittent Recovery Test Level 1 for direct measurement or indirect estimation of maximal oxygen uptake in female soccer players. Int J Sports Physiol Perform. 2014;9: 825–831. doi:10.1123/ijspp.2013-0313

24.   Grant S, Corbett K, Amjad AM, Wilson J, Aitchison T. A comparison of methods of predicting maximum oxygen uptake. Br J Sports Med. 1995;29: 147–152. doi:10.1136/bjsm.29.3.147

25.   Panasci M, Romuald L, La Torre A, Bonato M, Assadi H. Physiological responses during intermittent running exercise differ between outdoor and treadmill running. Appl Physiol Nutr Metab. 2019;42: 1–26.

26.   Hatamoto Y, Yamada Y, Sagayama H, Higaki Y, Kiyonaga A, Tanaka H. The relationship between running velocity and the energy cost of turning during running. PLoS One. 2014;9: 1–8. doi:10.1371/journal.pone.0081850

27.   Zamparo P, Zadro I, Lazzer S, Beato M, Sepulcri L. Energetics of shuttle runs: The effects of distance and change of direction. Int J Sports Physiol Perform. 2014;9: 1033–1039. doi:10.1123/ijspp.2013-0258

28.   Tønnessen E, Hem E, Leirstein S, Haugen T, Seiler S. Maximal Aerobic Power Characteristics of Male Professional Soccer Players, 1989 – 2012. Int J Sports Physiol Perform. 2013;8: 323–329. doi:10.1123/ijspp.8.3.323

29.　Stølen T, Chamari K, Castagna C, Wisløff U. Physiology of soccer: an update. Sport Med. 2005;35: 501–536.

30.　Duthie G, Pyne D, Hooper S. Applied Physiology and Game Analysis of Rugby Union. Sport Med. 2003;33: 973–991. doi:10.2165/00007256-200333130-00003

31.　Flouris AD, Metsios GS, Koutedakis Y. Enhancing the efficacy of the 20 m multistage shuttle run test. Br J Sports Med. 2005;39: 166–170. doi:10.1136/bjsm.2004.012500

32.　Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciana J. Criterion-related validity of the distance- and time-based walk/run field tests for estimating cardiorespiratory fitness: A systematic review and meta-analysis. PLoS One. 2016;11: 1–24. doi:10.1371/journal.pone.0151671

33.　Mulazimoglu O, Afyon YA, Keser AY. A comparison of three different shuttle run tests for the prediction of Vo2max. Int J Phys Educ Sport Heal. 2018;5: 112–117.

34.　Póvoas S, Krustrup P, Castagna C. Submaximal field testing validity for aerobic fitness assessment in recreational football. Scand J Med Sci Sport. 2020;30: 680–689. doi:10.1111/sms.13606

35.　Cooper KH. A Means of Assessing Maximal Oxygen Intake. Jama. 1968;203: 201. doi:10.1001/jama.1968.03140030033008

36.　Clark IE, West BM, Reynolds SK, Murray SR, Pettitt RW. Applying the critical velocity model for an off-season interval training program. J Strength Cond Res. 2013;27: 3335–3341.

37.　Thomas EJ, Pettitt RW, Kramer M. High-Intensity Interval Training Prescribed Within the Secondary Severe-Intensity Domain Improves Critical Speed But Not Finite Distance Capacity. J Sci Sport Exerc. 2020;2: 154–166. doi:10.1007/s42978-020-00053-6

38.　Dicks ND, Joe T V, Hackney KJ, Pettitt RW. Validity of Critical Velocity Concept for Weighted Sprinting Performance. Int J Exerc Sci. 2018/08/28. 2018;11: 900–909.

39.　de Aguiar RA, Salvador AF, Penteado R, Faraco HC, Pettitt RW, Caputo F. Reliability and validity of the 3-min all-out running test. Rev Bras Ciências do Esporte. 2018;40: 288–294. doi:10.1016/j.rbce.2018.02.003

40.　Moritani T, Nagata A, Devries HA, Muro M. Critical power as a measure of physical work capacity and anaerobic threshold. Ergonomics. 1981;24: 339–350.

41.　Jenkins DG, Quigley BM. The influence of high-intensity exercise training on the W-T relationship. Med Sci Sports Exerc. 1993;25: 275–282.

42.　Burnley M, Doust JH, Vanhatalo A. A 3-min all-out test to determine peak oxygen uptake and the maximal steady state. Med Sci Sports Exerc. 2006;38: 1995–2003. doi:10.1249/01.mss.0000232024.06114.a6

43.　Jones AM, Wilkerson DP, Dimenna F, Fulford J, Poole DC. Muscle metabolic responses to exercise above and below the " critical power " assessed using 31 P-MRS. Am J Physiol Regul Integr Comp Physiol. 2008;294: 585–593. doi:10.1152/ajpregu.00731.2007.

44.　Vanhatalo A, Doust JH, Burnley M. A 3-min All-out Cycling Test Is Sensitive to a Change in Critical Power. Med Sci Sports Exerc. 2008;40: 1693–1699. doi:10.1249/MSS.0b013e318177871a

45.　Muniz-Pumares D, Karsten B, Triska C, Glaister M. Methodological approaches and related challenges associated with the determination of critical power and curvature constant. J Strength Cond Res. 2019;32: 584–596.

46.　Jones AM, Vanhatalo A. The 'Critical Power' Concept: Applications to Sports Performance with a Focus on Intermittent High-Intensity Exercise. Sport Med. 2017;47: 65–78. doi:10.1007/s40279-017-0688-0

47.　Craig JC, Vanhatalo A, Burnley M, Jones AM, Poole DC. Critical Power: Possibly the Most Important Fatigue Threshold in Exercise Physiology. Zoladz A, editor. Muscle and Exercise Physiology. London, UK: Elsevier Inc.; 2019. doi:10.1016/B978-0-12-814593-7.00008-6

48.　Jones AM, Burnley M, Black MI, Poole DC, Vanhatalo A. The maximal metabolic steady state: redefining the ' gold standard .' Physiol Rep. 2019;7: 1–16. doi:10.14814/phy2.14098

49.　Poole DC, Jones AM. Oxygen Uptake Kinetics. Compr Physiol. 2012;2: 933–996. doi:10.1002/cphy.c100072

50.   Kramer M, Du Randt R, Watson M, Pettitt RW. Bi-exponential modeling derives novel parameters for the critical speed concept. Physiol Rep. 2019;7: e13993. doi:10.14814/phy2.13993

51.   Kramer M, Clark IE, Jamnick N, Strom C, Pettitt RW. Normative data for critical speed and D' for high-level male rugby players. J Strength Cond Res. 2018;32: 783–789.

52.   Kramer M, Thomas EJ, Pettitt RW. Critical speed and finite distance capacity: norms for athletic and non-athletic groups. Eur J Appl Physiol. 2020;120: 861–872. doi:10.1007/s00421-020-04325-5

53.   Dexheimer JD, Brinson SJ, Pettitt RW, Schroeder ET, Sawyer BJ, Jo E. Predicting Maximal Oxygen Uptake Using the 3-Minute All-out Test in High-Intensity Functional Training Athletes. Sports. 2020;8: 155–168.

54.   Kramer M, Du Randt R, Watson M, Pettitt RW. Energetics of male field-sport athletes during the 3-min all-out test for linear and shuttle-based running. Eur J Appl Physiol. 2019;119: 477–486. doi:10.1007/s00421-018-4047-0

55.   Bangsbo J. Fitness Training in Football: A Scientific Approach. Bagsværd, Denmark: HO≼Storm: August Krogh Institute, University of Copenhagen; 1994.

56.   Metaxas TI, Koutlianos NA, Kouidi EJ, Deligiannis AP. Comparative study of field and laboratory tests for the evaluation of aerobic capacity in soccer players. J Strength Cond Res. 2005;19: 79–84.

57.   Buchheit M, Laursen PB, Millet GP, Pactat F, Ahmaidi S. Predicting Intermittent Running Performance: Critical Velocity versus Endurance Index. Int J Sports Med. 2008;29: 307–315. doi:10.1055/s-2007-965357

58.   Berthoin S, Baquet G, Dupont G, Van Praagh E. Critical velocity during continuous and intermittent exercises in children. Eur J Appl Physiol. 2006;98: 132–138. doi:10.1007/s00421-006-0253-2

59.   Fukuda DH, Smith AE, Kendall KL, Hetrick RP, Hames RL, Cramer JT, et al. The reliability of the intermittent critical velocity test and assessment of critical rest interval in men and women. Eur J Appl Physiol. 2012;112: 1197–1205. doi:10.1007/s00421-011-2076-z

60.   Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. J Sports Sci. 1988;6: 93–101. doi:10.1080/02640418808729800

61.   Paradisis GP, Zacharogiannis E, Mandila D, Smirtiotou A, Argeitaki P, Cooke CB. Multi-stage 20-m shuttle run fitness test, maximal oxygen uptake and velocity at maximal oxygen uptake. J Hum Kinet. 2014;41: 81–87. doi:10.2478/hukin-2014-0035

62.   Liu NY, Plowman SA, Looney MA. The Reliability and Validity of the 20-Meter Shuttle Test in American Students 12 to 15 Years Old. Res Q Exerc Sport. 1992;63: 360–365. doi:10.1080/02701367.1992.10608757

63.   Kelly V, Wood A. The correlation between the 30-15 intermittent fitness test and a novel test of running performance. J Aust Strength Cond. 2013;21: 91–94.

64.   Brew D, Kelly V. The reliability of the 1 . 2km shuttle run test for intermittent sport athletes . J Aust Strength Cond. 2014;22: 127–131.

65.   Kempton T, Coutts AJ. Factors affecting exercise intensity in professional rugby league match-play. J Sci Med Sport. 2016;19: 504–508. doi:10.1016/j.jsams.2015.06.008

66.   Kramer M, Thomas EJ, Van Aswegen M. Functional Speed Reserve as a Proxy for the Anaerobic Speed Reserve Using the Critical Speed Concept. Ann Appl Sport Sci. 2020;in Press: e928.

67.   Lakens D, Scheel AM, Isager PM. Equivalence Testing for Psychological Research: A Tutorial. Adv Methods Pract Psychol Sci. 2018;1: 259–269. doi:10.1177/2515245918770963