"IDO MOVEMENT FOR CULTURE. Journal of Martial Arts Anthropology",
Vol. 20, no. 3 (2020), pp. 38–46
DOI: 10.14589/ido.20.3.6

COACHING & KINESIOLOGY

Jonatas Ferreira da Silva Santos[1,2(A-G)], Joao Paulo Lopes-Silva[2,3(A-E)], Irineu Loturco[4(A-E)], Emerson Franchini[2(A-G)]

1 Physical Training and Sport Performance Research Group, Physical Education Department, Healthy and Biological Science Faculty, Federal University of the Jequitinhonha and Mucuri Valleys, Diamantina – Minas Gerais (Brazil)
2 Martial Arts and Combat Sports Research Group, Sport Department, School of Physical Education and Sport, University of Sao Paulo, Sao Paulo (Brazil)
3 School of Physical Education, CESMAC University Center, Maceio, Alagoas (Brazil)
4 Nucleus of High Performance in Sport – NAR, Sao Paulo (Brazil)
Corresponding Author: Jonatas Ferreira da Silva Santos (Orcid code: 0000-0002-3309-4731)
UFVJM - Campus JK: Physical Education Department, Healthy and Biological Science Faculty, Federal University of the Jequitinhonha and Mucuri Valleys, Diamantina – Minas Gerais, Brazil.
Rodovia MGT 367 - Km 583, nº 5000
Bairro Alto da Jacuba, Diamantina/MG, ZIP code: 39100-000
e-mail: jonatas_contato@hotmail.com

Test-retest reliability, sensibility and construct validity of the frequency speed of kick test in male black-belt taekwondo athletes

Submission: 8.05.2019; acceptance: 21.11.2019

Key words: physical fitness, specific evaluation, turning kick, combat sports, martial arts

Abstract

Aim. The purpose of the present study was to describe the reliability, sensibility and construct validity (discriminant) of male taekwondo athletes in the frequency speed of kick test (FSKT) using single and multiple sets.

Material, Methods and Results. The study was divided in two parts. The first part consisted of fourteen male black-belt taekwondo athletes. Test and retest were carried out on separate occasions 1 week apart. The total number of kicks generated during $FSKT_{10s}$ and each set of $FSKT_{mult}$, the sum of kicks during five sets, and kicks decrement did not display any difference. The ICC and SEM for absolute and relative performance during $FSKT_{10s}$ were 0.95 and 0.60, respectively. The ICC and SEM for absolute and relative performance during $FSKT_{mult}$ were between 0.63-0.83 and 0.78-3.99, respectively, for almost all measures. The limits of agreement analyses reported a mean difference between -0.1 and 0.5. The SEM was lower than $SWC_{0.6}$ in almost all variables, classifying the tests as "good". This study showed that the $FSKT_{10s}$ and $FSKT_{mult}$ are reliable taekwondo-specific field tests for assessing taekwondo performance. The second part was made up of 153 male taekwondo athletes divided into National/International (n=45), Regional/State (n=55) and non-competitor (n=53 groups. The National/International group presented superior performance compared to the non-competitors in $FSKT_{10s}$ and $FSKT_{mult}$ ($FSKT_3$, $FSKT_5$ and $FSKT_{total}$) ($p<0.05$).

Conclusions. Considering scientific criteria as an important characteristic, the results presented herein should be of interest to coaches, trainers, practitioners and *taekwondo* athletes, as these are the only short duration and intermittent taekwondo tests, using specific action, currently available in the scientific literature.

Introduction

Taekwondo is an Olympic combat sport, characterized by high-intensity actions followed by low-intensity actions or pauses [Bridge *et al.* 2014; Campos *et al.* 2012]. The predominant energy metabolism is oxidative, however, the main actions which decide a match are performed with the supply from the phosphagen metabolism [Campos *et al.* 2012]. Taekwondo athletes predominantly use kicks to hit the opponent in the trunk or head, with the main objective being a knockout or score [Bridge *et al.* 2014; Kazemi, Perri, Soave 2010]. The most commonly used technical action is a turning kick, named the bandal tchagui [Kwok 2012]. This blow has the purpose of hitting the opponent's trunk as powerfully and fast as possible [Pieter, Pieter 1995]. Thus, a taekwondo-specific

test that mimics these characteristics could help coaches and strength and conditioning professionals to monitor training adaptations and athletes' competitiveness.

Since taekwondo became an Olympic combat discipline, there has been an increase in the quantity of scientific information regarding performance in this sport; however, the majority of data is collected from "non-specific tasks", such as strength-power tests (1RM, SJ, CMJ and Sit-ups), aerobic (maximum oxygen uptake on a treadmill, cycle ergometer or shuttle run test) and anaerobic assessments (Wingate test), and flexibility evaluations (sit-and-reach test) [Bridge *et al*. 2014]. Nonetheless, tests that mimic combat situations seem to be more appropriate to assess these athletes as well as being able to measure and evaluate "specific taekwondo skills", which directly influence combat performance [Bridge *et al*. 2014]. Recently, a few tests using specific taekwondo actions have been proposed in the literature [Sant'Ana *et al*. 2014; Sant'Ana, Silva, Guglielmo 2009; Santos, Valenzuela, Franchini 2015; Santos, Franchini 2016; Santos *et al*. 2016; Santos, Loturco, Franchini 2018; Santos, Franchini 2018]. For instance, aerobic performance has been measured using a progressive test, performed applying the bandal tchagui technique [Sant'Ana, Silva, Guglielmo 2009; Araujo *et al*. 2017]. In addition, power performance has been assessed using time to execution of the same technique [Sant'Ana *et al*. 2014; Sant'Ana, Silva, Guglielmo 2009; Santos, Franchini 2016; Santos *et al*. 2016; Santos, Valenzuela, Franchini 2015; Araujo *et al*. 2017] and impact generated by a blow [Sant'Ana *et al.* 2014]. Lastly, anaerobic performance has been measured using modified tests [Sant'Ana *et al*. 2014; Santos, Valenzuela, Franchini 2015], conducted with the same duration as the Wingate test [Bar-Or 1987; Inbar, Bar-Or, Skinner 1996]. These tests are applied continuously for 30s, performing the semicircle bandal tchagui kick alternating right and left legs. It is important to note that tests with these characteristics may decrease ecological validity, since a taekwondo match is intermittent in nature [Bridge *et al*. 2014], alternating short high-intensity actions and low-intensity actions [Campos *et al*. 2012; Kazemi, Perri, Soave 2010; Kwok 2012].

Another limitation of these tests is that they have not been validated for taekwondo. More recently, the frequency speed of the kick test has been used in two versions, 10s ($FSKT_{10s}$) and 90s ($FSKT_{mult}$) [Santos, Valenzuela, Franchini 2015; Santos, Franchini 2016; Santos *et al*. 2016; Santos, Loturco, Franchini 2018; Santos, Franchini 2018]. The $FSKT_{10s}$ is composed of kicks alternating right and left legs as powerfully as possible throughout the time. The second version, $FSKT_{mult}$, is performed in five sequential sets of 10s maximal kicks applications following a 10s interval between sets until the fifth set is finished. The number of blows (kicks and punches) executed during a taekwondo match is an important characteristic, which can be decisive to the match outcome. This can happen as in the case of a draw during the fourth round, the highest impact number is the decisive factor to determine the winner. During the FSKT, the number of kicks is performed in a similar intermittent manner as presented during an official match, following the high-intensity and low-intensity actions.

Importantly, $FSKT_{10s}$ and $FSKT_{mult}$ have been shown to be highly sensitive to identify alterations in performance of taekwondo athletes after nine-weeks of training [Santos *et al*. 2016] and differences between female taekwondo athletes grouped by competitive levels [Santos, Franchini 2018]. However, although the FSKT appears to be an appropriate test to measure taekwondo athlete performance, no studies have described its reliability, sensibility and construct validity in male taekwondo black belt. Scientific criteria like reliability, sensibility and construct validity are necessary characteristic to measure and make decisions [Chaabene *et al*. 2012; Weir 2005]. Changes performance in test and retest pre and post a period of training can be confident only if the test is reliable, the sensibility and construct validity is knowledge. Therefore, the purpose of this study was to investigate different scientific criteria like reliability, sensibility and construct validity in $FSKT_{10s}$ and $FSKT_{mult}$ of taekwondo athletes. It was hypothesized that the variables generated (kicks in each set, total kick number, kick fatigue index) during $FSKT_{10s}$ and $FSKT_{mult}$ are reliable, sensible and can differentiate athletes and non-athletes.

Material and Methods

Participants

All athletes and practitioners were free from any lower-body injury or neuromuscular disorder. All provided written consent after being informed about the purpose, procedures, and associated risks. A parental written consent has been obtained for athletes under 18 years. The research was approved by the local Ethics and Research Committee.

Part 1

Fourteen black-belt male taekwondo athletes (M±SD, age: 20.6±4.2 years; height: 180.4±7.0 cm; body mass: 70.7±11.8 kg; practice time: 7.8±4.7 years) volunteered to participate in this study. The athletes were competing at regional or more prominent levels (international: 21%; national: 36%; state: 36%; regional: 7%). The athletes participated in ~7 taekwondo sessions and ~3 strength and conditioning sessions weekly. Athletes were in the competitive phase of their periodization.

Part 2

One-hundred fifty-three male taekwondo athletes divided in National/International (n=45; National: 25;

International: 20) and Regional/State (n=55; Regional: 20; State: 35) and non-competitors (n=53) (median (interquartile range); age: 22 (17;26) years; body mass: 67 (60;79) kg; height: 175 (169;181) cm; practice time: 5 (2;10) years) volunteered to participate in this study. The athletes participated in ~7 taekwondo sessions and ~3 strength and conditioning sessions weekly. Athletes were in the competitive phase of their periodization. Non-competitors participated only in ~3 taekwondo sessions weekly.

Procedures

The present study was conducted in two separated parts as described below.

Part 1
First part was a test-retest study where the $FSKT_{10s}$ and $FSKT_{mult}$ were performed twice times with one week of rest between test and retest. The measurements were recorded in the same period of day to minimize circadian variation influence on performance. Each session was composed by $FSKT_{10s}$ and after 40-min was performed $FSKT_{mult}$. The purpose was performing the maximal quantity of kicks during tests. The variables generated were: number of kicks during each set and sum of five sets and kick decrement index. All tests were recorded and the videos were used posteriorly to count the kicks applied during the FSKT.

Part 2
This study was a cross-sectional comparison of male taekwondo athletes classified in national/international, regional/state and non-competitor performing $FSKT_{10s}$ and $FSKT_{mult}$. All athletes involved in this study had experience with FSKT testing procedures. The tests were executed in the training center where each athlete or trained, during the competitive period. The assessment was performed at the same period (starting 9 a.m.) for national/international, regional/state and non-competitor group, following the $FSKT_{10s}$ and $FSKT_{mult}$ order. A rest interval of 20-min between tests were applied. Before the testing session, a general and specific warm-up routine was performed. It was composed by running, stretching and low-intensity kicks and punches totalizing approximately 15-min.

Test and measurements

Part 1 and 2
Frequency Speed of Kick Test. The $FSKT_{10s}$ was conducted as described previously [Santos, Valenzuela, Franchini 2015; Santos, Franchini 2016; Santos, Loturco, Franchini 2018; Santos, Franchini 2018]. The FSKT had duration of 10s and was performed in front of a stand bag. After the sound signal, the athletes performed the maximal number of kicks possible, alternating right and left legs. The total number of kicks applied during the test determined the performance.

Multiple Frequency Speed of Kick Test. The $FSKT_{mult}$ was executed as described previously [Santos, Franchini 2016; Santos *et al.* 2016; Santos, Loturco, Franchini 2018; Santos, Franchini 2018]. Each FSKT had duration of ten seconds. During this study, each athlete performed five FSKT bouts with a 10s rest interval between repetitions. To perform the FSKT each athlete was placed in front of the stand bag equipped with a *taekwondo* trunk protector. The protector was positioned in same height of the performer trunk. After the sound signal, the subject performed the maximal number of kicks possible, alternating right and left legs. The performance was determined by the number of kicks in each set, total number of kicks and kick decrement index (KDI) during the test. The KDI indicates performance decreases during the test. To calculate the KDI the number of kicks applied during the multiple FSKT was taken into account. The calculation was performed using an equation that takes into account the results of all FSKT sets (Equation 1) [Girard, Mendez-Villanueva, Bishop 2011].

$$KDI\ (\%) = \left[1 - \frac{FSKT1 + FSKT2 + FSKT3 + FSKT4 + FSKT5}{\text{Best FSKT} \times \text{Number of Sets}}\right] \times 100$$

(Equation 1)

Video Analysis. All tests were recorded and the videos were used posteriorly to count the kicks applied during the FSKT. Kinovea software (Kinovea®, Version 0.8.15, Joan Charmant & Contributors, Bordeaux, France) was used to count the kicks applied during the FSKT. The count started when the athlete moved the attack foot. If the athlete started the kick before completing 10s but reached the target only after 10s the kick was not considered valid. All tests were administered by same researcher (JFSS), presenting ICC intra-measurer 1.00 for $FSKT_{10s}$ and 1.00, 0.99, 1.00, 1.00, 0.99, 1.00 and 0.99 for $FSKT_1$, $FSKT_2$, $FSKT_3$, $FSKT_4$, $FSKT_5$, $FSKT_{total}$ and KDI, respectively, during $FSKT_{mult}$.

Statistical analysis

Part 1
The normality distribution of each variable was tested using the Kolmogorov-Smirnov test. All variables presented normal distribution. To compare the mean in the test and retest the paired Student t test was used and 95% confidence interval (95% CI). To evaluate the magnitude of difference, *Cohen´s* effect size was calculated using the formula: *d* = Mean of differences/SD of differences. Threshold values to determine the effect size were <0.2 (small), 0.2 and 0.8 (moderate) and >0.8 (large) [Cohen 1988]. The intraclass correlation coeffi-

cient (ICC) was used to investigate the reliability of the FSKT. The ICC were interpreted using the following criteria proposed [Hopkins *et al*. 2009]: trivial: ≤0.1, small: >0.1-0.3, moderate: >0.3-0.5, large: >0.5-0.7, very large: >0.7-0.9, nearly perfect: >0.9. Limits of agreement (LOA), standard error of measurement (SEM), and coefficient of variation (CV) were calculated as indicators of the absolute reliability. The standard error of measurement (SEM) was calculated using the square root of the mean square error (MSE) ($SEM=SD_{diff}\sqrt{2}$) [Hopkins, 2000]. The SEM as coefficient of variation (CV) was calculated using the formula: CV=(SEM/M)*100 [Hachana *et al*. 2014]. The $MDC_{95\%}$ was calculated using the formula: $MDC_{95\%}=SEM*\sqrt{2}*1.96$ [Weir 2005; Haley, Fragala-Pinkham 2006]. To evaluate the usefulness of the FSKT, the smallest worthwhile change (SWC) was calculated as 0.2*between-subject standard deviation (SD) of that test [Atikinson, Nevill, 1998; Hopkins et al., 2001]. However, if the SEM was greater than the SWC for most variables, then another magnitude was used to identify minimal change possible to be identified by the $FSKT_{10s}$ and $FSKT_{mult}$, as follows: moderate: 0.6 ($SWC_{0.6*SD}$), large: 1.2 ($SWC_{1.2*SD}$), very large: 2.0 ($SWC_{2.0*SD}$), and extra-large: 4.0 ($SWC_{4.0*SD}$), using the same equation as that used in SWC (magnitude*SD) [Hopkins, 2004]. If the SEM was lower, greater or the same as the SWC, then the test was rated as "good", "marginal", or "satisfactory", respectively [Impelizzeri, Marcora 2009; Spencer *et al*. 2006]. The results of descriptive statistics are reported as M±SD. Significance was defined as $p<0.05$.

Part 2

The distribution of each variable was examined using the Shapiro Wilk test. Some variables did not present a normal distribution: $FSKT_{10s}$, $FSKT_2$ (I/N), $FSKT_4$ (S/R) and $FSKT_{total}$ (I/N). Thus, data are presented as median, interquartile range and non-parametric statistical test were used. The Mann-Whitney test was used like post-hoc of the Kruskal-Wallis test to compare groups. The Bonferroni correction was applied and all effects analyses were conducted using α= 0.0167%. The effect size (ES) was calculated using the following formula: $r=Z/\sqrt{N}$ [Hopkins 2004] and classified using the following scale [Hopkins 2002]: ≤0.2 (trivial); >0.2 - 0.6 (small); >0.6 - 1.2 (moderate); >1.2 - 2.0 (large); >2.0 (very large). The receiver operator characteristics (ROC) curve was used to determines the sensitivity and specificity and the curve was interpreted as the probability of discriminating taekwondo athletes between national/international and regional/state.

Results

Part 1

Results of the test-retest for ICC, TEM, LOA, SEM, SWC, and $MDC_{95\%}$ are presented in Figure 1 and Table 1.

Frequency Speed of Kick Test ($FSKT_{10s}$)

There was no significant difference between test and retest for $FSKT_{10s}$ (t= 1.88, p= 0.08, 95% CI= -0.063–0.920, *d*: 0.503 [moderate]; $FSKT_{10s}$ (test; retest): 21±2, 21±2). During $FSKT_{10s}$ the SEM was lower than $SWC_{0.6}$ and classified as "good".

Multiple Frequency Speed of Kick Test ($FSKT_{mult}$)

There was no significant difference between $FSKT_{mult}$ test and retest ($FSKT_1$: t=-0.33, p= 0.75, 95% CI= 1.073–0.787, *d*: -0.088 [small]; $FSKT_2$: t= 0.40, p= 0.70, 95% CI= -0.637–0.923, *d*: 0.106 [small]; $FSKT_3$: t= 0.21, p= 0.84, 95%

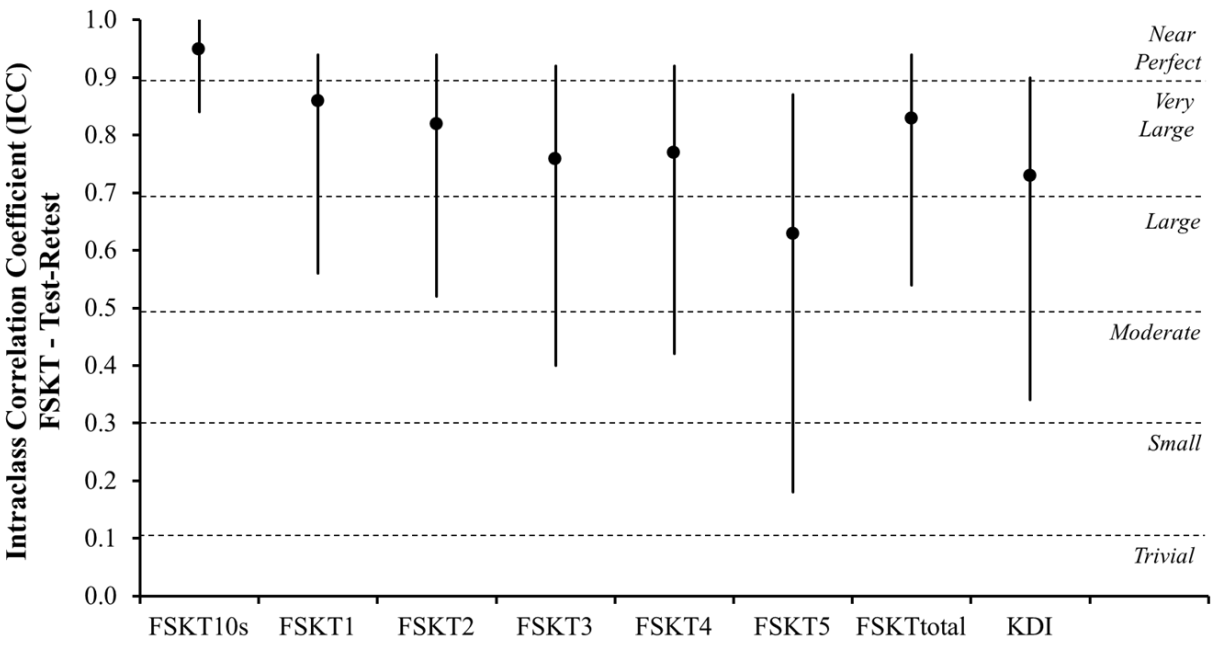

Figure 1. Intraclass correlation coefficient following classificatory criteria proposed by Hopkins et al., 2009.

Table 1. Performance and reliability results during frequency speed of kick test during test and retest (n = 14).

Variable	LOA (lowest and highest value)	SEM (kicks)	SEM as CV (%)	$SWC_{0.2}$ (kicks)	$SWC_{0.6}$ (kicks)	$MDC_{95\%}$ (kicks)
$FSKT_{10s}$	0.4 (-1.2; 2.1)	0.60	2.85	0.48	1.43	1.67
$FSKT_{mult}$						
$FSKT_1$	-0.1 (-3.3; 3.0)	1.14	5.44	0.49	1.48	3.16
$FSKT_2$	0.1 (-2.5; 2.8)	0.96	4.72	0.39	1.16	2.65
$FSKT_3$	0.1 (-2.4; 2.6)	0.90	4.69	0.33	0.99	2.49
$FSKT_4$	0.0 (-2.2; 2.2)	0.78	4.27	0.30	0.90	2.17
$FSKT_5$	0.2 (-2.4; 2.8)	0.93	5.03	0.24	0.73	2.57
$FSKT_{total}$	0.3 (-10.8; 11.4)	3.99	4.12	1.61	4.83	11.07
KDI (%)	0.5 (-4.2; 5.1)	1.69	20.41	0.63	1.88	4.68

LOA: limits of agreement; SEM: standard error of measurement; CV: coefficient of variation; SWC: smallest worthwhile change; $MDC_{95\%}$: minimal detectable change at 95% confidence interval.

Table 2. Performance during frequency speed of kick test (n = 153).

Variables	Non-competitors (n = 53)	Regional / State (n = 55)	National / International (n = 45)
	Median (IR) (Kicks)	Median (IR) (Kicks)	Median (IR) (Kicks)
$FSKT_{10s}$	19 (18;20)	20 (19;21)	20 (19;21)[a]
$FSKT_{mult}$			
$FSKT_1$	19 (18;20)	19 (18;21)	20 (19;22)
$FSKT_2$	18 (17;20)	19 (17;20)	19 (18;21)
$FSKT_3$	17 (16;19)	18 (17;19)	19 (17;20)[a]
$FSKT_4$	17 (16;18)	17 (16;19)	18 (16;19)
$FSKT_5$	16 (15;17)	17 (16;18)	17 (16;18)[a]
$FSKT_{total}$	87 (82;94)	90 (85;96)	93 (86;96)[a]
KDI (%)	8.0 (5.7-10.6)	7.8 (5.0;11.0)	8.3 (6.0;12.0)

IR: Interquartile range; a: Difference between National / International taekwondo athletes and non-competitors ($p < 0.05$).

CI= -0.661–0.804, *d*: -0.056 [small]; $FSKT_4$: t= 0.00, p= 1.00, 95% CI= -0.641–0.641, *d*: 0.000 [small]; $FSKT_5$: t= 0.61, p= 0.55, 95% CI= -0.543–0.971, *d*: 0.163 [small]; $FSKT_{total}$: t= 0.19, p= 0.85, 95% CI= -2.976–3.547, *d*: 0.051 [small]; KDI: t= 0.72, p= 0.59, 95% CI= -0.9211–1.8354, *d*: 0.192 [small]; $FSKT_{mult}$ (test; retest): $FSKT_1$: 21±2, 21±3; $FSKT_2$: 20±2, 20±2; $FSKT_3$: 19±2, 19±2; $FSKT_4$: 18±1, 18±2; $FSKT_5$: 18±1, 18±2; $FSKT_{total}$: 97±8, 97±10; KDI: 8.3±3.1, 7.8±3.0). The SEM was lower than $SWC_{0.6}$ in almost all variables, which were classified as "good" and the exception was in set 5. Results of the test-retest for ICC, TEM, LOA, SEM, SWC, and $MDC_{95\%}$ are presented in Table 1.

Part 2

The total blows applied during $FSKT_{10s}$ (H(2)= 6.80, p<0.05) and $FSKT_{mult}$ ($FSKT_3$: H(2)= 7.78, p<0.05; $FSKT_5$: H(2)= 6.62, p<0.05; $FSKT_{total}$: H(2)= 6.64, p<0.05) different among groups. The competitors of national/international level applied highest quantity of kicks in comparison of non-competitors' group (U= 1061.0, p= 0.010, r= -0.244 [small]) during $FSKT_{10s}$, but non-difference were presented among regional/state group and another's (p>0.05). Statistical difference was observed in $FSKT_{mult}$, with highest performance in national/international group compared with non-competitors group in $FSKT_3$ (U= 1041.0, p= 0.007, r= -0.257 [small]), $FSKT_5$ (U= 1067.0, p= 0.011, r= -0.242 [small]) and $FSKT_{total}$ (U= 1055.0, p= 0.010, r= -0.246 [small]), but there isn't presented statistical difference among regional/state group and another's (p>0.05). Results of the construct validity are presented in Table 2.

The ROC curve analysis was calculated between national/international and regional/state taekwondo athletes and the area under the curve was less than 0.70 for all sets, indicating that the discriminant validity was not good. The area under curve (AUC) for $FSKT_{10s}$ was 0.578 (p > 0.05) and FSKTmult was 0.571 (p > 0.05), 0.604 (p < 0.05), 0.604 (p > 0.05), 0.577 (p > 0.05), 0.576 (p > 0.05), 0.587 (p > 0.05) and 0.511 (p > 0.05) in $FSKT_1$, $FSKT_2$, $FSKT_3$, $FSKT_4$, $FSKT_5$, $FSKT_{total}$ and KDI, respectively (Figure 2).

Discussion

To our knowledge, this is the first study to investigate the reliability of FSKT by means of distinct, but complementary, statistical procedures (relative and absolute reliability). Of note, the "ICC" is one of the main statistical measures to examine the relative reliability of

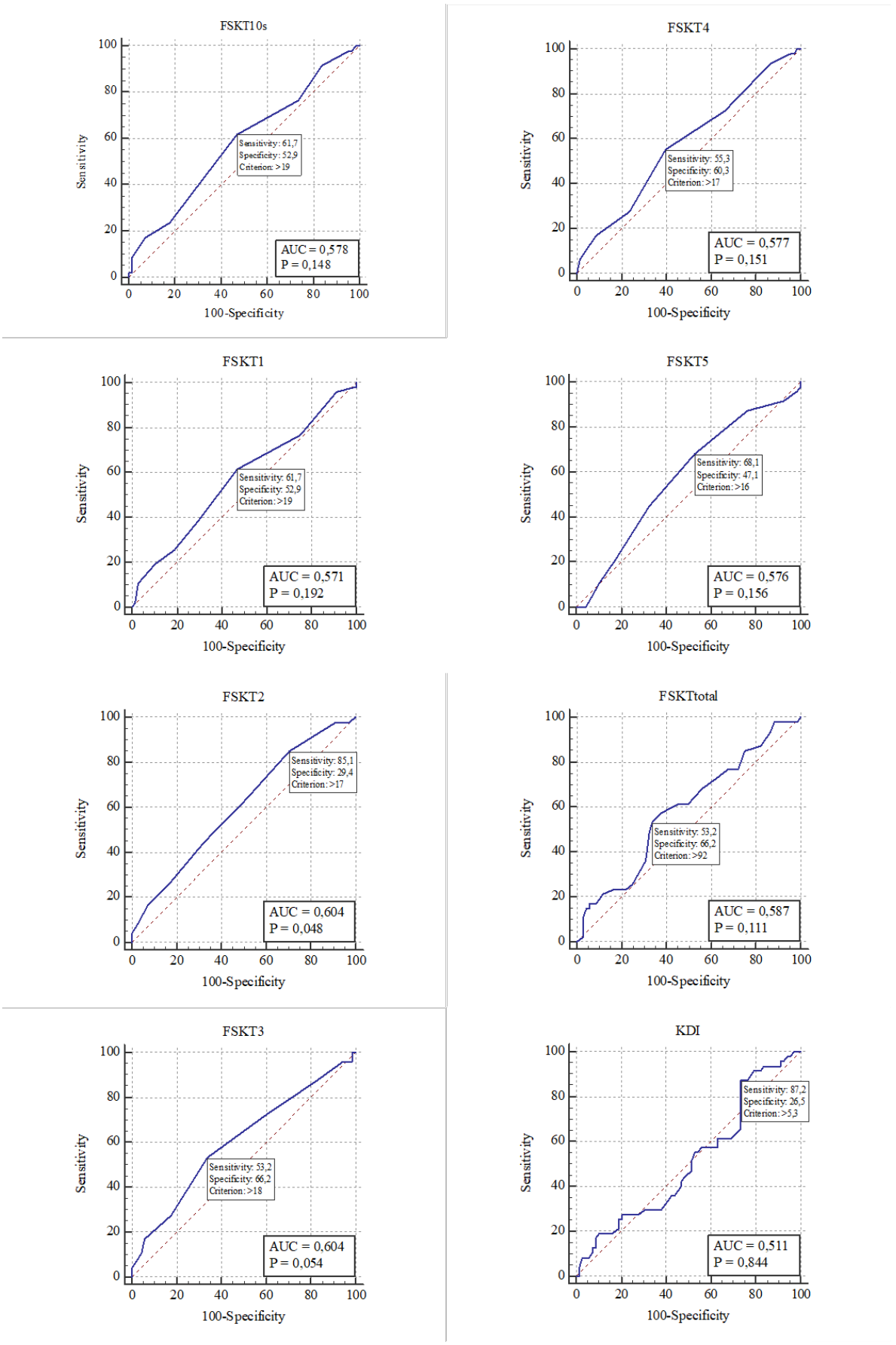

Figure 2. Receiver operating characteristics (ROC) curve for the frequency speed of kick test between regional/state and national/international taekwondo athletes.

a given test applied several times. In the present study, the relative reliability of $FSKT_{10s}$ using ICC was 0.95. The criterion adopted herein resulted in "near perfect" relative reliability classification [Hopkins *et al.* 2009]. The relative reliability of the $FSKT_{mult}$ using the ICC was 0.63 to 0.83 for the different measures of performance generated by the test. The relative reliability of these variables was classified as "large" to "very large" [Hopkins *et al.* 2009], evident by observing that the athletes performed the same average number of kicks during the test and retest in both tests. Additionally, it was presented difference between athletes and non-athletes but not difference between competitors' groups was presented.

The reliability can also be calculated absolutely, and used to estimate the amount of error in the result of a test. The measures used in the present study were LOA, SEM, SWC, and $MDC_{95\%}$. The absolute reliability presented by these measurements during FSKT was acceptable. The LOA presented high agreement between the test and retest for $FSKT_{10s}$ and $FSKT_{mult}$ (Table 1). Observing the LOA result it can be concluded that the random error of the tests is low. In a practical context, if a performer achieves a 23-kick performance in the first attempt of $FSKT_{10s}$, a performance between 21 (23-1.7) and 25 kicks (23+1.7) would be expected in the second attempt.

SEM can be expressed in an absolute or relative way. It has been recommended that SEM as CV should not be greater than 5% [Nevill, Atkinson 1997]. In the present study, the absolute reliability of $FSKT_{10s}$ was very good, with SEM as CV of 2.85%, which is well below the 5% limit. This percentage value represents approximately one kick. Changes in performance of this magnitude or greater were observed during the realization of $FSKT_{10s}$ [Santos *et al.* 2016; Araujo *et al.* 2017]. The absolute reliability of $FSKT_{mult}$ was good for almost all variables investigated. The SEM as CV presented variation of approximately 5%, a value equal to the limit indicated. This value represents approximately one kick variation for each FSKT series. This means that to detect actual improvements in performance, the changes need to be larger than one kick per set. A previous study reported changes in $FSKT_{mult}$ performance between three and four kicks after a nine-week training period [Santos *et al.* 2016]. For the total kicks executed during the five series ($FSKT_{total}$), the performance improvement would be relevant if there was an increase of four kicks. An average increase of 17 kicks has been described in the literature [Santos *et al.* 2016]. The only variation in SEM of more than 5% is associated with the KDI (%), which presented a variation of 20.4%, representing a variation of approximately 1.7% in the KDI. This measurement can oscillate more as it takes into account all series of the $FSKT_{mult}$ and consequently the possible changes that can occur in each of them.

With SEM it is possible to calculate the $MDC_{95\%}$, which is an estimate of the minimum detectable change to which the athlete's performance will need to be changed so that the change in performance is certainly greater than the error measurement [Haley, Fragala-Pinkham 2006]. In the present study the $MDC_{95\%}$ represents an amount that can vary between two and three kicks depending on the FSKT series considered. Other studies using FSKT are required for comparisons of $MDC_{95\%}$. However, it is worth noting that the limitation of this measure is that it can be affected by the variability of the sample from which it is generated [Starford *et al.* 1996]. Therefore, it is recommended that $MDC_{95\%}$ with athletes of different characteristics (sex, age, training level) and large samples be evaluated in future studies [Chaabene *et al.* 2012].

The SEM can be used in conjunction with the SWC to classify the test. In this regard, in the present study, two values of SWC were used, multiplied by 0.2 and 0.6 (Table 1). This strategy was used to enable establishment of the sensitivity of the tests, reducing the probability of an erroneous interpretation in changes across periods or test repetitions. Thus, the test was considered "good" for $FSKT_{10s}$, $FSKT_1$, $FSKT_2$, $FSKT_3$, $FSKT_4$, $FSKT_{total,}$ and KDI, as it presented higher values for $SWC_{0.6}$ compared to SEM. The test can be considered "satisfactory" for the $FSKT_5$, since the $SWC_{0.6}$ values were close to the SEM. This indicates that FSKT may not be a suitable test to detect small changes of magnitude 0.2, since the measurement error exceeds this value.

The FSKT cannot discriminate high level taekwondo athletes (national/international) of another competitors level (regional/state). In previous study was possible identify performance difference between female athletes, with superiority in national/international group compared to regional/state group [Santos, Franchini 2018]. The same characteristic not was presented between male competitors. Male *taekwondo* athletes grouped in national/international performed superior number kicks compared with non-competitors but not compared with regional/state athletes. A possible explication of this result is the volume and frequency adopted routinely by high level taekwondo athletes. Another possible explanation is that high level competitors athletes adapt to the required demand (physical and physiological) of the match. The lower volume and frequency were presented by practitioners' non-competitors. Finally, one possible limitation of the present study was that grouping national and international or regional and state levels eliminate possible differences, but in present study the number of athletes in each competitive level is low to compare separately. Future studies with separated groups (e.g. international vs. national) or between different levels of experience (e.g. novice and advanced athletes) should be conducted.

Conclusion

In conclusion, this study showed that FSKT could be considered a reliable and applied taekwondo field-test, with practical implications and scientific evidence to measure specific performance in taekwondo. The FSKT presented low to high reliability, as seen by the ICC and low dispersion, observed in LOA, SEM, SEM as CV, SWC, and MDC_{95}%. The FSKT can be used as a tool to measure taekwondo performance. Coaches and strength and conditioning professionals can apply FSKT knowing that a change greater than 2-5% among series is necessary to be 95% certain that the change in performance indicates improvement and exceeds the measurement error. The $MDC_{95\%}$ presented in this study was generated by experienced athletes, predominantly competitors at national and state level. Thus, sports professionals involved in the preparation of taekwondo athletes will be able to apply the FSKT and compare the variation in their performance with the $MDC_{95\%}$ presented to be 95% certain that performance changes actually point to performance improvements, above the error associated with the test. Finally, based in present study is possible suggest that coaches needs improve physical capacity of practitioners during taekwondo practice.

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

The authors are glad to the athletes who participated in the study.

References

1. Araujo M.P., Nobrega A.C., Espinosa G., Hausen M., Castro R.R., Soares P.P., Gurgel J.L. (2017), *Proposal of a new specific cardiopulmonar exercise test for taekwondo athletes*, "The Journal of Strength & Conditioning Research", vol. 31, no. 6, pp. 1525–1535.
2. Atikinson G., Nevill A.M. (1998), *Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine*, "Sports Medicine", vol. 26, no. 4, pp. 217–238.
3. Bar-Or O. (1987), *The Wingate anaerobic test an update on methodology, reliability and validity*, "Sports Medicine", vol. 4, no. 6, pp. 381–394.
4. Bridge C.A., Santos J.F.S., Chaabene H., Pieter W., Franchini E. (2014), *Physical and physiological profiles of taekwondo athletes*, "Sports Medicine", vol. 44, no. 6, pp. 713–733.
5. Campos F.A.D., Bertuzzi R., Dourado A.C., Santos V.G.F., Franchini E. (2012), *Energy demands in taekwondo athletes during combat simulation*, "European Journal of Applied Physiology", vol. 112, no. 4, pp. 1221–1228.
6. Chaabene H., Hachana Y., Attia A., Mkaouer B., Chaabouni S., Chamari K. (2012), *Relative and absolute reliability of karate specific aerobic test (KSAT) in experienced male athletes*, "Biology of Sport", vol. 29, no. 3, p. 211–215.
7. Cohen J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Routledge Academic, New York, NY.
8. Girard O., Mendez-Villanueva A., Bishop D. (2011), *Repeated-sprint ability - part I: factors contributing to fatigue*, "Sports Medicine", vol. 41, no. 8, pp. 673-694.
9. Hachana Y., Chaabene H., Rajeb G.B., Khlifa R., Aouadi R., Chamari K., Gabbett T.J. (2014), *Validity and reliability of new agility test among elite and subelite under 14-soccer players*, "Plos One" 9:e95773.
10. Haley SM, Fragala-Pinkham MA. (2006), *Interpreting change scores of tests and measures used in physical therapy*, "Physical Therapy", vol. 86, no. 5, pp. 735–743.
11. Hopkins W. (2002*), A Scale of Magnitudes for Effect Statistics*. Available at: http://www.sportsci.org/resource/stats/effectmag.html. Accessed October 10, 2019.
12. Hopkins W. (2004), *How to interpret changes in an athletic performance test*, "Sportscience", vol. 8, pp. 1–7.
13. Hopkins W.G. (2000), *Measures of reliability in sports medicine and science*, "Sports Medicine", vol. 30, no. 1, pp. 1–15.
14. Hopkins W.G., Elske J., Schabort E.J., Hawley J.A. (2001), *Reliability of power in physical performance tests*, "Sports Medicine", vol. 31, no. 3, pp. 211–234.
15. Hopkins W.G., Marshall S.W., Batterham A.M., Hanin J. (2009), *Progressive statistics for studies in sports medicine and exercise science*, "Medicine & Science in Sports & Exercise", vol. 41, no. 1, pp. 3–13.
16. Impelizzeri F.M., Marcora S.M. (2009), *Test validation in sport physiology: lessons learned from clinimetrics*, "The International Journal of Sports Physiology and Performance", vol. 4, no. 2, pp. 269–277.
17. Inbar O., Bar-Or O., Skinner J.S. (1996), *The Wingate anaerobic test*, Human Kinetics, Champaign, IL.
18. Kazemi M., Perri G., Soave D. (2010), *A profile of 2008 Olympic taekwondo competitors*, "The Journal of the Canadian *Chiropractic Association",* vol. 54, no. 4, pp. 243–249.
19. Kwok H.H.M. (2012), *Discrepancies in fighting strategies between taekwondo medalists and non-medalists*, "Journal of Human Sport & Exercise", vol. 7, no. 4, pp. 806–814.
20. Nevill A.M., Atkinson G. (1997), *Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science*, "British Journal of Sports Medicine", vol. 31, no. 4, pp. 314–318.
21. Pieter F., Pieter W. (1995), *Speed and force in selected taekwondo techniques*, "Biology of Sport", vol. 12, pp. 257–266.
22. Sant'Ana J., Silva J.F., Guglielmo L.G.A. (2009), *Physiological variables identified in progressive specific test for taekwondo*, "Motriz", vol. 15, no. 3, pp. 611–620.

23. Sant'Ana J., Diefenthaeler F., Dal Pupo J., Detanico D., Guglielmo L.G.A., Santos S.G. (2014), *Anaerobic evaluation of taekwondo athletes*, "The International SportMed Journal", vol. 15, no. 4, pp. 492–499.
24. Santos J.F.S., Franchini E. (2016), *Is frequency speed of kick test responsive to training? A study with taekwondo athletes*, "Sport Sciences for Health", vol. 12, no. 3, pp. 377–382.
25. Santos J.F.S., Franchini E. (2018), *Frequency speed of kick test performance comparison between female taekwondo athletes*, "The Journal of Strength & Conditioning Research", vol. 32, no. 10, pp. 2934-2938.
26. Santos J.F.S., Herrera-Valenzuela T., Mota G.R., Franchini E. (2016), *Influence of half-squat intensity and volume on the subsequent countermovement jump and frequency speed of kick test performance in taekwondo athletes*, "Kinesiology", vol. 48, no. 1, pp. 95–102.
27. Santos J.F.S., Loturco I., Franchini E. (2018), *Relationship between frequency speed of kick test performance, optimal load, and anthropometric variables in black-belt taekwondo athletes*, "Ido Movement for Culture. Journal of Martial Arts Anthropology", vol. 18, no. 1, p. 39–44.
28. Santos J.F.S., Valenzuela T.H., Franchini E. (2015), *Can different conditioning activities and rest intervals affect the acute performance of taekwondo turning kick?*, "The Journal of Strength & Conditioning Research", vol. 29, no. 6, pp. 1640–1647.
29. Spencer M., Fitzsimons M., Dawson B., Bishop D., Goodman C. (2006), *Reliability of a repeated-sprint test for field-hockey*, "The Journal of Science and Medicine in Sport", vol. 9, no. 1-2, pp. 181–184.
30. Starford P.W., Binkley J., Solomon P., Finch E., Gill C., Moreland J. (1996), *Defining the minimum level of detectable change for the Roland-Morris questionnaire*, "Physical Therapy", vol. 76, no. 4, pp. 359–365.
31. Weir J.P. (2005), *Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM*, "The Journal of Strength & Conditioning Research", vol. 19, no. 1, pp. 231–240.

Niezawodność powtórnych testów, czułość i poprawność konstrukcyjna testu częstotliwości kopnięcia u zawodników *taekwondo* płci męskiej posiadających czarny pas

Słowa kluczowe: sprawność fizyczna, ocena szczegółowa, kopnięcie obrotowe, sporty walki, sztuki walki

Streszczenie

Cel. Celem niniejszej pracy było określenie wiarygodności, czułości i poprawności konstrukcyjnej (wyróżniającej) zawodników *taekwondo* płci męskiej w teście częstotliwości prędkości testu kopnięcia (FSKT) z wykorzystaniem pojedynczych i wielokrotnych zestawów.

Materiał, metody i wyniki. Badanie zostało podzielone na dwie części. W pierwsza części brało udział czternastu zawodników taekwondo z czarnym pasem. Próby i powtórne testy zostały przeprowadzone osobno w odstępie 1 tygodnia od siebie. Łączna liczba kopnięć wygenerowanych podczas $FSKT_{10s}$ i każdego setu $FSKT_{mult}$, suma kopnięć podczas pięciu setów oraz zmniejszenie ilości kopnięć nie wykazywały różnicy statystycznej. ICC (współczynnik korelacji wewnątrzklasowej) i SEM (standardowy błąd pomiaru) dla bezwzględnej i względnej wydajności podczas zmniejszenia ilości kopnięć (*kicks decrement*) wynosiły odpowiednio 0,95 i 0,60. ICC i SEM dla bezwzględnej i względnej wydajności podczas $FSKT_{mult}$ wynosiły odpowiednio 0,63-0,83 i 0,78-3,99 dla prawie wszystkich pomiarów. Umowne granice analiz wykazywały średnią różnicę pomiędzy -0,1 i 0,5. Standardowy błąd pomiaru był niższy od najmniejsza znacząca zmiana ($SWC_{0.6)}$ w prawie wszystkich zmiennych, klasyfikując testy jako "dobre". Badanie to wykazało, że $FSKT_{10s}$ i $FSKT_{mult}$ są wiarygodnymi, przystosowanymi dla taekwondo testami terenowymi służącymi do oceny wydajności taekwondo. Druga część badania obejmowała stu pięćdziesięciu trzech sportowców taekwondo płci męskiej podzielonych na krajowych/międzynarodowych (n=45), regionalnych/stanowych (n=55) i grupę, która nie startowała w zawodach (n=53). Grupa narodowa/międzynarodowa osiągnęła lepsze wyniki niż grupa, która nie startowała w zawodach w zakresie $FSKT_{10s}$ i $FSKT_{mult}$ ($FSKT_3$, $FSKT_5$ i $FSKT_{total}$) ($p<0.05$).

Wnioski. Biorąc pod uwagę kryteria naukowe jako ważną cechę, przedstawione wyniki powinny być interesujące dla trenerów, trenerów, praktyków i zawodników taekwondo, ponieważ są to jedyne testy taekwondo o krótkiej i przejściowej długości, wykorzystujące specyficzne działania, dostępne obecnie w literaturze naukowej.