VALIDITY AND RELIABILITY OF AN ON-COURT FITNESS TEST FOR ASSESSING AND MONITORING AEROBIC FITNESS IN SQUASH

CARL JAMES,[1] FLORENCIO TENLLADO VALLEJO,[1] MELVIN KANTEBEEN,[2] AND SARO FARRA[1]

[1]*Institut Sukan Negara (National Sports Institute), Bukit Jalil Sport City, Sri Petaling, Kuala Lumpur, Malaysia; and* [2]*ConTest–Sports Performance Testing, Sport Medical Center Amsterdam, Olympiaplein, Amsterdam, Netherlands*

ABSTRACT

James, C, Tenllado Vallejo, F, Kantebeen, M, and Farra, S. Validity and reliability of an on-court fitness test for assessing and monitoring aerobic fitness in squash. *J Strength Cond Res* 33(5): 1400–1407, 2019—Current on-court assessments of aerobic fitness in squash are not designed to yield a wealth of physiological data. Moreover, tests may require complex computer equipment or involve simulated racket strokes, which are difficult to standardize at high intensities. This study investigated the validity and reliability of a squash-specific fitness test which can yield both a standalone performance score, as well as pertinent physiological markers such as $\dot{V}O_2$max, the lactate turnpoint and oxygen cost, in a sport-specific environment. Eight national squash players completed 3 tests in a counterbalanced order: an incremental laboratory treadmill test (LAB) and 2 on-court fitness tests (STs) that involved repeated shuttle runs at increasing speeds. $\dot{V}O_2$max during ST was agreeable with LAB (typical error [TE] = 3.3 $ml \cdot kg^{-1} \cdot min^{-1}$, $r = 0.79$). The mean bias between LAB and ST was 2.5 $ml \cdot kg^{-1} \cdot min^{-1}$. There were no differences in maximum heart rate, postexercise blood lactate concentration, or end of test rating of perceived exertion between LAB and ST ($p > 0.05$). The ST was highly reliable, with 74 (10) laps completed in ST1 and 75 (12) laps in ST2 (mean bias = 1 lap, TE = 3 laps, $r = 0.97$). Physiological markers were also reliable, including $\dot{V}O_2$max, (TE = 1.5 $ml \cdot kg^{-1} \cdot min^{-1}$, $r = 0.95$), the lap number at 4 $mMol^{-1}$ (TE = 4 laps, $r = 0.77$), and average $\dot{V}O_2$ across the first 4 stages (TE = 0.94 $ml \cdot kg^{-1} \cdot min^{-1}$, $r = 0.95$). We observed good agreement between LAB and ST for assessing $\dot{V}O_2$max and between both on-court trials for assessing test performance and selected physiological markers. Consequently, we recommend this test for monitoring training adaptations and prescribing individualized training in elite squash players.

KEY WORDS fitness testing, physiology, sport-specific, squash training

Address correspondence to Dr. Carl James, carlalexanderjames@gmail.com.

33(5)/1400–1407

Journal of Strength and Conditioning Research

INTRODUCTION

Squash is an intermittent, high-intensity racket sport that places great demands on aerobic metabolic pathways (9,16,22). Girard et al. (9) reported the average match intensity to be 86% of $\dot{V}O_2$max (range: 67–94%) and that 24 ± 9% of total match time was spent above 90% of $\dot{V}O_2$max. These indicated players consistently exercise above their lactate turnpoint during matches, suggesting that the routine training and monitoring of both aerobic and anaerobic metabolic pathways are critical when preparing players for competition.

In squash, the value of laboratory testing procedures that involve continuous steady-state or progressive incremental exercise to monitor changes in sport-specific performance, and pertinent physiological variables, seems limited. Typical laboratory tests of aerobic fitness fail to discriminate between the fitness of squash and nonsquash players, thereby not demonstrating construct validity (29). This likely reflects the lack of frequent changes of direction, multiple accelerations and decelerations, or regular recovery periods in laboratory tests, making these protocols less representative of the specific physiological demands of squash. Reflecting these limitations, several investigations have sought to develop sport-specific protocols to simulate match play demands (25,28) as well as measure squash-specific fitness using progressive, multistage, court-based protocols to exhaustion (8,10,18,27,29).

Existing squash-specific fitness tests demonstrate high construct validity (8,10,18,27,29). However, some of these tests require sophisticated computer software and equipment to be implemented (8,27), which precludes widespread use and testing across a number of locations. Some assessments also include racquet strokes, which are prompted by visual or auditory commands in a randomized fashion (8,10,29). Not only are these strokes difficult to standardize at higher intensities, but

their use has been criticized because they eliminate the various cues players use to anticipate their next shot (18). Therefore, tests that do not involve racket strokes or require lighting and computer software seem favorable. Reflecting these concerns, Micklewright and Papapdopoulou (18) validated a simpler, on-court running test. However, this test only provides a performance measure of finishing stage and does not yield a wealth of submaximal physiological data that could be used to individualize training prescription.

Currently, there is no validated, on-court assessment of aerobic fitness that yields a wealth of physiological data comparable with a laboratory procedure in an ecologically valid manner, while remaining simple to implement. However, the test structure of the incremental, multistage assessment of Gouttebarge et al. (10) is well suited for physiological assessment because it includes prolonged, discontinuous stages of 3–4 minutes. As such, it can provide both meaningful submaximal and maximal physiological data, as well as a sport-specific measure of performance, allowing sport scientists to understand subtle training improvements by differentiating between the various physiological mechanisms that contribute to test and sport performance. However, this test included racket strokes, which may decrease the retest reliability, a factor the study did not assess.

The aim of this study was to validate and assess the retest reliability of a modified squash-specific exercise test (10), without racket strokes, which yields both a maximal performance score, as well as pertinent physiological markers such as $\dot{V}O_2$max, the lactate turnpoint and oxygen cost, in a sport-specific environment.

Methods

Experimental Approach to the Problem

The study comprised 2 parts; first, validate the squash test against a laboratory treadmill test (LAB) and second, investigate the reliability of the squash test when completed 48 hours later. Subjects reported for testing on 3 occasions over a 5-day period. Sessions were separated by 48 hours and occurred at the same time of day. Subjects were randomly assigned to 2 groups where the testing order was counterbalanced. Four subjects completed the squash test on Monday and Wednesday, with the treadmill test on Friday, whereas the remaining subjects completed the tests in the opposite order. Validity comparisons were made between the treadmill test and the squash test in which each individual recorded the highest performance score. The existing training program continued throughout the week of the tests. All subjects had completed the squash test at least once in the previous 6 months, before the study commenced.

Subjects

Eight subjects (6 male and 2 female) volunteered for this study (± *SD* age: 20.3 [2.1] years, height: 171 [7], body mass 64.7 [6.3], and $\dot{V}O_2$max 48.8 [5] $ml \cdot kg^{-1} \cdot min^{-1}$). Subjects were nationally ranked squash players based at the National Squash Center of Malaysia and had been part of the national program for at least 2 years. They typically completed 9 training sessions per week (~11 hours) and regularly competed in international Professional Squash Association events. All subjects provided written, informed consent and the study received institutional ethical approval from the Institute Sukan Negara.

Players were instructed to rest the evening before a morning test and on the morning of an afternoon test. Players were asked to maintain their normal dietary habits for the duration of the testing week, replicate their diet for 24 hours before each test, and avoid caffeine and alcohol during this period. On each testing day, players avoided consuming sports drinks, supplements, and sports gels. Finally, players were asked to arrive hydrated, verified through urine analysis using refractometer (<1.020, Specific Gravity Refractometer Master series, Atago, USA) (24). Before all tests, players completed a wellness questionnaire, where they indicated fatigue, sleep quality, muscle soreness, stress, and mood on a 9-point scale (17).

Procedures

Treadmill Test. The LAB was performed on a motorized treadmill (Quasar treadmill; HP Cosmos, Nussdorf, Germany). Starting speed was determined from recent running performances in training, with a speed chosen that was expected to elicit steady-state blood lactate responses during the first 2 stages. Each stage was 3 minutes in duration, followed by 1-minute rest for blood sampling. Exercise intensity increased every stage by 1 $km \cdot h^{-1}$, until blood lactate concentration exceeded 6 $mMol^{-1}$ or physiological and perceptual responses indicated that they were exercising in the severe-intensity domain (6) and therefore would be unlikely to complete another 3-minute stage. At this point, speed was held constant and gradient increased by 1% every minute until volitional exhaustion. After a 5-minute active recovery (treadmill walking at 5 $km \cdot h^{-1}$), $\dot{V}O_2$max was verified as subjects ran for as long as they could during a square-wave bout of supramaximal exercise. The gradient was increased by 1% above that achieved during the final stage of the incremental test (19). Treadmill tests were deemed maximal if the verification protocol failed to elicit a change in oxygen uptake of >2 $ml \cdot kg^{-1} \cdot min$ (16,21).

Squash Test. The squash fitness test (ST) was a modification of a previously described procedure (10), after significant pilot testing and discussion with national squash coaches. Modifications included the removal of racket strokes and squash-specific movement patterns, as well as increasing shuttle distances by ~10% to account for the elimination of simulated shots. The test involved subjects completing repeated, multidirectional shuttle runs on the squash court, in accordance with audio beeps, until volitional exhaustion (Figure 1).

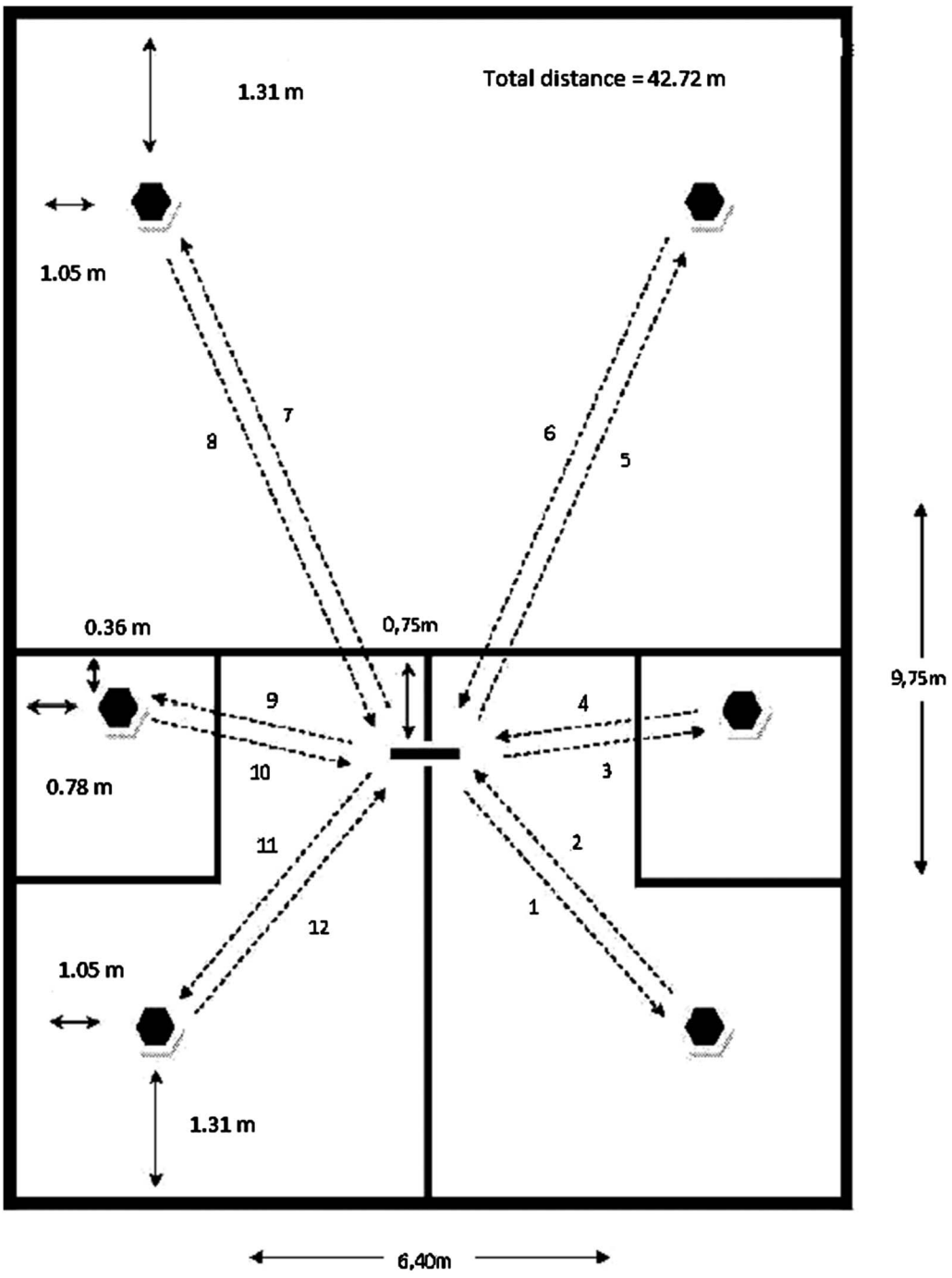

Figure 1. Layout of test. Test can be completed with or without physiological measurements.

The starting stage was chosen based on pilot testing data, in accordance with the guidelines presented in Table 1.

On average, subjects reached exhaustion within 5–8 stages. Each stage lasted between 3 and 4 minutes and speed increased by 0.19 $m \cdot s^{-1}$ between stages (Table 2). Participants ran between 6 and 10 "laps" of the court every stage, with each lap being 42.72 m. Reflecting the intermittent nature of the sport, subjects rested for 10 seconds after each lap and 30 seconds between stages. The stage that preceded each player's starting stage was completed as a standardized warm-up.

Table 1. Recommended starting stages for different standards of squash players.

Player	Starting stage	Warm-up stage
Youth female (~<15 y)	4	3
Junior female (~15–18 y)	5	4
Senior female (>18 y)	5	4
World-class female	6	5
Youth male (~<15 y)	5	4
Junior male (~15–18 y)	5	4
Senior male (>18 y)	6	5
World-class male	7	6

During the test, audio beeps indicated when subjects should be back at the "T" after each shuttle run, as well as the start and end of each lap. Subjects were told to run in accordance with the audio and to place a foot on each marker, which was closely monitored throughout. Subjects were not required to adopt squash-specific movements unless they believed it to be beneficial for changing direction and moving between markers. The test ended when the participant failed to reach the "T" at the required time for 2 consecutive laps (>1 m short), with the last successfully completed lap used as the player's score.

Data Acquisition. Subjects wore a portable metabolic cart (K5 gas analyzer; Cosmed K5, Rome, Italy) during all testing procedures to measure respiratory gases (breath by breath) and heart rate (HR). Before every test, the metabolic cart was calibrated in accordance with manufacturer's instructions and the lactate analyzer was checked against calibration standards. Capillary blood samples were collected from a fingertip for immediate analysis of blood lactate concentration using a Lactate Scout analyzer (Lactate Scout; EKF Diagnostics, Cardiff, United Kingdom). Blood lactate and rating of perceived exertion (RPE, 6–20 scale) (4) were taken at the end of every stage, during both protocols, and again at exhaustion. For the ST, the number of laps at 4 $mMol^{-1}$ was calculated by solving the polynomial regression equation for blood lactate concentration versus lap number, as a measure of the lactate turnpoint (14,23).

Statistical Analyses

All outcome variables were assessed for normality and sphericity before further analysis. Data were analyzed using SPSS (Version 22; SPSS, Inc., Chicago, IL, USA) with statistical significance set at $p \leq 0.05$. Data are presented as mean $\pm SD$. The following battery of statistics was calculated on data derived from the squash test chosen for analysis and compared against the laboratory test, which represented the criterion measure for assessing validity. Such a battery of relative and absolute validity or reliability statistics provides a more robust analysis than significance testing alone or any singular test (1,11,13). These calculations were: mean bias (MB), Pearson's correlation coefficient (r), typical error (TE) of the estimate (TEE), providing standard or TE of the predicted y-value for each x and this was also expressed as a mean coefficient of variation (TEE CV%) (12). Finally, the limits of agreement (LOA) were calculated, reflecting the range within which 95% of differences between the 2 measurement methods would be expected to fall (3). Similar absolute and relative statistics were calculated to assess the reliability of performance and physiological variables: MB, TE of the measure (TEM), calculated from the SD of the mean difference for each pair of trials using the formula $TE = SD(\text{diff})/\sqrt{2}$, Pearson's r, and LOA. Finally, paired-samples T-tests are used to identify MB between and within exercise tests. Relative validity and reliability are presented as Pearson's r, as data concerns no more than 2 trials (12).

Given the intermittent nature of the ST, 3 rolling averages (10, 20. and 30 seconds) were calculated to identify which would be the most valid for assessing $\dot{V}O_2max$ and which provided the most reliable measure. For submaximal intensities, taken from the first 4 stages of the test, again 3 different averaging intervals were calculated to indicate which duration yielded the most reliable $\dot{V}O_2$ data. These were fixed-time intervals covering the final 10, 30, and 60 seconds of the stage.

Table 2. Overview of test structure and stage lengths.

Stage	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Laps	5	5	6	6	6	7	7	7	8	8	8	9	9	9	10
Stage length (min)	03:35	03:13	03:33	03:17	03:04	03:24	03:13	03:04	03:22	03:15	03:08	03:25	03:19	03:14	03:32

Results

Validity Between Squash Test and Laboratory Test

Pretest Measures. All subjects arrived similarly hydrated for both trials (LAB = 1.010 [0.008], ST = 1.009 [0.006], $p > 0.05$). Pretest wellness questionnaires revealed no differences between trials (for all variables $p > 0.05$). Environmental conditions were not different between trials for ambient temperature (LAB = 24.1 [0.7]° C, ST = 24.5 [0.5]° C); relative humidity (LAB = 51 [2.7]%, ST = 53 [2.2]%); or wet bulb globe temperature (LAB = 20.3 [0.9]° C, ST = 20.8 [0.4]° C).

Exercise Measures. During LAB, all subjects displayed a $\dot{V}O_2$ plateau and this was matched by a verification within 2 $ml \cdot kg^{-1} \cdot min^{-1}$ for 6 subjects, whereas 2 subjects terminated exercise before the verification was within 2 $ml \cdot kg^{-1} \cdot min^{-1}$ of the existing $\dot{V}O_2max$ value. Table 3 shows the agreement of $\dot{V}O_2max$ between LAB and ST with different sampling periods. Using a 10-second rolling average, the MB in $\dot{V}O_2max$ between LAB and ST was the smallest and did not result in a significant difference between tests when compared with 20 and 30 seconds' averaging.

No statistical difference was observed between tests for maximum HR (LAB = 195 [7] $b \cdot min^{-1}$, ST = 200 [7] $b \cdot min^{-1}$, $d = 0.6$). There was no difference in postexercise blood lactate concentration (LAB = 9.2 [1.9] $mMol^{-1}$, ST = 8.6 [1.9] $mMol^{-1}$, $d = 0.28$) or end of test RPE (LAB = 19.6 [0.5] $mMol^{-1}$, ST = 19.1 [0.8] $mMol^{-1}$, $d = 0.74$) between LAB and ST.

Reliability Between Squash Tests

One participant could not complete ST2 due to injury, therefore reliability analyses are from 7 subjects.

Pretest Measures. All subjects arrived similarly hydrated for both trials (ST1 = 1.010 [0.006] and ST2 = 1.010 [0.006]). Pretest wellness questionnaires revealed no differences between trials (for all variables $p > 0.05$). Environmental conditions were not different between trials for ambient temperature (ST1 = 24.5 [0.8]° C, ST2 = 24.5 [0.6]° C), relative humidity (ST1 = 52.9 [4.3]%, ST2 = 53.3 [2.8]%), or wet bulb globe temperature (ST1 = 20.9 [1.2]° C, ST2 = 20.8 [0.5] ° C).

Maximal Exercise Measures. One player performed better in ST1, 5 players performed better in ST2, and one player reached the same stage on each test. Overall, test performance during ST1 and ST2 was closely matched, with 74 (10) laps completed in ST1 and 75 (12) laps completed in ST2. This equated to reaching Level 12 lap 1 during ST1 and Level 12 lap 2 during ST2. The MB between tests was 1 lap, with a TE of 3 laps (TEM 4%) and correlation coefficient of $r = 0.97$. The LOA were 7 laps (−9: 6 laps).

Different data averaging intervals were adopted to indicate the most reliable measure of $\dot{V}O_2max$ and revealed similar variability across 10, 20, and 30 seconds' rolling averages. These data are shown in Table 4.

Maximum HR was similar between tests (ST1 = 201 [9] $b \cdot min^{-1}$; ST2 = 201 [7] $b \cdot min^{-1}$; MB = 0 $b \cdot min^{-1}$, TEM = 3 $b \cdot min^{-1}$, TEM [CV%] = 1%, $r = 0.98$), as was postexercise blood lactate concentration (ST1 = 9.7 [1.3] $mMol^{-1}$; ST2 = 9.9 [2.2] $mMol^{-1}$; $p = 0.80$) and postexercise RPE (ST1 = 19 [1]; ST2 = 19 [1]; $p = 0.74$).

Submaximal Exercise Measures. Lap number at 2 $mMol^{-1}$ was similar during ST1 (47 [7]) and ST2 (49 [9]). This resulted in an MB of 2 laps, TEM of 4 laps, and a TEM (CV%) of 8%. This provided a correlation of $r = 0.74$, whereas the LOA [lower: upper] were −13: 8 laps and there was no statistical difference between trials ($p = 0.29$). Total laps at 4 $mMol^{-1}$ was also highly agreeable, occurring at 59 (8) laps in ST1 and 60 (8) laps in ST2. This resulted in a MB of 1 lap, TEM of 4 laps, and a TEM (CV%) of 7%. The correlation coefficient was $r = 0.77$, the LOA [lower: upper] were −12: 10 laps, and there was no statistical difference between trials ($p = 0.54$).

Average $\dot{V}O_2$ across the first 4 stages was lower in ST2, which was statistically significant across data derived from the final 10, 30, and 60 seconds of each stage. However, for all

Table 3. Agreement of $\dot{V}O_2max$ derived from 10, 20, and 30 seconds' rolling averages during laboratory treadmill test and squash fitness test ($n = 8$).*

	Rolling average					
	10 s		20 s		30 s	
	LAB	ST	LAB	ST	LAB	ST
Mean ($ml \cdot kg^{-1} \cdot min^{-1}$)	48.8	46.2	48.4	44.5†	47.9	43.7†
SD	5.0	4.1	5.1	4.1	4.9	4.2
Mean bias ($ml \cdot kg^{-1} \cdot min^{-1}$)	2.5		3.9		4.2	
TEE ($ml \cdot kg^{-1} \cdot min^{-1}$)	3.3		3.2		3.3	
TEE (CV%)	7.0		6.9		7.2	
LOA (lower:upper) ($ml \cdot kg^{-1} \cdot min^{-1}$)	−3.5:8.6		−1.8:9.7		−1.81:10.2	
r	0.79		0.82		0.79	
d	0.56		0.86		0.91	

*LAB = laboratory treadmill test; ST = squash fitness test; TEE = typical error of the estimate, where LAB represents the criterion measure; LOA = limits of agreement; *r* = Pearson's correlation coefficient; *d* = Cohen's *d* effect size.
†Statistically significant difference vs. LAB ($p \leq 0.05$).

Table 4. Agreement of $\dot{V}O_2$max derived from 10, 20, and 30 seconds' rolling averages during 2 squash fitness tests (ST1 and ST2) ($n = 7$).†

	Rolling average					
	10 s		20 s		30 s	
	ST1	ST2	ST1	ST2	ST1	ST2
Mean ($ml \cdot kg^{-1} \cdot min^{-1}$)	47.7	46.2	45.4	44.3	45.0	43.5
SD	5.4	6.8	4.4	5.1	4.7	5.4
Mean bias ($ml \cdot kg^{-1} \cdot min^{-1}$)	1.0		1.1		1.5	
TEM ($ml \cdot kg^{-1} \cdot min^{-1}$)	1.5		1.2		1.4	
TEM (CV%)	3.2		2.8		3.1	
LOA (lower:upper) ($ml \cdot kg^{-1} \cdot min^{-1}$)	−3.2:5.1		−2.2:4.5		−2.8:5.3	
r	0.95		0.94		0.93	
d	0.25		0.24		0.29	

*TEM = typical error of the measure; LOA = limits of agreement; *r* = Pearson's correlation coefficient; *d* = Cohen's *d* effect size.

†Statistically significant difference between trials ($p \leq 0.05$).

averaging intervals, the MB and TEM between tests was <2 $ml \cdot kg^{-1} \cdot min^{-1}$, indicating a modest magnitude of difference. The most agreeable data were derived from 60-second averages, whereby ST1 = 38.4 (4.2) $ml \cdot kg^{-1} \cdot min^{-1}$, compared with 36.5 (4.1) $ml \cdot kg^{-1} \cdot min^{-1}$ in ST2 (MB = 1.9 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM = 0.94 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM [CV%] = 2.5%, $r = 0.95$, LOA [lower: upper] = −0.7: 4.5 $ml \cdot kg^{-1} \cdot min^{-1}$, $d = 0.46$, $p = 0.005$). This compared with 30-second averaging whereby ST1 = 38.4 (4.1) $ml \cdot kg^{-1} \cdot min^{-1}$ and ST2 = 36.4 (4.2) $ml \cdot kg^{-1} \cdot min^{-1}$ (MB = 2 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM = 1 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM (CV %) = 2.7%, $r = 0.94$, LOA [lower: upper] = −0.8: 4.8 $ml \cdot kg^{-1} \cdot min^{-1}$, $d = 0.48$, $p = 0.006$). Finally, data from the final 10 seconds of each stage revealed ST1 = 40 (4.8) $ml \cdot kg^{-1} \cdot min^{-1}$, compared with 38.1 (5.1) $ml \cdot kg^{-1} \cdot min^{-1}$ in ST2 (MB = 1.9 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM = 1.2 $ml \cdot kg^{-1} \cdot min^{-1}$, TEM (CV%) = 3.1%, $r = 0.94$, LOA [lower: upper] = −1.4: 5.3 $b \cdot min^{-1}$, $d = 0.39$, $p = 0.02$).

Average HR across the first 4 stages during ST1 (169 [8] $b \cdot min^{-1}$) and ST2 (167 [8] $\cdot b \cdot min^{-1}$) was also highly agreeable (MB = −2 $b \cdot min^{-1}$, TEM = −3 $b \cdot min^{-1}$, TEM [CV%] = 2%, $r = 0.84$, LOA [lower: upper] = −6: 11 $b \cdot min^{-1}$, $d = 0.31$, $p = 0.166$).

Discussion

We present a modified on-court squash fitness test that allows for measurements of pertinent submaximal and maximal physiological markers, as well as performance, which can be implemented without complex equipment. We observed agreement between LAB and ST for assessing $\dot{V}O_2$max (TEE = 3.3 $ml \cdot kg^{-1} \cdot min^{-1}$), when a rolling 10-second average was adopted, as well as maximum HR, demonstrating criterion validity. Test performance was highly reliable, with a TEM of 3 laps when repeated after 48 hours. Physiological responses such as $\dot{V}O_2$max and lap number at 4 $mMol^{-1}$ were also highly reliable, with TEs of <1.5 $ml \cdot kg^{-1} \cdot min^{-1}$ and 4 laps, respectively. Consequently, we recommend the use of this test for assessing and monitoring physical performance and prominent physiological markers in amateur and elite squash players.

The validity of generic fitness tests that do not include frequent changes of direction and multiple accelerations and decelerations must be considered when assessing fitness in squash. Existing squash-specific fitness tests contain short stages that limit inferences regarding submaximal physiology, or require computer equipment to create randomization of shuttle runs (8,27). Although our test can be completed using only the audio track to acquire a measure of performance, the 30-second rest periods at the end of each stage facilitate blood sampling as well as the collection of other physiological and perceptual data useful for optimizing and individualizing training. The simplified format would seem suitable for junior or amateur athletes, requiring only a mobile phone to play the audio and complete the assessment. We did not include racquet strokes in this on-court assessment as pilot testing revealed high levels of variation during the latter stages of the maximal test. Moreover, pilot testing suggested that some players may fall behind the required speed because of a slower shot technique, rather than being fatigued, which confounds inference about critical physical qualities. Similarly, we removed any requirements to undertake squash-specific movements, such as backward running, after pilot testing revealed tests ended before exhaustion.

Criterion validity is demonstrated by the correlations between $\dot{V}O_2$max scores from the ST and LAB ($r = 0.8$ for all averaging intervals), a recognized method of $\dot{V}O_2$max assessment. Although the MB across the group was acceptable (2.5 $ml \cdot kg^{-1} \cdot min^{-1}$), the largest difference for any participant was 6 $ml \cdot kg^{-1} \cdot min^{-1}$, which could indicate that other physical limitations may have prevented maximizing of aerobic capacity. For example, change of direction ability and associated attributes such as reactive strength, speed, and anthropometry may influence performance in this highly representative sport-specific test. This is an area for future research to consider, which will help inform training prioritization.

The most agreeable $\dot{V}O_2$max data were from 10-second averaging, with 20 and 30 seconds demonstrating larger errors, which may be inappropriate averaging durations for measuring $\dot{V}O_2$max in an intermittent test such as this. Although 10-second averaging is shorter than the relatively

ubiquitous rolling 30-second average, it is not without precedence for measuring $\dot{V}O_2$max (20). Indeed, recommendations are available for steady-state exercise (26), but do not seem to exist for intermittent exercise, despite tests such as the Yoyo (15) or the 30-15 (5) being used to estimate aerobic capacity, with the use of a portable metabolic cart. Given the relatively low work: rest ratio in this test, where stages above level 7 result in lap durations of <20 seconds before each 10-second rest, a rolling 10-second average contains only exercising data, rather than a combination of exercise and the subsequent rest period, providing agreement with the laboratory test. The attainment of a $\dot{V}O_2$ plateau and verifications for 6 of these tests provide confidence that treadmill tests were maximal. Therefore, although the intermittent nature of ST negates the attainment of a $\dot{V}O_2$ plateau, the close association of the 2 tests is likely comparable with the measurement error of the metabolic cart (7) and supported by the similarity of maximum HR and postexercise blood lactate concentrations. Unlike Wilkinson et al. (29) and Girard et al. (8), we did not observe a greater $\dot{V}O_2$max during ST compared with LAB. The reason for this is unclear; however, our squash test has a different intensity profile, with longer stages, and we did not include racquet strokes, which may have reduced the overall energy expenditure.

Test performance was highly reliable, indicating that inferences regarding player fitness can be made within stages, based on the number of laps they complete at a given stage. The observed MB (1 lap) and TE (3 laps) suggest that a change of 4 laps can confidently be interpreted as a real change in test performance. Indeed, the smallest worthwhile change was 2 laps (2). End users should be aware, however, that we cannot discount a learning effect when the test is first performed because all players had completed this test previously. Alongside the performance outcome, physiological markers such as $\dot{V}O_2$peak, maximum HR, and postexercise blood lactate concentration were all highly reliable. As shown in Table 4, the close agreement between $\dot{V}O_2$peak across averaging intervals of 10, 20, and 30 seconds' rolling averages, indicates reliability of $\dot{V}O_2$peak, irrespective of which rolling average is adopted. This supports the use of a 10-second rolling average to track maximal aerobic capacity during the test. At submaximal intensities, the number of laps completed at both 2 and 4 $mMol^{-1}$ was highly reliable between tests, with a smaller MB and stronger correlation at 4 $mMol^{-1}$ compared with 2 $mMol^{-1}$. Nevertheless, that the TE for both markers was 4 laps suggests that metabolic transitions, that these fixed blood lactate concentrations are used to demarcate, can be tracked relatively accurately when suitable precautions are taken to control for typical factors that can influence the blood lactate response such as recovery and nutritional status, as we did before testing. There was a consistent, albeit modest (<2 $ml \cdot kg^{-1} \cdot min^{-1}$) reduction in average $\dot{V}O_2$ across the first 4 stages in ST2, which could represent a learning effect on the test at lower intensities. This was supported by similarly modest reduction in HR across these stages (-2 $b \cdot min^{-1}$) and difference in lap number at 2 $mMol^{-1}$, but the trend does not seem to continue as exercise intensity increases, with data more agreeable for 4 $mMol^{-1}$ and $\dot{V}O_2$peak. Fixed-time averages of the final 10, 30, and 60 seconds of each stage provided similarly agreeable responses, with 60 seconds showing the strongest agreement across our battery of statistics. Because 60-second sampling interval captures more data from the test, we recommend users adopt this sampling period.

Practical Applications

Coaches and practitioners may use this test in a simplified format (audio only), to provide a measure of performance, or with accompanying physiological measures such as HR, blood lactate, and oxygen uptake, to understand training responses. The simplified performance test may be appropriate for amateur and subelite players, whereas elite players will benefit from physiological data from which training can be planned on an individual basis. When physiological measures are taken during the test, the attainment of a $\dot{V}O_2$peak during a representative, sport-specific test, such as ST, may be just as relevant as tracking $\dot{V}O_2$max in a laboratory, to understand training responses. Future research may therefore want to investigate the appropriateness of adding a verification procedure to ST to increase confidence of data interpretation.

Girard et al. (9) provided training recommendations, which can be used to prescribe highly specific squash training. Data from the current test can be used in a similar way, for example to individualize training and target specific physiological adaptations that coaches can have high levels of confidence that will transfer to the court, given the close association of our test design with characteristics of elite squash match play. We have shown that a sport-specific blood lactate profile can be developed, which allows training to be individualized on an athlete's needs. For example, the maximum sustainable exercise intensity on the court i.e., the speed at which LTP occurs, can be identified from the test, and then coach ball-feeding sessions or individual, simulated match-play repetitions ("ghosting") using an accompanying audio track, can be based around this intensity.

In conclusion, we present a modified squash-specific fitness test that can be used both as a performance measure, whereas also allowing for pertinent physiological measures to be taken. Test results were both valid and reliable and we recommend the use of the test for both subelite and elite squash players.

Acknowledgments

The authors thank the Malaysian squash players who participated in this study, as well as national coaches Peter Genever and Andrew Cross for their help and advice throughout the development of this test. The authors have no funding or conflicts of interest to disclose.

References

1. Atkinson, G and Nevill, A. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sport Med* 26: 217–238, 1998.
2. Batterham, AM and Hopkins, WG. Making meaningful inferences about magnitudes. *Int J Sports Physiol Perform* 1: 50–57, 2006.
3. Bland, JM and Altman, DG. Statistical methods for assessing agreement between two methods of Clinical measurement. *Lancet* 327: 307–310, 1986.
4. Borg, G. *Borg's Perceived Exertion and Pain Scales*. Champaign, IL: Human Kinetics, 1998. p. 104.
5. Buchheit, M. The 30-15 intermittent fitness test: Accuracy for individualizing interval training of young intermittent sport players. *J Strength Cond Res* 22: 365–374, 2008.
6. Burnley, M and Jones, AM. Oxygen uptake kinetics as a determinant of sports performance. *Eur J Sport Sci* 7: 63–79, 2007.
7. Garcia-Tabar, I, Eclache, JP, Aramendi, JF, and Gorostiaga, EM. Gas analyzer's drift leads to systematic error in maximal oxygen uptake and maximal respiratory exchange ratio determination. *Front Physiol* 6: 308, 2015.
8. Girard, O, Sciberras, P, Habrard, M, Hot, P, Chevalier, R, and Millet, GP. Specific incremental test in elite squash players. *Br J Sports Med* 39: 921–926, 2005.
9. Girard, O, Chevalier, R, Habrard, M, Sciberras, P, Hot, P, and Millet, GP. Game analysis and energy requirements of elite squash. *J Strength Cond Res* 21: 909–914, 2007.
10. Gouttebarge, V, Moen, M, Boschman, J, and Kantebeen, M. Squash-specific exercise test for elite squash players: Development and validity. *Eur J Sport Exerc Sci* 2: 1–6, 2013.
11. Harper Smith, AD, Crabtree, DR, Bilzon, JLJ, and Walsh, NP. The validity of wireless iButtons and thermistors for human skin temperature measurement. *Physiol Meas* 31: 95–114, 2010.
12. Hopkins, WG. Measures of reliability in sports medicine and science. *Sport Med* 30: 1–15, 2000.
13. James, CA, Richardson, AJ, Watt, PW, and Maxwell, NS. Reliability and validity of skin temperature measurement by telemetry thermistors and a thermal camera during exercise in the heat. *J Therm Biol* 45: 141–149, 2014.
14. James, CA, Willmott, AGB, Richardson, AJ, Watt, PW, and Maxwell, NS. Ischaemic preconditioning does not alter the determinants of endurance running performance in the heat. *Eur J Appl Physiol* 116: 1735–1745, 2016.
15. Krustrup, P, Mohr, M, Amstrup, T, Rysgaard, T, Johansen, J, Steensberg, A, Pedersen, PK, and Bangsbo, J. The Yo-Yo intermittent recovery test: Physiological response, reliability, and validity. *Med Sci Sports Exerc* 35: 697–705, 2003.
16. Locke, S, Colquhoun, D, Briner, M, Ellis, L, O'Brien, M, Wollstein, J, and Allen, G. Squash racquets. A review of physiology and medicine. *Sports Med* 23: 130–138, 1997.
17. McLean, BD, Coutts, AJ, Kelly, V, McGuigan, MR, and Cormack, SJ. Neuromuscular, endocrine, and perceptual fatigue responses during different length between-match microcycles in professional rugby league players. *Int J Sports Physiol Perform* 5: 367–383, 2010.
18. Micklewright, D and Papapdopoulou, E. A new squash specific incremental field test. *Int J Sports Med* 29: 758–763, 2008.
19. Midgley, AW and Carroll, S. Emergence of the verification phase procedure for confirming "true" VO2max. *Scand J Med Sci Sport* 19: 313–322, 2009.
20. Midgley, AW, Mcnaughton, LR, and Carroll, S. Effect of the V̇o2 time-averaging interval on the reproducibility of V̇o2max in healthy athletic subjects. *Clin Physiol Funct Imaging* 27: 122–125, 2007.
21. Mier, CM, Alexander, RP, and Mageean, AL. Achievement of VO2max criteria during a continuous graded exercise test and a verification stage performed by college athletes. *J Strength Cond Res* 26: 2648–2654, 2012.
22. Montpetit, RR. Applied physiology of squash. *Sport Med* 10: 31–41, 1990.
23. Saunders, PU and Green, DJ. Runners and walkers. In: *Physiological Tests for Elite Athletes*. RK Tanner and CJ Gore, eds. Champaign, IL: Human Kinetics, 2013. pp. 397–409.
24. Sawka, MN, Burke, LM, Eichner, ER, Maughan, RJ, Montain, SJ, and Stachenfeld, NS; American College of Sports Medicine position stand. Exercise and fluid replacement. *Med Sci Sports Exerc* 39: 377–390, 2007.
25. Sherman, RA and Creasey, TJBA. An on-court, ghosting protocol to replicate physiological demands of a competitive squash match. In: *Science and Racket Sports 3*. AKJ Lees, ed. Oxon, United Kindom: Routledge, 2004. p. 3–8.
26. Spurway, J and Jones, AM. Lactate testing. In: *Sport and Exerice Physiology Testing Guidelines: The British Association of Sport and Exercise Sciences Guide*. EM Winter, AM Jones, RCR Davison, PD Bromley and T Mercer, eds. (Vol 1). Oxon, United Kingdom: Sport Testing, 2007. pp. 112–119.
27. Steininger, K and Wodick, RE. Sports-specific fitness testing in squash. *Br J Sports Med* 21: 23–26, 1987.
28. Todd, MK, Mahoney, CA, and Wallace, WFM. The efficacy of training routines as a preparation for competitive squash. In: *Science and Racket Sports II*. A Lees, I Maynard, M Hughes, and T Reilly, eds. London, United Kingdom: Taylor and Francis, 1998.
29. Wilkinson, M, Leedale-Brown, D, and Winter, EM. Validity of a squash-specific fitness test. *Int J Sports Physiol Perform* 4: 29–40, 2009.