Journal of Human Kinetics volume 72/2020, 265-278 DOI: 10.2478/hukin-2019-0112
Section III – Sports Training

Preliminary Validation of Mirrored Scales for Monitoring Professional Soccer Training Sessions

by
Rodrigo F. Morandi[1], *Eduardo M. Pimenta*[1], *André G. P. Andrade*[1], *Tane K. F. Serpa*[1], *Eduardo M. Penna*[1,2], *Charles O. Costa*[3], *Mário N. S. O. Júnior*[4], *Emerson S. Garcia*[1,3,4]

We aimed to create a single subjective method to assess both internal training loads and subsequent fatigue. This new training-fatigue (dose-response) scale (TFS) was composed of two similar scales with the same properties, metrics and construction criteria. These two scales were designed to rate the perceived exertion (RPETFS) and perceived fatigue (RPFTFS) in professional soccer players. Twenty-two athletes participated to establish reliability, and 15 participated to establish validity. For reliability, the intraclass correlation coefficient (ICC) and standard error of measurement (SEM) were used. For criterion validity, the Spearman's correlation coefficient and linear regression analyses were applied. Associations between RPETFS and RPFTFS were verified by a chi square test, and a further factorial exploratory analysis was conducted. RPETFS and RPFTFS were found to be reliable (ICC 0.74 and 0.77, SEM 0.30 and 0.30, respectively) and valid. RPETFS was best explained by the internal load of the Banister training impulse ($p < 0.001$), while RPFTFS was best explained by the internal load of the Stagno training impulse ($p < 0.001$). An association was found between the scales (RPETFS and RPFTFS) in which training duration had a more substantial impact on these subjective perceptions than did training intensity ($p < 0.01$). RPETFS and RPFTFS scales are reliable and valid for monitoring training sessions in Brazilian professional soccer players. The simultaneous oscillations of the RPETFS and RPFTFS scores can be used by staff members to better plan weekly training programs based on dose-response ratings. Finally, training duration must be carefully controlled because it has a greater impact than intensity on subjective perceptions.

Key words: *soccer, rating of perceived exertion, training impulse, recovery, ordinal subjective scales.*

Introduction

A rating of perceived exertion (RPE) is a practical tool for evaluating perceived exercise intensity. The RPE integrates different types of sensorial information, including signals elicited from peripheral working muscles and the cardiovascular, respiratory and central nervous systems (Borg, 1998). The most popular RPE scales are the Borg RPE scale (Borg, 1998), which has 15 levels (6 to 20), the category ratio scales CR10 (Borg, 1998) and CR100 (Fanchini et al., 2016).

In an attempt not only to evaluate training loads, but also subsequent recovery, Finnish Kenttä and Hassmén (1998) proposed a subjective recovery scale that mirrored the Borg's scale (1998). In this scale, only the verbal anchors were changed. It was intended to be used in combination with the original Borg RPE scale and

[1] - *Federal University of Minas Gerais, Belo Horizonte, Brazil.*
[2] - *Federal University of Pará, Castanhal, Brazil.*
[3] - *Cruzeiro Esporte Clube, Belo Horizonte, Brazil.*
[4] - *Federal University of Maranhão, São Luís, Brazil.*

Authors submitted their contribution to the article to the editorial board.
Accepted for printing in the Journal of Human Kinetics vol. 72/2020 in April 2020.

a recovery-related questionnaire. However, the need for the daily application of a long and time-consuming questionnaire kept people from using it in competitive sports. Additionally, although their report was a review paper, the authors did not test or validate the scale. To simplify their measurements, Laurent et al. (2011) proposed excluding the daily recovery questionnaire and mirrored the recovery scale by Kenttä and Hassmén (1998) in the Borg CR10 scale (with only the verbal anchors changed). Despite the fact that this was a promising development, it would not be methodologically correct to use a mirrored CR10 recovery scale because Borg´s CR10 is a psychophysical scale with unique properties and metrics. In addition, this recovery scale possesses verbal anchors for different topics in the same level, such as recovery, fatigue as well as metabolic and performance expectations. Therefore, it could be confusing and might result in the misinterpretation of the data. Moreover, neither of the two studies (Laurent et al., 2011; Sikorski et al., 2013) that used this recovery scale revealed how it was elaborated nor evaluated its nature (ordinal, psychophysical, etc.) or reliability.

Since the subjective perception of recovery can only indirectly estimate the actual fatigue of an individual, we designed a scale for rating the subjective perception of fatigue instead of the subjective perception of recovery. We used CR10 as a reference to design two similar scales. These were created to be used together as a training-fatigue scale (TFS), with one designed for rating perceived exertion (RPE_{TFS}) and the other designed for rating perceived fatigue (RPF_{TFS}). Both were created to be similar (mirrored), to have the same metrics (0 to 10) and the same format (blank levels intercepted by levels with verbal anchors) and to be ordinal in nature (same linear levels hierarchy, with level 1 higher than 0, 2 higher than 1, and so on). Using this methodological procedure, both scales have the same properties, metrics and construction criteria.

In team sports or high-performance environments, fast decisions must be made every day to determine how an athlete will participate in subsequent training. On the one hand, measuring blood marker levels, neural fatigue or any other clinical variable on a daily basis in every player is impractical, expensive, time-consuming and uncomfortable for the athletes. On the other hand, subjective scales are a cheap, fast and easy alternative for identifying which athletes are able to perform. Therefore, the purpose of the present study was to adapt the previous idea proposed by Kenttä and Hassmén (1998) to create a single subjective method to assess both internal loads and subsequent fatigue. There is no such a scale in the literature, yet the development of such a scale would optimize staff members' decision making processes in daily practice because they would no longer need to use multiple different scales or ensure that all the athletes are acquainted with each separate scale.

Considering that soccer is a sport with a high number of official matches (60 to 70 annually), time for recovery is seldom sufficient (Lazarim et al., 2009), and it is therefore of utmost importance to optimize training and recovery. Hence, professional soccer represents a perfect environment for validating this new scale because in this sport, fast decisions must be made in a large number of players in daily practice. The TFS presented in this study simultaneously analyzes score oscillations from RPE_{TFS} and RPF_{TFS} (dose-response) and could help coaches and staff members identify which athletes undergoing the same training sessions have lower or higher fatigue levels relative to their perceived exertion. With this information, individual plans could be made for the athletes. Those with lower fatigue could be pushed a little harder during training sessions. On the other hand, it could support investigations aimed at determining the reasons for higher fatigue levels in those following a *light* training session.

Hence, the present study had three objectives: 1) to verify the reliability of the RPE_{TFS} and RPF_{TFS} scales; 2) to establish validity criteria for the RPE_{TFS} and RPF_{TFS} scales; and 3) to analyze the possible practical applications of the RPE_{TFS} and RPF_{TFS} scales in monitoring training sessions in professional soccer players.

Methods

This study was approved by the Ethics Committee (protocol number: 27609814.3.0000.5149 – 2014 Jun 06) and respected all guidelines of the National Health Council (Res. 466/12) regarding studies involving human subjects. All risks and benefits were properly explained to the volunteers before they provided written consent in accordance with the Declaration of Helsinki.

The methodological procedures followed in the present study were divided into two parts: reliability testing (part A) and determining criterion validity (part B).

Five similar training sessions (each starting at 16:00 during the third and fourth weeks of the preseason) were used to verify the reliability of the RPE_{TFS} and RPF_{TFS} scales. Twenty-two athletes rated their RPE_{TFS} 30 minutes after their afternoon training session. The RPF_{TFS} was rated the following morning at 09:00 prior to their morning training session. The training sessions consisted of large-sided games (65 x 68 m) involving two teams of 11 players each. The main objective was to score the highest number of goals with no more than two consecutive touches of the ball by each player. These training sessions were divided into five 10-min games with four 3-min passive recovery intervals for a total training time of 60 minutes.

After reliability was verified, the validity of the RPE_{TFS} and RPF_{TFS} scales was determined during the first month of the competitive season. Because RPE measures the internal training load (Borg, 1998), criterion validity was evaluated using other internal training loads, such as the mean heart rate (HR_{mean}), the percentage of the maximum heart rate ($\%HR_{max}$), the percentage of the heart rate that corresponded to the onset of blood lactate accumulation ($\%HR_{OBLA}$), the training impulse created by Banister (1991) ($TRIMP_B$), the training impulse modified by Stagno et al. (2007) ($TRIMP_{MOD}$) and the session rating of perceived exertion ($sRPE_{TFS}$). The total duration of the training session in minutes was also obtained.

These TRIMP measurements were used throughout the study as both are valid measures of the internal load (Borresen and Lambert, 2009). $TRIMP_B$ was created first and is used to measure the internal load of steady state aerobic activities using the heart rate reserve of the whole training session (Borresen and Lambert, 2009). $TRIMP_{MOD}$ indicates a modification of the original $TRIMP_B$ and is used in exercises that have higher intensity variations (Stagno et al., 2007). This variable was created based on the lactate response curve to increasing exercise intensity (Stagno et al., 2007). Therefore, the $TRIMP_{MOD}$ score was adjusted exponentially according to the number of minutes that were spent in different HR intensity zones (Stagno et al., 2007). Since soccer training can include both the aerobic steady state and intermittent activities, these two TRIMPs were chosen to cover most of the observed exercises.

Fifteen of the athletes participating in part A (reliability) were also used in part B (criterion validity) during nine training sessions, resulting in 135 data points. As in the reliability testing stage, all training sessions occurred in the afternoon (16:00), and the rating of subsequent fatigue was obtained the next morning before the training session (09:00). These nine training sessions consisted of at least one of the following routine soccer activities: a simulated game, offensive and defensive situations (2 vs 2, 3 vs 3), small-sided games, and tactical and technical training (crossing, heading and kicking). During all encounters included in parts A and B, players were wearing heart rate monitors, and hydration was provided *ad libitum.*

Participants

The sample size was calculated using an equation [1] obtained from a pilot study. Considering a significance level of 5% and a statistical power of 0.90, the correlation coefficients were used among the two scales (RPE_{TFS} and RPF_{TFS}) with $TRIMP_{MOD}$ and $TRIMP_B$ (Zar, 2010):

$$n = \left(\frac{Z_\beta + Z_\alpha}{\zeta_0}\right)^2 + 3, \qquad [1]$$

where "Z_β" represents the z-value corresponding to a β area of 0.1 (statistical power of 0.9), "Z_α" is the z-value corresponding to an area of significance of 0.05, and "ζ_0" is the Fisher's transformation equation [2] of Spearman's correlation coefficient:

$$\zeta_0 = 0.5 \ln\left(\frac{1+r}{1-r}\right). \qquad [2]$$

Fifteen individuals were required to achieve an adequate sample size. However, because of difficulties in conducting investigations in professional sports and the possible loss of volunteers, twenty-two professional male athletes from a first division Brazilian soccer team (age, 27.01 ± 4.38 years; body height, 178.36 ± 6.34 cm; body mass, 69.82±6.66 kg; body fat content, 9.86 ± 3.03% and maximum oxygen uptake, 54.24 ± 2.57 $mLO_2 \cdot kg^{-1} \cdot min^{-1}$) participated in part A of the present study. However, due to losses resulting from injury or absence, fifteen of the original 22 athletes (age, 27.75 ± 4.84 years; body height, 178.20 ± 6.29 cm; body mass, 68.91 ± 7.40 kg; body fat content, 9.36 ± 2.46% and maximum oxygen uptake, 54.37 ± 2.57 $mLO_2 \cdot kg^{-1} \cdot min^{-1}$) were available to participate in part B.

The following inclusion criteria were applied: soccer players (except goalkeepers) from a first division Brazilian soccer team and regular participation by athletes in official competitions organized by the Brazilian Soccer Confederation.

The exclusion criterion was absence from any of the monitored training sessions conducted to evaluate reliability or criterion validity.

Design and procedures

In the first week, exercise testing and anthropometric measurements were performed in the morning (08:00-09:00). Measurements included body mass, body height, body fat content, maximal oxygen uptake (VO_{2max}) and the onset of blood lactate accumulation (OBLA). During the 10 training sessions observed in this first week, the athletes were familiarized with both scales (RPE_{TFS} and RPF_{TFS}) of the TFS, and all participants received explanations of the experimental protocol.

VO_{2max} was evaluated indirectly using a soccer-specific field test (Yo-Yo Endurance Test) (Bangsbo, 1994) in which the total distance covered is positively related with VO_{2max} (Castagna et al., 2006). OBLA was assessed using a field test (Coelho et al., 2015). Two to five 1000 m runs were performed at an initial speed of 10 km/h. Sixty to 90 seconds after finishing the run, 25 μL of capillary blood was collected from the fingertip to measure blood lactate concentration (Accusport®). The test was considered completed once the blood lactate level reached or surpassed 4 mM. If lactate levels did not exceed this amount, one more run was performed at a speed 1 km/h faster than the previous one. HR_{OBLA} was obtained by linear interpolation (Microsoft Excel®). Temperature and relative humidity data were obtained from the Meteorological Department of the Nuclear Technology Development Center. The environmental conditions recorded during the Yo-Yo Endurance Test and OBLA measurements were as follows: 24.98 ± 2.45°C and 49.76 ± 12.94% relative air humidity, 23.45 ± 2.93°C and 55.17 ± 12.87% relative air humidity, respectively.

The heart rate was measured and recorded continuously using a heart rate monitor (Polar®, Team System 2®, Finland) during all tests and training sessions. A player's maximum HR was the highest HR recorded during the study (Antonacci et al., 2007). To estimate energy expenditure (kcal) during the training sessions, a linear relationship between HR and VO_{2max} was applied according to methods previously used in professional soccer players (Coelho et al., 2012).

To establish reliability and criterion validity, heart rate monitors were delivered to the athletes 30 minutes before and retrieved at the end of each training session. Since conflicting results have been reported in previous studies (Fanchini et al., 2016; McLaren et al., 2016), the RPE_{TFS} data were collected 30 minutes after the end of the training session in a separate room with athletes answering the question "How was your perceived exertion in this training session?". The next morning, in a separate room and before the beginning of the training session, the athletes were instructed to rate their perceived fatigue (RPF_{TFS} scale) by answering the question "How are you feeling now?".

Initially, regarding the construction of the RPE_{TFS} and RPF_{TFS} scales, both presented only five levels and all possessed verbal anchors (the same as those in the final outcome). This first format was used during two consecutive years (seasons) to verify the verbal anchor of the highest incidence, resulting in *medium* for RPE_{TFS} and *tired* for RPF_{TFS}. Therefore, the final outcomes of the RPE_{TFS} and RPF_{TFS} scales presented these verbal anchors in the middle (level 5).

The CR10 scale was used as a reference to design an RPE_{TFS} scale with levels from 0 to 10. To make the scale ordinal (linear level hierarchy) instead of psychophysical, the verbal anchors were arranged in a linear fashion and separated by one blank level each. After level 5 was selected as *medium*, levels 3, 1, 7 and 9 were named *light*, *very light*, *hard* and *maximal*, respectively. This design was maintained to ensure the inclusion of extreme levels (0 and 10) without verbal anchors. According to Borg (1998), it is possible to have unexpected perceptions of exertion that are higher than the maximum already experienced.

Therefore, these situations could be rated as 10 (above maximum) and could then be updated to represent the athlete's new maximum. A level of 0 was used when there was no activity (i.e., a day off or treatment due to an injury).

Based on the previous idea proposed by Kenttä and Hassmén (1998) and the above statements, an RPF_{TFS} scale was created that had the same ordinal design as the RPE_{TFS} scale. The verbal anchor *tired* was placed at the middle (level 5), and *normal*, *rested*, *very tired* and *exhausted* were included as levels 3, 1, 7, and 9, respectively. As discussed previously, extreme levels (0 and 10) were left blank and were updated when a higher or lower perception of fatigue was reported in comparison with previous experiences.

Borg (1998) states that volunteers with a low educational level may present some limitations and have difficulty in understanding standard scales. Brazilian soccer players were identified to have low economic and educational profiles, and this sport is ranked as 31st out of a total of 33 regarding the scholarship of their athletes. Therefore, as described in the previous literature (Serafim et al., 2014), chromatic colors were added to the TFS. They were placed in ascending thermal order from the lowest to the highest perception of effort/fatigue. The lowest levels were associated with the primary color blue (cold), and the highest levels were associated with the primary color red (hot). The colors associated with the lowest to the highest levels were connected by gradually changing secondary colors so that there was a continuous spectrum from cold (minimum) to hot (maximum) (Figure 1).

Statistical Analysis

Descriptive analyses of HR_{mean}, $\%HR_{max}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$, RPE_{TFS} and RPF_{TFS} were performed, and the results were expressed as the mean, standard deviation (SD) and coefficients of variation (CV). Reliability (Part A) was calculated using the intraclass correlation coefficient ($ICC_{3,1}$) and standard error of measurement (SEM) (Weir, 2005).

The Shapiro-Wilk test was used to verify the normality of the data. Descriptive analyses of $\%HR_{max}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$, RPE_{TFS} and RPF_{TFS} were performed, and the results were expressed as the mean and SD. Spearman's correlation coefficients followed by a multiple stepwise linear regression were used to verify any correlations among the variables listed above (Part B).

To validate the $sRPE_{TFS}$ obtained from the RPE_{TFS}, Spearman's correlation coefficients were calculated to verify any correlations among RPE_{TFS}, $TRIMP_B$ and $TRIMP_{MOD}$, and a simple linear regression was performed for each TRIMP (Part B).

After the validation process, statistical procedures were performed to verify any practical applications of the TFS. For this purpose, RPE_{TFS} and RPF_{TFS} ratings were divided into three groups (below median, median and above median) to perform a chi-square test (χ^2) of associations. Afterwards, a factorial exploratory analysis (FEA) was used to identify the number of factors and the distributions of the variables. A principal component analysis was used with varimax rotation, Bartlett's sphericity (BTS) and Kaiser-Meyer-Olkin (KMO) descriptive tests. Hair et al. (2013) proposed that five observations for each variable analyzed was the minimum number of observations that should be required to use FEA, although 10 observations per variable were recommended. Our study included 135 observations, with almost 14 observations per variable, and this was higher than the proportion suggested in the literature (Hair et al., 2013).

All statistical analyses were performed using SPSS software version 18.0 (SPSS Inc., Chicago IL), and the significance level adopted was $p < 0.05$.

Results

Part A

Table 1 presents the descriptive analyses and reliability of HR_{mean}, $\%HR_{max}$, $TRIMP_B$, $TRIMP_{MOD}$, RPE_{TFS}, RPF_{TFS} and $sRPE_{TFS}$ during small-sided games (24.58 ± 2.59°C and 51.79 ± 12.20 relative air humidity).

Part B

To establish criterion validity, the mean intensity of the training sessions was 73.80 ± 9.83% of the HR_{max}. The mean $TRIMP_B$ and $TRIMP_{MOD}$ were 81.76 ± 52.24 and 132.00 ± 91.14 arbitrary units, respectively. For the subjective measurements, the mean RPE_{TFS}, RPF_{TFS} and $sRPE_{TFS}$ were 5.21 ± 1.32, 4.45 ± 0.91 and 323.07 ± 199.63, respectively.

Table 2 presents the Spearman's

correlation coefficients for RPE_{TFS} with HR_{mean}, $\%HR_{max}$, $\%HR_{OBLA}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$ and RPF_{TFS} (23.44 ± 2.81°C and 62.43 ± 13.83 relative air humidity).

Table 3 presents the Spearman's correlation coefficients for RPF_{TFS} with HR_{mean}, $\%HR_{max}$, $\%HR_{OBLA}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$ and RPE_{TFS}.

Multiple stepwise linear regression revealed that 56.7% of the adjusted variance of the RPE_{TFS} could be explained by $TRIMP_B$ ($p < 0.001$). In addition, at least 56% of the adjusted variance of the RPF_{TFS} reported the next morning could be explained solely by the RPE_{TFS} or the $TRIMP_{MOD}$ ($p < 0.001$) of the previous training session. Taken together, the RPE_{TFS} and $TRIMP_{MOD}$ explained 65.5% of the adjusted variance of the RPF_{TFS} recorded the following morning ($p < 0.001$).

The Spearman's correlation coefficients for the $sRPE_{TFS}$ obtained from the RPE_{TFS} scale with $TRIMP_B$ (r = 0.90) and $TRIMP_{MOD}$ (r = 0.92) were statistically significant ($p < 0.01$) for all individuals and ranged from 0.81 to 0.99 and from 0.79 to 0.98, respectively. A simple linear regression showed that 84.8% of the variance of the $TRIMP_B$ ($p < 0.001$) and 78.1% of the variance of the $TRIMP_{MOD}$ ($p < 0.001$) could be explained by $sRPE_{TFS}$.

Table 4 presents the results of the chi-square test ($\chi^2 = 74.28$, $p < 0.001$) used to verify any statistical associations between RPE_{TFS} and RPF_{TFS}.

The results of the BTS and KMO descriptive tests used in the factorial exploratory analysis were satisfactory (0.872 and 0.001, respectively). In total, 10 components were identified, and the first two explained 89.76% of the accumulated variance. Table 5 presents the distribution of the factorial loads for HR_{mean}, $\%HR_{max}$, $\%HR_{OBLA}$, $TRIMP_B$, energy expenditure, $TRIMP_{MOD}$, RPE_{TFS}, RPF_{TFS}, session duration and $sRPE_{TFS}$ between these two components.

It should be noted that each component had at least three variables with a factorial load above 0.3 (the minimum requirement based on previous research) (Hair et al., 2013). Additionally, $TRIMP_B$, energy expenditure, $TRIMP_{MOD}$, session duration and $sRPE_{TFS}$ had higher factorial loads in the first component than in the second. Activity duration had a substantial impact on all of these variables. In addition, session duration presented the lowest factorial load in component 2 and was even lower than that recommended by the literature, indicating very little association (Hair et al., 2013).

The variables of intensity (HR_{mean}, $\%HR_{max}$ and $\%HR_{OBLA}$) presented the highest factorial loads in component 2. Conversely, in component 1, the same variables had factorial loads close to the minimum recommended (0.3) (Hair et al., 2013). Therefore, component 1 was designated "volume" and component 2 "intensity". Interestingly, RPE_{TFS} and RPF_{TFS} showed higher factorial loads in component 1 than in component 2, suggesting that training duration had a greater impact than intensity on these subjective ratings when measured by HR_{mean}, $\%HR_{max}$ and $\%HR_{OBLA}$.

Table 1

Descriptive analysis and reliability (95% Confidence Intervals) during small-sided games; p < 0.001 for all ICC results

Variable	Mean	SD	CV	$ICC_{(3,1)}$	SEM
HR_{mean}	152.55	8.33	0.05	0.89 (0.81 – 0.95)	0.27 (0.01 – 3.63)
$\%HR_{max}$	79.44	4.34	0.05	0.89 (0.81 – 0.95)	0.27 (0.01 – 1.89)
$TRIMP_B$	163.31	19.81	0.12	0.85 (0.74 – 0.92)	0.29 (0.03 – 10.10)
$TRIMP_{MOD}$	168.78	22.55	0.13	0.87 (0.78 – 0.94)	0.28 (0.03 – 10.58)
RPE_{TFS}	5.22	0.54	0.10	0.74 (0.59 – 0.86)	0.30 (0.04 – 0.35)
RPF_{TFS}	4.88	0.60	0.12	0.77 (0.63 – 0.88)	0.30 (0.04 – 0.36)
$sRPE_{TFS}$	324.09	35.60	0.11	0.63 (038 – 081)	0.23 (0.05 – 28.03)

SD = standard deviation; CV = coefficient of variation; ICC = intraclass correlation coefficient; SEM = standard error of measurement; HR_{mean} = mean heart rate (bpm); $\%HR_{max}$ = percentage of maximum heart rate; $TRIMP_B$ = training impulse from Banister et al. (1991); $TRIMP_{MOD}$ = training impulse from Stagno et al. (2007); RPE_{TFS} = rating of perceived exertion; RPF_{TFS} = rating of perceived fatigue; $sRPE_{TFS}$ = session RPE.

Table 2

Spearman's correlation coefficients for RPE_{TFS} with HR_{mean}, $\%HR_{max}$, $\%HR_{OBLA}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$ and RPF_{TFS}

Subject	HR_{mean}	$\%HR_{max}$	$\%HR_{OBLA}$	$TRIMP_B$	$TRIMP_{MOD}$	$sRPE_{TFS}$	RPF_{TFS}
S1	0.65	0.65	0.79*	0.84**	0.85**	0.86**	0.60
S2	0.69*	0.69*	0.95**	0.95**	0.95**	0.95**	092**
S3	0.84**	0.80**	0.91**	0.86**	0.86**	0.88*	0.87**
S4	0.41	0.41	0.41	0.82**	0.94**	0.91**	0.86**
S5	0.85**	0.85**	0.80**	0.80**	0.80**	0.79*	0.55
S6	0.82**	0.82**	0.82**	0.93**	0.96**	0.96**	0.64
S7	0.64	0.64	0.64	0.91**	0.86**	0.98**	0.84**
S8	0.37	0.61	0.18	0.70*	0.54	0.75*	0.68*
S9	0.89**	0.90**	0.89**	0.98**	0.92**	0.97**	0.88*
S10	-0.13	0.16	0.14	0.76*	0.76*	0.85**	0.66
S11	0.86**	0.86**	0.85**	0.94**	0.90**	0.94**	0.79*
S12	0.55	0.55	0.55	0.73*	0.82**	0.83**	0.98**
S13	0.61	0.61	0.61	0.63	0.74*	0.76*	0.91**
S14	0.81**	0.88**	0.88**	0.95**	0.95**	0.93**	0.96**
S15	0.61	0.61	0.74*	0.84**	0.84**	0.90**	0.87**
Mean	0.57**	0.64**	0.63**	0.77**	0.76**	0.88**	0.74**

*HR_{mean} = mean heart rate (bpm); $\%HR_{max}$ = percentage of maximum heart rate; $\%HR_{OBLA}$ = percentage of the heart rate corresponding to the onset of blood lactate accumulation; $TRIMP_B$ = training impulse from Banister et al. (1991); $TRIMP_{MOD}$ = training impulse from Stagno et al. (2007); $sRPE_{TFS}$ = session RPE; RPF_{TFS} = rating of perceived fatigue; *statistically significant at $p < 0.05$; **statistically significant at $p < 0.01$.*

Table 3

Spearman's correlation coefficients for RPF_{TFS} with HR_{mean}, $\%HR_{max}$, $\%HR_{OBLA}$, $TRIMP_B$, $TRIMP_{MOD}$, $sRPE_{TFS}$ and RPE_{TFS}

Subject	HR_{mean}	$\%HR_{max}$	$\%HR_{OBLA}$	$TRIMP_B$	$TRIMP_{MOD}$	$sRPE_{TFS}$	RPE_{TFS}
S1	0.79*	0.79*	0.79*	0.84**	0.79*	0.74*	0.60
S2	0.50	0.50	0.79*	0.84**	0.90**	0.90**	0.92**
S3	0.82**	0.77*	0.83**	0.88**	0.83**	0.90**	0.87**
S4	0.49	0.49	0.49	068*	0.84**	0.76*	0.86**
S5	0.26	0.26	0.34	0.41	0.45	0.58	0.55
S6	0.60	0.60	0.60	0.77*	0.70*	0.70*	0.64
S7	0.71*	0.71*	0.71*	0.92**	0.90**	0.90**	0.84**
S8	0.83**	0.95**	0.76*	0.95**	0.84**	0.81**	0.68*
S9	0.85**	0.84**	0.85**	0.83**	0.77*	0.82**	0.88**
S10	-0.17	0.20	0.14	0.71*	0.71*	0.71*	0.66
S11	0.61	0.61	0.63	0.84**	0.79*	0.79*	0.79*
S12	0.55	0.55	0.55	0.73*	0.82**	0.83**	0.98**
S13	0.82**	0.82**	0.82**	0.82**	0.82**	0.73*	0.91**
S14	0.80*	0.85**	0.85**	0.94**	0.94**	0.94**	0.96**
S15	0.64	0.64	0.83**	0.73*	0.82**	0.73*	0.87**
Mean	0.56**	0.62**	0.59**	0.74**	0.75**	0.76**	0.74**

*HR_{mean} = mean heart rate (bpm); $\%HR_{max}$ = percentage of maximum heart rate; $\%HR_{OBLA}$ = percentage of the heart rate corresponding to the onset of blood lactate accumulation; $TRIMP_B$ = training impulse from Banister et al. (1991); $TRIMP_{MOD}$ = training impulse from Stagno et al. (2007); $sRPE_{TFS}$ = session RPE; RPE_{TFS} = rating of perceived exertion; *statistically significant at $p <0.05$; **statistically significant at $p < 0.01$.*

Table 4

Results of the Chi-Square association test.

RPE_{TFS}		RPF_{TFS}			TOTAL
		Rested	Tired	Very tired	
Light	Observed frequency	41 (95.5%)	2 (4.5%)	0 (0.0%)	43 (100%)
	Expected frequency	24.8	11.8	6.4	43.0
	Adjusted residual	6.0	-4.1	-3.3	
Medium	Observed frequency	29 (70.73%)	12 (29.27%)	0 (0.0%)	41 (100%)
	Expected frequency	23.7	11.2	6.1	41.0
	Adjusted residual	2.0	0.3	-3.2	
Hard	Observed frequency	8 (15.69%)	23 (45.09%)	20 (39.22%)	51 (100%)
	Expected frequency	29.5	14.0	7.6	51.0
	Adjusted residual	-7.7	3.6	6.2	
Total	Observed frequency	78	37	20	135
	Expected frequency	78.0	37.0	20.0	135.0

RPE_{TFS} = rating of perceived exertion; RPF_{TFS} = rating of perceived fatigue.

Table 5

Distribution of the factorial loads between the two components

Variables	Component 1	Component 2
HR_{mean}	0.312	**0.931**
$\%HR_{max}$	0.380	**0.916**
$\%HR_{OBLA}$	0.371	**0.895**
$TRIMP_B$	**0.831**	0.515
Energy expenditure	**0.835**	0.476
$TRIMP_{MOD}$	**0.831**	0.451
RPE_{TFS}	**0.771**	0.357
RPF_{TFS}	**0.781**	0.317
Session Duration	**0.924**	0.244
$sRPE_{TFS}$	**0.949**	0.269

The highest factorial load in each component is presented in bold. HR_{mean} = mean heart rate (bpm); $\%HR_{max}$ = percentage of maximum heart rate; $\%HR_{OBLA}$ = percentage of the heart rate corresponding to the onset of blood lactate accumulation; $TRIMP_B$ = training impulse from Banister et al. (1991); $TRIMP_{MOD}$ = training impulse from Stagno et al. (2007); RPE_{TFS} = rating of perceived exertion; RPF_{TFS} = rating of perceived fatigue; $sRPE_{TFS}$ = session RPE.

	0		0
VERY LIGHT	1	RESTED	1
	2		2
LIGHT	3	NORMAL	3
	4		4
MEDIUM	5	TIRED	5
	6		6
HARD	7	VERY TIRED	7
	8		8
MAXIMUM	9	EXHAUSTED	9
	10		10

Figure 1

Training-fatigue scale composed of the RPE_{TFS} (left) and the RPF_{TFS} (right) scales.

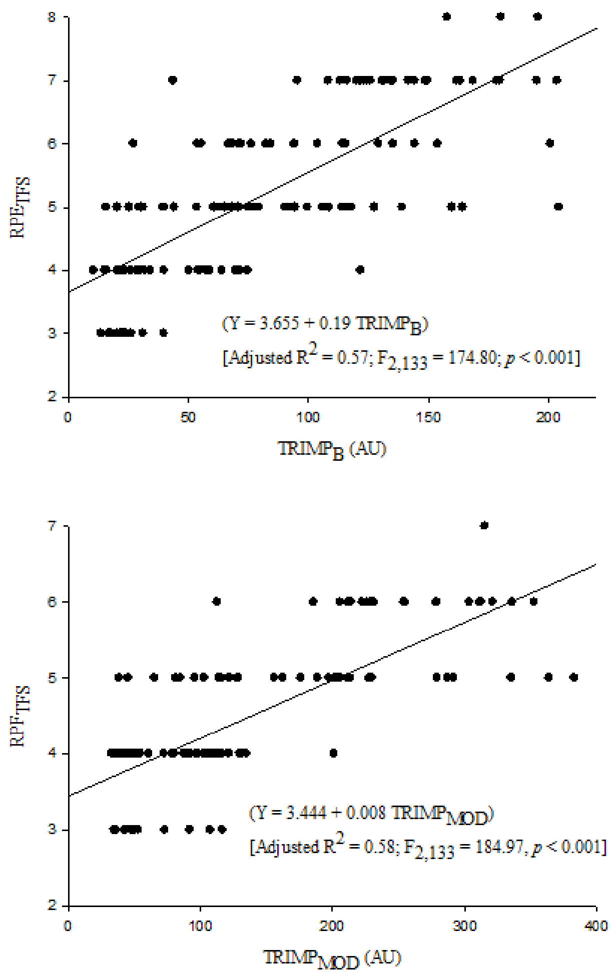

Figure 2

Multiple stepwise linear regression between RPE_{TFS} and $TRIMP_B$ (first) and between RPF_{TFS} and $TRIMP_{MOD}$ (second).

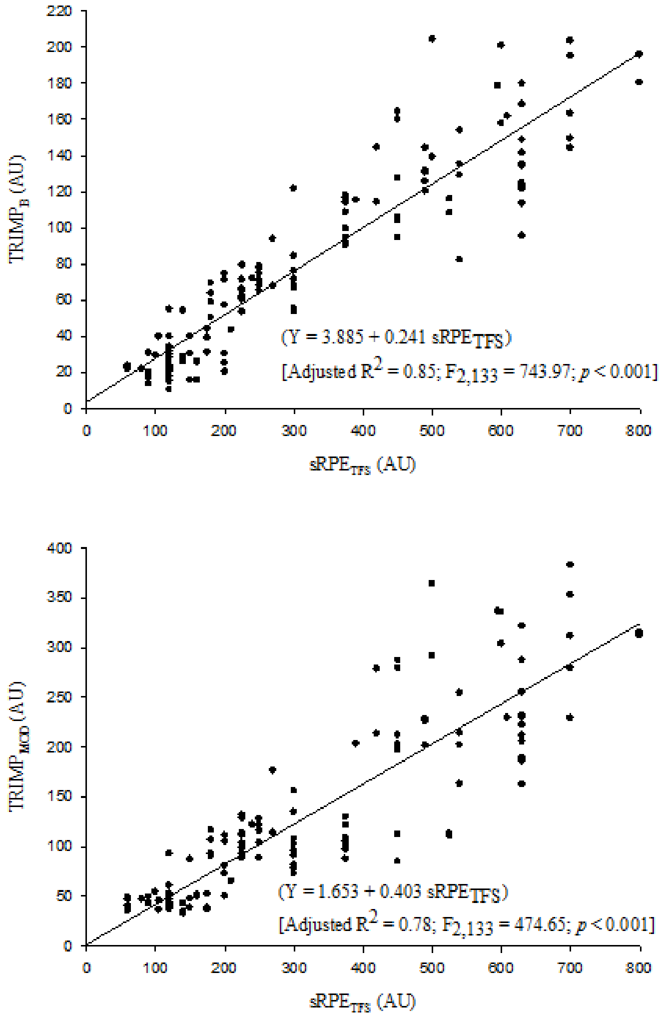

Figure 3

Simple linear regression between $TRIMP_B$ and $sRPE_{TFS}$ (first) and between $TRIMP_{MOD}$ and RPF_{TFS} (second).

Discussion

The aim of the present study was to validate the RPETFS and RPFTFS scales from the TFS for Brazilian professional soccer players. Both scales were verified to be reliable and valid, and RPETFS and RPFTFS were best explained by the internal training load TRIMPB and TRIMPMOD, respectively ($p < 0.01$). An association was also found between both scales (RPETFS and RPFTFS), with training duration having a greater impact than training intensity on these subjective perceptions ($p < 0.01$).

Despite the recommendation (Weir, 2005) that ICCs and SEMs should be used to determine reliability, different methods are used for this purpose in CVs, Pearson's correlation coefficients and Bland-Altman plots. For example, the reliability of the Borg RPE scale (Borg, 1998) and the CR10 scale was verified by Pearson's correlation coefficients (Borg, 1998). While Scott et al. (2013) and Wallace et al. (2014) showed ICC values similar to those found in the present study

(0.66 and 0.78, respectively), they supported their findings related to the reliability of the CR10 scale using %CV values. This makes it difficult to compare the results of the present investigation with those obtained in the previously cited studies.

Moreover, another study (Haddad et al., 2013) used ICCs, CVs and SEMs to measure CR10 reliability in junior soccer players. They found the CR10 scale to be reliable and obtained ICC and SEM values similar to those obtained in the present study (0.77 and 2.24, respectively). Therefore, considering the fact that there is no single method for verifying the reliability of subjective scales (Scott et al., 2013), the multifactorial nature (psychological and physiological) of the RPE scale (Borg, 1998; McLaren et al., 2016), the similarities between the findings of the present study and those in the literature (Haddad et al., 2013; Scott et al., 2013; Wallace et al., 2014), the reliability recommendations of Weir (2005) and the lack of this measure in some papers (Laurent et al., 2011; Sikorski et al., 2013), the authors of the present study view both the RPETFS and RPFTFS scales as reliable.

After reliability was determined, criterion validity was assessed. As expected, RPETFS had the highest correlations with sRPETFS, TRIMPB and TRIMPMOD (0.88, 0.77 and 0.76, respectively). sRPETFS is calculated by multiplying the RPETFS score by the total training duration. TRIMPB and TRIMPMOD are considered measures of the internal training load (Nakamura et al., 2010), and they take into account both the intensity (HR) and the duration (min) of the training session. Because perceived exertion is also a variable of the internal training load that considers the intensity and duration of the training session (Borg, 1998), the findings in the present study are well justified.

However, it must be noted that the RPETFS found in approximately eight individuals was not significantly correlated with their HR_{mean}, $\%HR_{max}$ and %HROBLA. A previous meta-analysis (Chen et al., 2002) demonstrated that correlations between the HR and perceived exertion were not as high as expected. One possible explanation for this finding is that training sessions for this sport are characterized by intermittent exercises in which both aerobic and anaerobic energetic pathways are active (Bangsbo, 1994). Therefore, the anaerobic contribution may be associated with ratings of perceived exertion that are not reflected by changes in the HR (Impellizzeri et al., 2004). This does not necessarily mean that the RPE scale is not valid under these conditions. In contrast, it has been suggested that the HR expressed as HR_{mean}, $\%HR_{max}$ or %HROBLA may not be suitable for measuring internal training loads under specific soccer training conditions (Nakamura et al., 2010).

The findings reported for RPFTFS were similar to those for RPETFS and resulted in higher correlations with sRPETFS, TRIMPB and TRIMPMOD (0.76, 0.74 and 0.75, respectively). Again, in eight individuals, RPFTFS showed no significant correlations with HR_{mean}, $\%HR_{max}$ or %HROBLA. Since Spearman's correlation coefficients and multiple regression analysis showed that the previous RPETFS had an impact on the RPFTFS reported the next morning, the same point that was previously made regarding RPETFS can be extrapolated to explain the RPFTFS findings.

In an attempt to develop a single rating for perceived exertion that corresponds to the whole training session, Foster et al. (1995) created the sRPE. This new variable was validated (Foster et al., 2001) and found to be very easy and practical to use, and it therefore became extremely widespread in sports (Impellizzeri et al., 2004; Nakamura et al., 2010; Scott et al., 2013).

In the present investigation, we also aimed to validate the use of sRPE from the RPETFS scale. As expected, the correlations among sRPETFS, TRIMPB and TRIMPMOD were higher (0.92 and 0.90, respectively) and even slightly higher than those reported in other studies (Campos-Vazquez et al., 2015; Impellizzeri et al., 2004) that used sRPE from the CR10 scale. In addition, a simple linear regression analysis demonstrated high predictability for TRIMPB and TRIMPMOD values obtained from our sRPETFS measurements (84.8% and 78.1% of the variance, respectively). Collectively, these results show that using the sRPETFS from the RPETFS scale is a valid method for monitoring soccer training sessions.

After we verified the reliability and validity of the RPETFS and RPFTFS scales obtained from the TFS, we explored the third objective of this study, which was to analyze any potentially practical implications for the proposed scales in monitoring soccer training. Chi-square association tests showed that 4.5% of the athletes found themselves feeling tired (RPFTFS = 5) on the next morning (15 h later) despite rating their previous training session as light (RPETFS < 5, adjusted residual = -4.1). This finding suggests that for some reason, some athletes did not achieve an adequate recovery. Recovery has been shown to depend on many factors (e.g., age, quality of sleep, adequate food and water intake, and stress). Therefore, the oscillations observed in the scores obtained from these two scales (dose-response) might help coaches and staff members make better decisions regarding individual training plans.

Another interesting result was observed when the RPETFS was considered hard (RPETFS > 5). When ratings were assessed the next morning (15 h later), eight of the athletes felt rested (RPFTFS < 5, adjusted residual = -7.7), 23 felt tired (RPFTFS = 5, adjusted residual = 3.6), and another 20 felt very tired (RPFTFS > 5, adjusted residual = 6.2). The results showed that intermittent training sessions might require 72 h for a full recovery (Twist and Eston, 2005). Furthermore, genetics may also play a role in individual recovery kinetics after plyometric training (Pimenta et al., 2011). Therefore, oscillations in RPETFS and RPFTFS scores could potentially be used to subjectively rank the athletes according to their ability to perform the next training session and determine which athletes still need more time to recover. With that in mind, whenever a training session is considered hard by the athletes, coaches can anticipate that most of them will feel tired if the recovery time is less than 15 h. Hence, the next training session could be adjusted accordingly or rescheduled for the afternoon to allow more time to recover. Altogether, the findings of the present study reveal a practical, cheap, fast and easy method for monitoring soccer training sessions when other technological procedures are not available.

In an attempt to better understand how RPETFS and RPFTFS are correlated with monitored variables, a factorial exploratory analysis was performed. Total training duration was found to have a greater impact than intensity on both scales. One possible explanation is that soccer-specific training tends to have an intensity similar to that of an official match (85% of HR_{max}) (Coelho et al., 2012). This notion is supported by the results of the present study, which showed that the CVs obtained for training intensity, expressed as $\%HR_{max}$, were lower than the CVs obtained for the total duration of the training sessions (0.13 and 0.44, respectively).

Since variability was lower for training intensity than for training duration, greater energy expenditure was required during longer training sessions. This observation is supported by the finding that the factorial load was higher for the "volume component" of energy expenditure than for its "intensity" component (0.835 and 0.476, respectively). This suggests that in comparison to short sessions, long training sessions (with a higher energy expenditure) lead to a higher subjective level of fatigue (Marcora, 2009; Noakes, 2012), resulting in higher RPETFS after training and higher RPFTFS the following morning. With this information in mind, during the preseason, in which at least two training sessions per day are performed for several days with a reduced recovery time, coaches can plan their training sessions with a focus on attenuating the athlete's subjective ratings of fatigue. Coaches can benefit from planning intense, short training sessions instead of long, moderate sessions because they produce smaller reductions in the athletes' willingness to train during this period and increase the probability of maintaining training quality in subsequent days.

However, the subjective nature of the variables must be considered. Regardless of the scale used, caution must be taken to avoid being manipulated when only subjective variables are used. Therefore, it is always recommended that subjective and objective variables should be used together to achieve better accuracy when monitoring and planning training sessions.

In summary, the results of this study indicate that both the RPETFS and RPFTFS scales of the TFS are reliable and valid for monitoring training sessions in Brazilian professional soccer players. Oscillations observed in the RPETFS and RPFTFS scores can be used by coaches to better plan their weekly programs based on dose-response ratings. Additionally, we found that subjective perceptions are more influenced by training duration than training intensity.

Acknowledgements

The authors thank the athletes involved in this study and CAPES, Fapemig, Fapema and Ministério do Esporte for funding this research.

References

Antonacci L, Mortimer LF, Rodrigues VM, Coelho DB, Soares DD, Silami-Garcia E. Competition, estimated, and test maximum heart rate. *J Sport Med Phys Fit*, 2007; 47: 418–421

Bangsbo J. The physiology of soccer--with special reference to intense intermittent exercise. *Acta Physiol Scand*, 1994; 619: 1–155

Banister E. Modeling elite athletic performance. In J. D. MacDougall, H. A. Wenger, & H. J. Green (Eds.), *Physiological testing of the high-performance athlete.* Champaign, IL: Human Kinetics Books, 403–425; 1991

Borg G. *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics; 1998

Borresen J, Lambert MI. The quantification of training load, the training response and the effect on performance. *Sports Med*, 2009; 39: 779–795

Campos-Vazquez MA, Mendez-Villanueva A, Gonzalez-Jurado JA, León-Prados JA, Santalla A, Suarez-Arrones L. Relationships between rating-of-perceived-exertion- and heart- rate-derived internal training load in professional soccer players: A comparison of on-field integrated training sessions. *Int J Sport Physiol*, 2015; 10: 587–592

Castagna C, Impellizzeri FM, Chamari K, Carlomagno D, Rampinini E. Aerobic fitness and yo-yo continuous and intermittent tests performances in soccer players: A correlation study. *J Strength Cond Res*, 2006; 20: 320–325

Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: A meta-analysis. *J Sports Sci*, 2002; 20: 873–899

Coelho D, Coelho LG, Mortimer LA, Hudson ASR, Marins JCB, Soares DD, Silami-Garcia E. Energy demand and heart rate evaluation at different phases during a match along an official soccer competition. *Rev Bras Cineantropomet Desemp Humano*, 2012; 14: 419–427

Coelho DB, Morandi RF, Paixão RC, Mendes TT, Borba DA, De-Barros CLM, Garcia ES. Anaerobic threshold of 4.0mM is able to estimate the maximal lactate steady state of soccer players on the field tests. *Rev Bras Ciênc Mov*, 2015; 23: 32–39

Fanchini M, Ferraresi I, Modena, R, Schena F, Coutts AJ, Impellizzeri FM. Use of the CR100 scale for session rating of perceived exertion in soccer and its interchangeability with the CR10. *Int J Sport Physiol*, 2016; 11: 388–392

Foster C, Florhaug JA, Franklin J, Gottschall L, Hrovatin LA, Parker S, Dodge C. A new approach to monitoring exercise training. *J Strength Cond Res*, 2001; 15: 109–115

Foster C, Hector LL, Welsh R, Schrager M, Green MA, Snyder AC. Effects of specific versus cross-training on running performance. *Eur J Appl Physiol O*, 1995; 70: 367–372

Haddad M, Chaouachi A, Wong DP, Castagna C, Hambli M, Hue O, Chamari K. Influence of fatigue, stress, muscle soreness and sleep on perceived exertion during submaximal effort. *Physiol Behav*, 2013; 119: 185–189

Hair J, Black W, Babin B, Anderson R. *Multivariate data analysis*. Upper Saddle River, NJ: Pearson Education Limited; 2013

Impellizzeri FM, Rampinini E, Coutts AJ, Sassi A, Marcora SM. Use of RPE-based training load in soccer. *Med Sci Sport Exer*, 2004; 36: 1042–1047

Kenttä G, Hassmén P. Overtraining and recovery. *Sports Med*, 1998; 26: 1–16

Laurent CM, Green JM, Bishop PA, Sjökvist J, Schumacker RE, Richardson MT, Curtner-Smith M. A practical approach to monitoring recovery: Development of a perceived recovery status scale. *J Strength Cond Res*, 2011; 25: 620–628

Lazarim FL, Antunes-Neto JMF, da Silva FOC, Nunes LAS, Bassini-Cameron A, Cameron L.C, Macedo DV. The upper values of plasma creatine kinase of professional soccer players during the Brazilian National Championship. *J Sci Med Sport*, 2009; 12: 85–90

Marcora S. Perception of effort during exercise is independent of afferent feedback from skeletal muscles, heart, and lungs. *J Appl Physiol*, 2009; 106: 2060–2062

McLaren SJ, Graham M, Spears IR, Weston M. The sensitivity of differential ratings of perceived exertion as measures of internal load. *Int J Sport Physiol*, 2016; 11: 404–406

Nakamura FY, Moreira A, Aoki MS. Training load monitoring: is the session rating of perceived exertion a reliable method? *Rev Ed Fis*, 2010; 21: 118-126

Noakes TD. Fatigue is a brain-derived emotion that regulates the exercise behavior to ensure the protection of whole body homeostasis. *Front Physiol*, 2012; 3: 1–13

Pimenta EM, Coelho DB, Cruz IR, Morandi RF, Veneroso CE, Azambuja P, Paz-Fernández JA. The ACTN3 genotype in soccer players in response to acute eccentric training. *Eur J Appl Physiol*, 2011; 112: 1495–1503

Scott TJ, Black CR, Quinn J, Coutts AJ. Validity and reliability of the session-RPE Method for quantifying training in australian football. *J Strength Cond Res*, 2013; 27: 270–276

Serafim THS, Tognato AC, Nakamura PM, Queiroga MR, PereiraG, Nakamura FY, Kokubun E. Development of the color scale of perceived exertion: Preliminary validation. *Percept Motor Skill*, 2014; 119: 884–900

Sikorski EM, Wilson JM, Lowery RP, Joy JM, Laurent CM, Wilson SMC, Gilchrist P. Changes in perceived recovery status scale following high-volume muscle damaging resistance exercise. *J Strength Cond Res*, 2013; 27: 2079–2085

Stagno KM, Thatcher R, van Someren KA. A modified TRIMP to quantify the in-season training load of team sport players. *J Sport Sci*, 2007; 25: 629–634

Twist C, Eston R. The effects of exercise-induced muscle damage on maximal intensity intermittent exercise performance. *Eur J Appl Physiol*, 2005; 94: 652–658

Wallace LK, Slattery KM, Impellizzeri FM, Coutts AJ. Establishing the criterion validity and reliability of common methods for quantifying training load. *J Strength Cond Res*, 2014; 28: 2330–2337

Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 2005; 19: 231–240

Zar JH. *Biostatistical analysis*. Upper Saddle River, NJ: Prentice Hall; 2010

Corresponding author:

Eduardo Macedo Penna,

Universidade Federal do Pará, Campus Universitário de Castanhal. Avenida dos Universitários, S/N, - Jaderlândia, Castanhal, Pará, Brasil. 68746-360

E-mail: eduardomp@ufpa.br