|  |  |
| --- | --- |
| Validity evidence of sample time series graphs for American Football athletes | For correspondence: marciotasinafojr@gmail.com |

Márcio Fernando Tasinafo Júnior1-2; Leonardo Santos Lopes da Silva3; Leonardo da Silva Gonçalves3; André Pereira dos Santos1-2,4-5; Pedro Pugliesi Abdalla2,6; and Dalmo Roberto Lopes Machado1-2,4.

1School of Physical Education and Sport of Ribeirão Preto (EEFERP/USP), University of São Paulo, Ribeirão Preto, Brazil;

2Study and Research Group in Antropometry, Training and Sport (GEPEATE), University of São Paulo, Ribeirão Preto, Brazil;

3Department of Internal Medicine, Ribeirão Preto Medical School (FMRP/USP), University of São Paulo, Ribeirão Preto, Brazil;

4Ribeirão Preto College of Nursing (EERP/USP), University of São Paulo, Ribeirão Preto, Brazil;

5Department of Child, Family and Population Health Nursing, (CFPHN), Seattle, United State of America.

6Education Network, Physical Education Course - Claretiano, Batatais, Brazil.

*Please cite as*: Tasinafo-Júnior, M.F., Da Silva, L. S., Gonçalves, L. S., et al. (2025). Validity evidence of sample time series graphs for American Football athletes. *SportRχiv*.

# ABSTRACT

American football is a high-contact sport characterized by intermittent bouts of intense physical activities. Effective monitoring of training stress is critical for optimizing athletic performance and minimizing injury risk in this demanding sport. Our objective in this observational study was to investigate the validity of evidence of Sample Time Series Graphs (STSG) as a multifactorial stress monitoring tool in Brazilian American football athletes. Thirty-eight athletes underwent anthropometric evaluations, and pre- and post-competition assessments, including STSG, Daily Analysis of Life Demands for Athletes (DALDA), Session Rating of Perceived Exertion (sRPE-TL), and motor performance tests. STSG's internal structure was verified through exploratory factor analysis and Cronbach's alpha. The significance level was set at (p<0.05). Concurrent validity was analyzed by correlating STSG items with DALDA, weekly sRPE-TL, monotony, and strain. Sensitivity was evaluated using generalized estimating equations (GEE) and generalized mixed models (GMM). Our findings showed that STSG demonstrated robust factorial structure (χ² = 12.4, df = 4, p = 0.02; RMSEA = 0.0724; BIC = -11.6; TLI = 0.93), adequate internal consistency (Cronbach’s α = 0.648), and relevant concurrent validity (ρ > 0.39; ρ < -0.41). Analyses indicated increased stress near competitions, primarily in muscle soreness and training readiness. STSG was proved as a valid and reliable tool for monitoring stress indicators in Brazilian American football athletes, potentially optimizing performance and well-being. Further research should explore the STSG’s applicability across diverse sports and competitive levels.

**Keywords:** American football; overtraining; training load control; validity evidence; psychometrics.

# INTRODUCTION

American football is an intense, continuous contact sport involving substantial impact forces exceeding 10g (gravitational force) and high-intensity intermittent activities with long recovery periods (41,52). To meet the demands of the sport, American football athletes undergo intense training loads (23,53). Individuals participating in exercise programs aim to develop their physical capacities, motor skills, and consequently improve their performance in specific motor tasks. However, it is now well understood that physical exercises can function as stress factors, disrupting the body's homeostasis. This disruption triggers a series of adaptive responses aimed at restoring physiological homeostasis and preventing future stress (45,49). At times, exercise-induced stress can exceed the available adaptive responses to restore homeostasis. If this stressor is intense, chronic, and without adequate recovery periods, athletes may develop what is known as overtraining syndrome, increasing the likelihood of injuries, and decreasing performance (5,35). Often, American football athletes are affected by both conditions (10,15,16,19,31).

To prevent overtraining and understand individual responses to training, it is essential to continuously quantify and monitor training load components (intensity, volume, frequency, recovery, and others) and their effects (3,7,9). Therefore, conducting formative assessments is important to ensure that training stimuli, recovery periods, nutrition, and leisure activities are adequate for these athletes (25,37). This control allows athletes and coaches to plan training periods, phases, and cycles to enhance performance and minimize the risk of overreaching, overtraining (7,8), injuries and excessive stress (14,20,32,40).

There are several instruments that measure stress indicators associated with overtraining and injuries, of physical, biochemical, psychological, immunological, behavioral, and physical nature (2,21,28). Most instruments present indicators of only one of these types; however, combining them is suggested for controlling and monitoring training loads (44,48). Nevertheless, many instruments have limited applicability for sports analysis practiced in some countries due to lack of validity evidence or excessive cost, as is the case in Brazil (1,47,54). Particularly, monitoring and controlling training of Brazilian Football Athletes still face several challenges. American football teams typically consist of numerous athletes (approximately 60), with limited financial resources and time to evaluate stress indicators for all players. In this case, an easy-to-use, rapid, and low-cost operational instrument capable of estimating various stress indicators (behavioral, physiological, and psychological) would be highly useful in the clinical practice of this sport.

Bompa and Haff (7,8) proposed a visually accessible and easy-to-use instrument to monitor training load components. This instrument, called Sample Time Series Graph (STSG)(7,8), encompasses indicators of various types related to training stress: behavioral indicators (sleep duration, readiness for training, and appetite); physiological indicators (muscle soreness, resting heart rate, and body mass); and psychological indicators (fatigue sensation, sleep quality, and competitive readiness). STSG were used over throughout a full season of competition in female soccer players (4), and the results indicate a tendency towards overload syndrome in the players during the season, highlighting the importance of using subjective and objective measures to monitor training load. However, the effectiveness of this tool remains unknown, particularly among Brazilian athletes. Furthermore, the STSG is an instrument that has not undergone the construction, translation, or cultural adaptation stages recommended by psychometric literature for effective validation (1,39). Therefore, it is necessary to investigate whether STSG shows validity evidence for monitoring training stress indicators. If the instrument does not demonstrate consistent validity evidence, its discontinuation or reconstruction would be advisable (1).

Therefore, this study aims to examine the validity evidence of the STSG. We will evaluate its construct validity by evaluating both the instrument's internal structure and its relationship with external measures, adhering to the guidelines proposed in specialized literature (1). Our initial hypothesis posits that the STSG proves to be a valid instrument for monitoring training stress in Brazilian American football athletes, supported by substantial evidence.

# METHOD

## **Experimental approach to the problem**

To verify the validity evidence of the STSG, we conducted a study with a longitudinal trial, with the following independent variables: Daily Analysis of Life Demands for Athletes (DALDA) questionnaire, training load indicators (calculated from session rating of perceived exertion [sRPE] and session time). Data collections occurred at two points, 15-week apart: seven days before the competitive period and immediately after the competition’s conclusion. The competitive season spanned from March 12 to July 2, 2023, with the team finishing second in the championship. During both week-long collections periods, we recorded daily STSG, sRPE, and training session time. On the final day of each period, participants completed the DALDA questionnaire. In February, prior to initial data collection, we conducted a 7-day familiarization phase with all instruments (STSG, DALDA, and PSE). We also collected sample characterization variables during this time. Figure 1 presents the flowchart of the data collection schedule. The STROBE checklist was used for this manuscript.

**Figure 1 -** Data collection flowchart.

Linha do tempo

Descrição gerada automaticamente

**Note:** DALDA = daily analysis of life demands for athletes; sRPE-TL= training load calculated from rate of perceived exertion.

## **Subjects**

Male amateur athletes, aged between 16 and 38 years, from the American football team of Ribeirão Preto-SP were participants in this study. All participants who met the elegibility criteria were included in the study (38 athletes out of a total of 58). Inclusion criteria were active team members who did not receive salaries (except for travel allowances) and signed the informed consent form. Athletes were excluded from this study if they withdrew, revoked their informed consent form, or failed to respond to questionnaires at least once during the study period. This study was approved by the research ethics committee of the School of Physical Education and Sport of Ribeirão Preto, University of São Paulo (CAEE: 51912321.0.0000.5659), in accordance with Resolution No. 466/12 of the National Health Council, 2012, which regulates research involving human subjects (equivalent to the Declaration of Helsinki in Brazil). Participants gave written consent for review and signature before starting the study. Permission was also obtained through the copyright clearance center for republication of the STSG from the Human Kinetics publisher (7,8).

## **Instruments**

### Body Measurements

For sample characterization, body mass and stature measurements were collected following Lohman et al. (29) standards. Body mass was measured in kilograms (kg) with gram accuracy using a Filizola® anthropometric scale (Personal model, maximum capacity of 180kg, Campo Grande, MS). Stature was recorded in meters (m) with centimeter accuracy using a Sanny® fixed wall stadiometer (Professional – ES2020). The athlete positioned themselves against the wall with upper limbs extended along the body, evenly distributed weight, and barefoot.

### Sample Time Series Graph (STSG)

The STSG required daily recording of resting heart rate in beats per minute (bpm) upon waking, changes in body mass (kg), and sleep duration (h), categorized as "no sleep", four to 11, or ≥12 hours per night. Qualitative variables were recorded using Likert-type scales, with 4 to 5 possible responses to assess sleep quality, sensation of fatigue, readiness for training, appetite, competitive readiness, and muscle soreness. Lower scores on the Likert-type items of STSG indicate higher training stress, as demonstrated in Table 1. Resting heart rate values were measured using finger oximeters (G-Tech, LED MD300C19), and for body mass measurement, players were asked to use the same weighing scale and always at the same time of day.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table 1** - Table of scores for the STSG likert items, adapted from Baghaei et al (2022). | | | | | |
|  | **Item 1 (score = 5)** | **Item 2 (score = 4)** | **Item 3 (score = 3)** | **Item 4 (score = 2)** | **Item 5 (score = 1)** |
| **Quality of sleep** | Very deep | Normal | Restless | Bad with breaks | No sleep |
| **Tiredness sensation** | Very rested | Normal | Tired | Very tired | Painful |
| **Training willigness** | Very deep | Good | Poor | Unwilling | Did not train |
| **Appetite** | Very good | Good | Poor | Ate because should | Did not eat |
| **Competitive willigness** | High | Average | Low | None at all | - |
| **Muscle Soreness** | No pain | Little pain | Moderate | Severe pain | - |
|  | | | | | |

During the familiarization week, athletes received detailed instructions on how to fill out the STSG. Daily alerts were set up on their cell phones to ensure regular completion. The graphs were individually filled out without external influences and collected independently from the coaching staff to maximize impartiality.

### Daily Analysis of Life Demands for Athletes (DALDA)

To validate convergent constructs and identify sources and symptoms of stress, the DALDA questionnaire (42), translated and validated into Portuguese (36), was used. DALDA consists of two parts: "Part A" - sources of stress and "Part B" - symptoms of stress, evaluating each item in three categories: worse than normal, normal, and better than normal. The frequency of "worse than normal" responses indicates elevated stress. The questionnaire was individually administered by a trained evaluator.

### Training indices: sRPE-TL (session Rating of Perceived Exertion – Training Load)

The sRPE-TL indices, monotony, and strain were calculated from the PSE (6 to 20 scale) and training session duration, as described by Foster (20). sRPE was recorded between 20 and 30 minutes after each training session. Athletes rated the overall intensity of the training session, including warm-up and cool-down, and data were recorded using Google Forms.

## **Statistical analysis**

Descriptive analysis (mean, standard deviation, and 95% confidence interval [95%CI]) was used for sample characterization. Data normality was checked using the Shapiro-Wilk test. For normally distributed variables, independent t-tests were used, while non-normally distributed variables were compared using the Mann-Whitney U test between non-starter and starter athletes. Statistical analyses were conducted using JAMOVI (Version 2.5.3) with a significance level of 5% (p < 0.05).

## **Validity evidence based on the internal structure of the instrument**

#### **Factor analysis**

#### Item validity was assessed through exploratory factor analysis using maximum likelihood method and Varimax rotation. Preconditions included Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity. Model fit was assessed by chi-square, Root Mean Squared Error of Approximation (RMSEA), Tucker-Lewis Index (TLI), and Bayesian Information Criterion (BIC). RMSEA values below 0.08 and TLI values above 0.90 (preferably above 0.95) were considered adequate (11,18).

***Internal consistency***

Internal consistency was evaluated using Cronbach's Alphas for all items and for each individual factor. Additionally, a Spearman correlation matrix between all Likert-type items was presented.

## **Validity evidence based on the relationship with other variables**

***Concurrent validity***

To assess concurrent validity, Spearman correlations were conducted between STSG items and DALDA, weekly sRPE-TL, monotony, and strain. Analyses were performed pre- and post-competition using data from the 7th and 14th days of STSG. Correlation values were interpreted as follows: r ≥ 0.10 = weak correlation; r ≥ 0.30 = moderate correlation; r ≥ 0.50 = strong correlation (13).

***Sensibility***

The sensitivity of STSG to temporal variations was evaluated using Generalized Estimating Equations (GEE) and Generalized Mixed Models (GMM). Models were created for each STSG item, considering fixed effects for the day, daily training load (pre-competition), non-starter/starter group (post-competition), and interaction between variables. Random effects included each athlete. Models were generated based on Akaike Information Criterion (AIC) and convergence, totaling 80 models.

# Results

# Out of 58 athletes, 38 responded to the questionnaires at least once during the study and were included in the analyses. Table 2 presents descriptive statistics of sample characteristics, including normality tests and distribution comparisons between starter (n=16) and non-starter(n=22) athletes in the study. Mean values, standard deviations, and 95% confidence intervals are provided for age, body mass, and stature variables.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2 -** Descriptive statistics of age, body mass and height with normality tests and distribution comparison tests between the reserve and starter groups for Brazilian American football athletes (n=38). | | | | | | | | | | | |
| **Variable** | **All sample (n=38)** | |  | **Non-starter (n= 22)** | | |  | **Starter (n= 16)** | | | p |
| **Mean (SD)** | **CI 95%** |  | **Mean (SD)** | **CI 95%** | **p-S.W.** |  | **Mean (SD)** | **CI 95%** | **p-S.W.** |
| **Age (years)** | 27.5(5.7) | 25.7-29.4 |  | 26.7(6.6) | 23.8-29.7 | 0.115 |  | 28.6(3.8) | 26.6-30.7 | 0.987 | 0.314 0 |
| **Body mass (kg)** | 101.1(23.9) | 109.0-97.5 |  | 96.4(24.9) | 85.3-107.4 | 0.009\* |  | 107.6(21.5) | 96.1-119.1 | 0.048\* | 0.139U0 |
| **Height (m)** | 1.82(0.6) | 1.80-1.85 |  | 1.80(0.06) | 1.77-1.82 | 0.846 |  | 1.86(0.06) | 1.82-1.89 | 0.023\* | 0.006U\* |
| **Note:** n = number of athletes; kg = kilograms; m = meters; SD = standard deviation; 95% CI = 95% Confidence interval; p-S.W. = p value for the Shapiro-Wilk normality test; p = p value for the distribution comparison test; U = the p value indicated in this cell is in relation to the Mann-Whitney U test; \* = p< 0.05. | | | | | | | | | | | |

# The average age of participants was 27.5 years (SD = 5.7), with a 95% CI between 25.7 and 29.4 years. No statistical differences were found regarding age between starters and non-starters. The overall mean body mass was 101.1 kg (SD = 23.9). Starters were taller (p = 0.006) than non-starters, 1.86 m (SD = 0.06) and 1.80 m (SD = 0.06), respectively. It is important to note that for body mass, both in the non-starter and starter groups, the Shapiro-Wilk normality test indicated a non-normal distribution (p < 0.05).

# Table 3 presents the distribution of practice time and athlete frequency per position for the entire sample, as well as for non-starter and starter athletes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 3 -** Distribution of Practice Time and Frequency of Athletes by Position for the Complete Sample, Reserves, and Starters of the Brazilian American Football Team (n=38). | | | |
| **Years of AF practice (n)** | Full sample  (n=38) | Non-starters (n=22) | Starters (n=16) |
| **I started this year** | 4 | 4 | 0 |
| **I’ve been playing for 1 year** | 3 | 2 | 1 |
| **It’s been 2 years** | 4 | 3 | 1 |
| **It’s been 3 years** | 2 | 1 | 1 |
| **It’s been 4 years** | 4 | 3 | 1 |
| **It’s been 5 years** | 5 | 3 | 2 |
| **It’s been 6 years** | 3 | 0 | 3 |
| **It’s been 7 years or more** | 13 | 6 | 7 |
| **Frequency of positions (n)** | Full sample  (n=38) | Non-starters (n=22) | Starters (n=16) |
| **Center** | 2 | 1 | 1 |
| **Offensive guard** | 2 | 1 | 1 |
| **Offensive tackle** | 1 | 0 | 1 |
| **Quarterback** | 3 | 2 | 1 |
| **Running back** | 3 | 2 | 1 |
| **Tight end** | 3 | 2 | 1 |
| **Wide receiver** | 4 | 2 | 2 |
| **Defensive tackle** | 4 | 2 | 2 |
| **Defensive end** | 3 | 2 | 1 |
| **Linebacker** | 5 | 3 | 2 |
| **Strong safety** | 3 | 2 | 1 |
| **Free safety** | 3 | 1 | 2 |
| **Cornerback** | 2 | 2 | 0 |
| **Note:** AF **=** american football;n = number of athletes. | | | |

# Most athletes have been practicing American football for 7 years or more. It was expected that starters would have the longest practice time, but this trend was not clear. Notably, all participating athletes who responded "started this year" were starters, emphasizing the complexity of learning and performance in American football. Regarding athlete positions, Linebackers were the most frequent, followed by Wide Receivers and Defensive Tackles. Most positions had a balanced distribution between non-starters and starters, with exceptions of Cornerbacks who were all non-starters (n=2), and the only Offensive Tackles (n=1), who were a starter.

# Table 4 and Figures 2 and 3 present the results of exploratory factor analysis of STSG, along with prerequisites and model fit measures for Brazilian football athletes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 4 -** Exploratory factor analysis, prerequisites and model adjustment measures for STSG in Brazilian American football athletes (n= 38). | | | | |
|  | **All sample (38 athletes - 400 responses)** | | | |
|  | **Factor 1** | **Factor 2** | **Singulariity** | **KMO** |
| **General** |  |  |  | 0.691 |
| **Quality of sleep** | 0.151 | **0.407** | 0.811 | 0.708 |
| **Tiredness sensation** | 0.261 | **0.637** | 0.527 | 0.722 |
| **Training willigness** | **0.646** | 0.461 | 0.370 | 0.724 |
| **Appetite** | **0.667** | -0.09 | 0.547 | 0.664 |
| **Competitive willigness** | **0.659** | 0.196 | 0.527 | 0.724 |
| **Muscle soreness** | -0.119 | **0.503** | 0.732 | 0.526 |
| ***Eigen value*** | 1.4 | 1.1 |  |  |
| **% of variance** | 23.4 | 18.1 |  |  |
| **Cumulative** | 23.4 | 41.4 |  |  |
| **Bartlett's test** | **χ²** | **df** | **p** | |
| 439 | 15 | <0.001 | |
| **Model fit measures** | | | | |
|  | **χ²** | **df** | **p** | |
| **Model testing** | 12.4 | 4 | 0.02 | |
| **RMSEA** | 0.0724 (90%CI: 0.029-0.120) | | | |
| **TLI** | 0.93 | | | |
| **BIC** | -11.6 | | | |
| **Note:** STSG = sample time series graphs; RMSE = root mean squared error aproximation; TLI = Tucker-lewis index; BIC = Bayesian information criterion; 90%CI = 90% confidence interval; KMO = measure of adequacy of sampling Kaiser-Meyer-Olkin; χ² = chi-square; df = degrees of freedom; p = p value. | | | | |

Bartlett's test of sphericity (χ² = 439, df = 15, p < 0.001) and the overall KMO (0.691) suggested that the correlation matrix of items is factorable. Parallel analysis suggested two factors as most representative for the data. Except for χ² of model fit measures, the factorial structure showed adequate fit indices (χ² = 12.4, df = 4, p = 0.02; RMSEA = 0.0724; BIC = -11.6; TLI = 0.93).

Figure 2, presenting the scree plot, supports the choice of two factors, showing a clear inflection after the second factor.

**Figure 2 -** Distribution of eigenvalues ​​(eigen values-scree plot) of the likert-type items of the training monitoring graphs, for the data and for the simulation.

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

# Figure 3 depicts the trajectory diagram of Likert-type item factors from STSG.

**Figure 3 -** Factor trajectory diagram of STSG likert-type items.

Diagrama

Descrição gerada automaticamente

**Note:** TW = training willingness; AP = appetite; CW = competitive willingness; QS = quality of sleep; TS = tiredness sensation; MS = muscle soreness; Ft1 = factor 1; Ft2 = factor 2.

The figure provides a visual representation of the relationships between the items and the identified factors, consistent with the interpretation above. Model fit measures indicate reasonable fit, with RMSEA = 0.0724 (90% CI: 0.029-0.120) and TLI = 0.93.

Table 5 presents the results of internal consistency analysis for all responses.

# 

# The Cronbach's Alpha was acceptable for Likert scale items of the instrument but varied from good for Factor 1 to suggesting revision for Factor 2. Excluding the muscle pain item could improve the Cronbach's Alpha. Three items showed item-total correlations below 0.3, indicating they could be excluded. Only two Spearman correlations were significant: competitive readiness and readiness for training; and readiness for training and fatigue sensation, both with R > 0.5.

Table 6 found significant correlations on different days of the study. Regarding the 7th day, moderate positive correlations were found between sleep hours, weekly sRPE-TL, and moderate negative correlations in the number of "worse than normal" responses for stress symptoms, and number of responses "to all responses of the DALDA".

On the 14th day, resting heart rate, sleep quality, fatigue sensation, readiness for training, appetite, and sum of scores showed significant correlation with the number of "better than normal" responses for at least one part of the DALDA. For training load indicators, only one significant correlation was found between body mass and monotony. The only significant negative correlation found was between resting heart rate and the number of "stress sources" responses in DALDA.



Table 7 demonstrates the mean, standard deviation, range, and 95% confidence interval for STSG variables over study days, with significant main effects for day, daily load (for pre-competition analysis), non-starter/starter group (for post-competition analysis) indicated by Omnibus test and significant differences indicated by Bonferroni post hoc test.

.

# 

Bonferroni-adjusted significance for:

d= day 8 non-starter group > day 9 non-starter group; p=0.021;

e= day 8 < day 10; p=0.005;

f= day 8 < day 11; p=0.032;

g= day 8 < day 13; p=0.016;

h= day 8 < day 14; p=0.016;

i= day 10 non-starter group > day 12 non-starter group; p=0.006;

j= day 10 non-starter group > day 14 non-starter group; p=0.018;

k= day 8 < day 10; p=0.012;

l= day 8 < day 11; p=<0.001;

m= day 8 < day 13; p=<0.001;

n= day 8 < day 14; p=0.015;

o= day 10 non-starter group > day 8 starter group; p=0.005;

p= day 11 non-starter group > day 8 starter group; p=<0.001;

q= day 13 non-starter group > day 8 starter group; p=0.002;

r= day 14 non-starter group > day 8 starter group; p=0.034;

s= day 8 non-starter group > day 8 starter group; p=0.046;

t= day 8 starter group < day 10 starter group; p=0.008;

u= day 8 starter group < day 11 starter group; p=0.002;

v= day 8 starter group < day 12 starter group; p=0.001;

w= day 8 starter group < day 13 starter group; p=<0.001;

x= day 8 starter group < day 14 starter group; p=<0.001;

y= day 9 non-starter group > day 8 starter group; p=0.005;

z= day 8 < day 13; p=0.031;

§= day 9 < day 13; p=0.004;

†= day 8 starter group < day 13 starter group; p=0.002.

The analysis revealed significant main effects for sRPE-TL, training day, and group (non-starter or starters) on various STSG items. Significant interactions were observed between day and group for some items. In pre-competition analyses, a main effect was found only for sRPE-TL in muscle pain. For post-competition data, a main effect was found for day in readiness for training, muscle pain, and sum of scores; a main effect for non-starter/starter group in muscle pain; and an interaction effect for sleep quality, competitive readiness, muscle pain, and sum of scores.

In post-competition analysis, the Bonferroni test indicated that readiness for training was lower on day 8 compared to days 10, 11, 13, and 14. Muscle pain was worse on day 8 compared to days 10, 11, 13, and 14. For the main effect of non-starter/starter groups post-competition, the post-hoc test indicated a difference only for muscle pain, suggesting that the starters group exhibits higher post-competition pain levels.

In the interaction between day and non-starter/starter groups post-competition, the sleep quality score, competitive readiness, muscle pain, and sum of scores on day 8 for the non-starter group were higher than day 9 for the same group. Comparing days in the starters group, day 8 had lower scores than days 10, 11, 12, 13, and 14. Regarding the sum of scores, only the score on day 8 for the starters group was lower than day 13 for the same group

# Discussion

# The focus of this study was to explore evidence of validity for the STSG, based on its internal structure and its relationship with other variables. Regarding internal structure evidence, the Likert-type items of the STSG demonstrated factorable structure with good fit indices. Two factors with moderate fit indices were identified, and internal consistency was confirmed with an acceptable Cronbach's alpha. Additionally, there were several positive associations among item scores, reinforcing the reliability of the instrument. Concerning evidence based on the relationship with other variables, concurrent validity analysis revealed numerous associations, all indicating directions that support the use of item scores, especially for post-competition periods. Sensitivity analysis indicated stability during the pre-competition period. However, during the post-competition period, several items and the sum of items were sensitive to competition, indicating higher levels of competitive stress in the days following competition, with a progressive decrease during the recovery period.

# The initial hypotheses that there is evidence that the STSG is a valid instrument for monitoring training stress in Brazilian American football athletes were largely confirmed. Due to the complex nature of the instrument, it is essential to analyze variables such as resting heart rate and body mass considering their variations over a longer period. It may be that the 7-day pre-competition and 7-day post-competition period was insufficient to provoke significant changes in these variables, as indicated by temporal analysis. Therefore, studies with longer monitoring periods using the STSG are suggested. Accumulated evidence suggests that the STSG can be used to monitor training stress in Brazilian American football athletes.

To our knowledge, only the study by Baghaei et al.(4) has used the STSG to monitor a competitive season. In this study, the authors longitudinally evaluated 12 female soccer players over a 6-month season. In addition to the STSG, cortisol data and the Société Française de Médecine du Sport Questionnaire (SFMS-Q) were collected. The SFMS-Q is a 54-item questionnaire with binary response options (“yes” or “no”), where higher score indicates more symptoms and advancement toward the overtraining continuum. All measures were collected pre-competition. Except for body mass, all STSG items were also collected over 7 consecutive days in each period, similar to our study, but they used the 7-day average for analysis. Key results for the STSG indicated higher stress levels during season compared to the beginning of the season for all items except for sleep hours. At the end of season, only resting heart rate and training willingness showed significant values indicating higher stress levels. The SFMS-Q showed a similar pattern to resting heart rate and training willingness in all comparisons, suggesting validity evidence for the STSG as a tool to monitor training stress in female soccer athletes. Additionally, cortisol levels were higher post-competition compared to pre-competition, reinforcing STSG validity evidence as it also signaled higher stress levels coinciding with the indicated moments in both subjective (SFMS-Q) and physiological (cortisol) measures. Although this study indicated evidence that the STSG could be a valid instrument for monitoring training stress, the authors did not verify the factorial structure, internal consistency, or the correlation between STSG items and other stress indicators, as done in our study.

While there are no other studies using the STSG in American football athletes, similar subjective instruments have been utilized in other studies (22,24,27,33,34,43,50,51). Interestingly, for most of these instruments, such as well-being perception, stress perception, self-reported thermal stress, and subjective evaluation of the training program, psychometric properties, and the rigorous steps for creating psychometric instruments as suggested by specialized literature are not reported or carried out (46,54). This may result from a lack of familiarity with psychometric terminology and concepts among sports and exercise scientists. However, although they often do not recognize it, these studies frequently verify validity evidence based on relationship with other variables (e.g., training load, oxidative stress, cortisol training period). For example, Sampson et al. (43), with 42 NCAA division 1 athletes, used a “wellness questionnaire” composed of Likert-type items scored from 1 to 5: pain, sleep, energy, and average scores. In this instrument, higher scores indicate higher wellness. The study lasted 17 weeks, and the questionnaire was completed daily on field training day. Concurrently, injury incidence and the acute-chronic workload ratio based on the Playerload variable from GPS (sum of 7-day Playerload divided by the sum of 21 previous days Playerload) were collected. Regarding injury incidence, the authors found that the wellness z-score (-1 = worse wellness, 0= normal wellness, and +1 better wellness) alone could not discriminate against different injury probabilities. However, when considering the acute-chronic workload ratio z-score, athletes had higher injury probabilities with z-scores of -2 and +2. When checking the interaction of wellness questionnaire items with the acute-chronic workload ratio only pain items and average scores showed a main effect. Both differences occurred only for the +2 z-score of acute-chronic workload ratio, indicating that athletes with better pain and average scores had a higher probability of injury. Paradoxically, this suggests that athletes who accumulate training load more quickly and feel well may be at a greater risk of injury. This may be associated with inadequate recovery and underestimation of subjective signs of fatigue, reinforcing the use of combined instruments to monitor training.

Using the same psychometric instrument, Fullagar et al. (22) assessed 52 NCAA Division 1 athletes, one day before each game, and 2, 3, and 4 days after all games, during two consecutive regular seasons (2x12 weeks). For the average wellness score, a small effect was found, with the likelihood of being lower on days 2, 3, and 4 compared to the day before the competition. For sleep, a small effect was also found, with a lower likelihood for the 3rd day after the competition compared to the day before the competition. For energy, a small effect was found, with a lower score likelihood for days 3 and 4 compared to the day before the competition. This study reinforces evidence that psychometric instruments are sensitive to post-competition stress.

Our study aligns with the methodology used in other studies that seek to verify validity evidence of psychometric instruments in the sports context. Several studies focus on the validation of instruments that assess cognitive constructs, such as Behavioral Regulation in Sport, Basic Needs Satisfaction in Sport, and Sport Multidimensional Perfectionism (17,30,38). For example, in Lonsdale, Hodge, and Rose's (30) study, the ideal steps for creating psychometric instruments, as suggested by specialized literature (1,46,54), were followed. The authors constructed the Behavioral Regulation in Sport Questionnaire (BRSQ), a tool to assess intrinsic motivation, extrinsic motivation, and amotivation in athletes. Initially, the authors conducted a detailed review of constructs and interviews with 15 athletes, resulting in 80 preliminary items. After review by seven experts, the first version of the BRSQ contained 42 items (42-BRSQ). This questionnaire was applied in a pilot study with 10 athletes and underwent confirmatory factor analysis with 382 athletes, resulting in the final version with 32 questions (32-BRSQ). Internal consistency was verified by Cronbach's alpha. Additionally, the authors conducted concurrent validity analysis with other related questionnaires and a test-retest reliability analysis. Similarly, in Ng, Lonsdale, and Hodge's (38) study, the Basic Needs Satisfaction in Sport Scale (BNSSS) was initially developed and validated to assess autonomy, competence, and relatedness satisfaction in competitive sports contexts. After interviewing 6 athletes and selecting 32 items, the agreement analysis reduced the instrument to 15 items (15-BNSSS). This questionnaire was applied to 273 athletes and subsequently, the autonomy subscale was reviewed, resulting in a new application to 371 athletes. Factorial and internal consistency analysis was performed as well as a concurrent validity analysis with other scales and a test-retest with 63 athletes. In Dunn et al.'s (17) study, the focus was on validating the Sport Multidimensional Perfectionism Scale (Sport-MPS). The instrument was applied in 4 distinct samples, totaling 847 athletes from different competitive levels. Evidence of validity based on internal structure and relationship with other variables was verified, despite some necessary adjustments in the confirmatory factor analysis. Like in our study, the mentioned works used robust techniques to verify different sources of evidence and validity of their psychometric instruments.

Among the techniques used to verify evidence of validity based on the internal structure of an instrument, factor analysis and internal consistency are widely used and were used in our study. In factor analysis, except for the χ² of model fit measures, the model fit indices showed good adequacy, suggesting that the proposed model is robust and adequate. The χ² of model fit measures significantly indicates that the model estimates did not sufficiently replicate the sample variances and covariances. However, the χ² of model the significance of this test is highly influenced by the large sample size and should be used in conjunction with other fit indices (11). The structure revealed two well-defined factors: Factor 1 included training readiness, appetite, and competitive disposition items. Factor 2 included sleep quality, tiredness sensation, and muscle pain items. This indicates that the Likert-type items of the STSG measure distinct constructs. Theoretically, Factor 1 items are related to the physical readiness and motivational construct for training, while Factor 2 items are related to the fatigue and recovery construct of athletes.

The analysis of internal consistency, measured by Cronbach's alpha, indicates that Likert-type items are consistent and represent the same construct. Excluding the muscle soreness item increases Cronbach's alpha, likely due to its high sensitivity to post-game changes, as shown in the sensitivity analysis (Table 7). This response to post-competition soreness aligns with the literature, as observed in Choi et al.'s study (12), which found a lower pain threshold after kendo competition. Therefore, removing or adapting the scoring scale for the muscle soreness item could enhance the internal consistency of the instrument.

When analyzing the internal consistency of each factor separately, only Factor 2 showed a moderate/low Cronbach's alpha value (0.482). This may be attributed to the large variation in muscle soreness scores, as noted in the sensitivity analysis (Table 7), which did not occur with the other items in this factor (sleep quality and fatigue sensation). The lack of significant associations between sleep quality and muscle soreness, appetite, and between competitive disposition and muscle soreness reinforces the internal structure consistency of the instrument.

Concurrent construct validation demonstrated significant correlations between STSG items and indicators: number of "a" responses (worse than normal); number of "c" responses (better than normal) from DALDA, weekly training load, and monotony. Significant associations differed between pre-competition (day 7) and post-competition (day 14) periods. Before competition, associations were found between hours of sleep and number of "a" responses for stress symptoms (r=-0.43), number of "a" responses for the complete DALDA (r=-0.41), and weekly training load (r=0.39). This indicates that more sleep is associated with lower stress levels and higher training loads accumulation, or vice versa. However, Fullagar et al.'s study (22) with 23 American football athletes found that objective sleep indicators were worse during intensified training periods with higher external training loads, measured by accelerometers (Playerload). It is important to note that this population consisted of young college students, and academic stress may have been a confounding factor in their conclusions.

In the post-competition period (day 14), only competitive disposition and muscle soreness items showed no significant associations with other training stress indicators. This contradicts our hypothesis about the concurrent validity of these items but underscores the need to evaluate them separately when monitoring training, as they measure different constructs. Regarding resting heart rate, other Likert-type items, and the sum of STSG scores, moderate significant correlations were found with number of "c" responses (better than normal) for at least one part of DALDA (stress sources, stress symptoms, and/or complete DALDA). This indicates that, in the post-competition, STSG was more associated with low stress indicators than high stress indicators.

The sensitivity analysis of STSG revealed significant variations in items over time, especially in the post-competition period. Before competition, muscle soreness was the only item that showed a main effect for sRPE-TL (training load), suggesting that pre-competition training volumes can directly impact athletes' perception of muscle soreness. The absence of association with other STSG items was expected at this time, as athletes were not under significant training or competition stress, as observed in the post-competition period. However, muscle soreness can assist in identifying training load accumulation.

After the competition, we found that readiness for training, muscle soreness, and the sum of scores showed a significant day effect, varying significantly over days. All analyses indicate that on the days following competition, training stress was significantly higher than in subsequent days. This shows that the instrument was able to detect the immediate impact of competition on athletes' physical and psychological states, aligning with literature suggesting a critical recovery period in the first day’s post-competition (22).

Regarding the main effect for non-starters and starters group, starters experienced higher levels of muscle soreness post-competition. This may be attributed to greater exposure and intensity of play for starters, emphasizing the importance of differentiated recovery and training strategies between these groups. Additionally, main effects were found for the interaction of day and non-starters/starters group for sleep quality, competitive disposition, muscle soreness, and sum of scores items. For sleep quality, a difference was found only between days 8 and 9 for the non-starters group, indicating that only these athletes slept better after competition. This likely occurred due to the lower physical and emotional burden experienced by non-starter athletes, allowing for better sleep quality on the night following competition. If the match result were positive, it is possible that starting athletes would also experience better sleep quality on the subsequent day. For competitive disposition, only the non-starter group showed higher levels of disposition on day 10 compared to days 12 and 14. Towards the end of the season, it is common for some athletes to change teams, need time to recover from injuries, or decide to quit/retire from the sport. Additionally, it is a cultural practice among coaches to offer a similar number of snaps to players (non-starters and starters) during the preparation period, to decide on the new starters for the next competition. These factors may have influenced the increase in competitive disposition only for the non-starter group.

The interaction main effect for muscle soreness showed that soreness was always higher for starters in all significant interactions between the two groups. For interactions that occurred between days within the starters group, all indicated higher levels of muscle soreness for the day following competition. No interaction was found between days within the non-starters group. These findings corroborate with previous main effects, suggesting that personalized interventions, considering athlete status (non-starter or starter) and specific time in the training and competition cycle, are important for stress management and athletic performance optimization. For the interaction main effects of the sum of Likert-type items scores from STSG, there was a lower score only for day 8 compared to day 13 within the starters group, indicating that the instrument detected higher stress on the day following competition. This aligns with our hypothesis that STSG can be used to detect stress, demonstrating sensitivity to variations in stress levels in interactions where higher stress levels were expected. These findings reinforce the validity of evidence of STSG as a tool for monitoring Brazilian American football athletes.

In the discussion of the secondary findings of this study, the differentiation between starters and non-starters regarding stature deserves attention. Previous studies support our findings, indicating that this physical characteristic can provide significant competitive advantages for American football players. For instance, a study by Black and Roundy (6), involving 1618 American football players, found that starters tended to be heavier, stronger, and faster than non-starters. Additionally, Hoffman et al. (26) conducted a longitudinal study with 221 NCAA Division III athletes, observing that American football players became significantly heavier by the 5th year compared to the 1st and 2nd years, while maintaining a constant body fat percentage. When stratifying into lineman (defensive lineman, centers, guards, tackles, defensive tackles, defensive ends, linebackers, and tight ends) and backs (running backs, quarterbacks, wide receivers, and defensive backs), it was noted that backs showed a significant increase in body mass from the second year onwards, continuing through the fifth year. In contrast, linemen showed a significant increase only in the fourth year compared to the first. This pattern suggests that weight gain is an adaptation required during American football seasons, with potential implications for performance and game strategy.

The main limitations of this study include the concentration of data from a single group of Brazilian American football athletes. These limitations were mitigated by the robustness of the statistical methods used and by the repeatability of measurements over time. The use of a longitudinal design also strengthened the conclusions, allowing for a more detailed analysis of variations over time. Furthermore, the primary goal of STSG is to be used over extended consecutive periods, enabling coaches and athletes to gather a substantial amount of information and observe fluctuations in indicators throughout the entire season. Nevertheless, the experimental designs of only seven consecutive days limit the applicability of all the advantages that STSG could offer.

# 

# Conclusion

We confirm the validity evidence of the STSG use. The results of this study provide robust support for the use of STSG as an effective tool in assessing training stress in American football athletes, contributing to the optimization of athletic performance and athletes' health. However, we recommend that future research explore longer periods with STSG, particularly considering variables such as resting heart rate and body mass, to better understand their variations over time and their relationship to training stress. Additionally, new studies could utilize other outcome variables, such as injuries, biomarkers, and performance decline and be conducted with populations from other sports. These new studies could further elucidate the clinical and practical utility of STSG in managing training stress. Additionally, new studies could utilize other outcome variables such as injuries and performance declines in other sports.

**Practical applications**

Considering the practical applicability of our results, it highlighted the utility of STSG in monitoring training stress in Brazilian American football athletes, especially post-competition. The instrument's sensitivity to post-competition stress variations, as evidenced by sensitivity analysis, reinforces its ability to capture the immediate impact of games on athletes' physical and psychological aspects. This is crucial for adjusting recovery and training strategies, customizing them according to individual needs and roles on the team (non-starter or starter). Finally, it is recommended that future research explore longer monitoring periods with STSG, especially considering variables such as resting heart rate and body mass, to better understand their variations over time and their relationship with training stress. These studies can further elucidate the clinical and practical utility of STSG in managing training stress in American football athletes and other sports.

# Declaration of conflicting interests

Nothing to declare.

# Contributions

Contributed to conception and design: MFTJr, DRLM, LSLS, PPA, and APS.

Contributed to acquisition of data: MFTJr, LSLS, LSG.

Contributed to analysis and interpretation of data: MFTJr.

Drafted and/or revised the article: MFTJr, LSLS, LSG, APS, PPA, and DRLM.

Approved the submitted version for publication: MFTJr, LSLS, LSG, APS, PPA, and DRLM.

# Acknowledgements

The authors would like to thank all the athletes, coaches and evaluators who participated in this study.

# Funding information

This study was financed by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-CAPES-Finance Code 001 (Bsc. Márcio Fernando Tasinafo Júnior 88887.717878/2022-00).

# REFERENCES

# 1. AERA, APA, and NCME. Standards for Educational and Psychological Testing. 6 th ed. Washington: American Educational Research Association; American Psychological Association and National Council on Measurement in Education, 2014.

# 2. Alves, RN, Costa, LOP, and Samulski, DM. Monitoring and prevention of overtraining in athletes. Rev Bras Med Esporte 2006; 12: 291–296. <https://doi.org/10.1590/S1517-86922006000500013>

# 3. Baechle, TR and Earle, RW. Essentials of Strength Training and Conditioning. 3 rd ed. Champaign: Human Kinetics, 2008.

# 4. Baghaei, S, Tadibi, V, Amiri, E, and Giboin, L-S. Subjective and objective variables of overtraining syndrome in female soccer players: A longitudinal study. Sci & Sports 2002; 37: 459–467. [10.1016/j.scispo.2021.05.005](https://doi.org/10.1016/j.scispo.2021.05.005)

# 5. Bellenger, CR, Fuller, JT, Thomson, RL, et al. Monitoring Athletic Training Status Through Autonomic Heart Rate Regulation: A Systematic Review and Meta-Analysis. Sports Med 2016; 46: 1461–1486. [10.1007/s40279-016-0484-2](https://doi.org/10.1007/s40279-016-0484-2)

# 6. Black, W and Roundy, E. Comparisons of Size, Strength, Speed, and Power in NCAA Division 1-A Football Players. J of Strength & Cond Res 1994; 8: 80-85. <10.1519/JSC.0000000000001486>

# 7. Bompa, T and Haff, GG. Periodização. Teoria e Metodologia do Treinamento. 5 th ed. São Paulo: Phorte, 2013.

# 8. Bompa, TO and Haff, G. Periodization: Theory and Methodology of Training. 5 th ed. Champaign: Human Kinetics, 2009.

# 9. Bourdon, PC, Cardinale, M, Murray, A, et al. Monitoring Athlete Training Loads: Consensus Statement. , Int J Sports Physiol Perform 2017; 12(s2): 161-170. [10.1123/IJSPP.2017-0208](https://journals.humankinetics.com/view/journals/ijspp/12/s2/article-pS2-161.xml).

# 10. Brandenburg, MA, Butterwick, DJ, Hiemstra, LA, Nebergall, R, and Laird, J. A comparison of injury rates in organised sports, with special emphasis on American bull riding : review article. International SportMed Journal 2007: 8: 78–86.

# 11. Brown, TA. Confirmatory Factor Analysis for Applied Research. 2 nd ed. New York: Guilford Publications, 2015.

# 12. Choi, JC, Min, S, Kim, YK, et al. Changes in pain perception and hormones pre- and post-kumdo competition. Horm Behav 2013: 64: 618–623. [10.1016/j.yhbeh.2013.08.013](https://doi.org/10.1016/j.yhbeh.2013.08.013).

# 13. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York: Routledge, 1988.

# 14. Dantas, EHM. A Prática da Preparação Física. 6 th ed. São Paulo: Roca, 2014.

# 15. Dick, R, Agel, J, and Marshall, SW. National Collegiate Athletic Association Injury Surveillance System Commentaries: Introduction and Methods. J Athl Train 2007; 42: 173–182, 2007.

# 16. Dick, R, Ferrara, MS, Agel, J, et al. Descriptive Epidemiology of Collegiate Men’s Football Injuries: National Collegiate Athletic Association Injury Surveillance System, 1988–1989 Through 2003–2004. J Athl Train 2007; 42: 221–233. [PMC1941300](https://pmc.ncbi.nlm.nih.gov/articles/PMC1941300/)

# 17. Dunn, JGH, Dunn, JC, Gotwals, JK, et al. Establishing construct validity evidence for the Sport Multidimensional Perfectionism Scale. Psychol of Sport and Exer 2006; 7: 57–79. [10.1016/j.psychsport.2005.04.003](https://doi.org/10.1016/j.psychsport.2005.04.003).

# 18. Fabrigar, L, Wegener, D, MacCallum, R, and Strahan, E. Evaluating the use of Exploratory Factor Analysis in psychological research. Psychological Methods 1999; 4: 272–299. [10.1037/1082-989X.4.3.272](https://doi.org/10.1037/1082-989X.4.3.272).

# 19. Feeley, BT, Kennelly, S, Barnes, RP, et al. Epidemiology of National Football League Training Camp Injuries from 1998 to 2007. Am J Sports Med 2008; 36: 1597–1603. [10.1177/036354650831602](https://doi.org/10.1177/0363546508316021).

# 20. Foster, C. Monitoring training in athletes with reference to overtraining syndrome. Med Sci Sports Exerc 1998; 30: 1164–1168. [10.1097/00005768-199807000-00023](https://doi.org/10.1097/00005768-199807000-00023).

# 21. Fry, RW, Morton, AR, and Keast, D. Overtraining in Athletes. Sports Med 1991; 12: 32–65. [10.2165/00007256-199112010-00004](https://doi.org/10.2165/00007256-199112010-00004).

# 22. Fullagar, HHK, Govus, A, Hanisch, J, and Murray, A. The Time Course of Perceptual Recovery Markers After Match Play in Division I-A College American Football. Sci and Med in Football, 2017; 3: 332-336. [10.1123/ijspp.2016-0550](https://doi.org/10.1123/ijspp.2016-0550).

# 23. Gabbett, TJ. The development and application of an injury prediction model for noncontact, soft-tissue injuries in elite collision sport athletes. J Strength Cond Res 2010; 24: 2593–2603. <10.1519/JSC.0b013e3181f19da4>.

# 24. Hamlin, MJ, Wilkes, D, Elliot, CA, Lizamore, CA, and Kathiravel, Y. Monitoring Training Loads and Perceived Stress in Young Elite University Athletes. Front Physiol 2019; 10(34): 1-12. [10.3389/fphys.2019.00034](https://doi.org/10.3389/fphys.2019.00034).

# 25. Heyward, VH and Gibson, AL. Advanced Fitness Assessment and Exercise Prescription. 9 th ed. Champaign: Human Kinetics, 2024.

# 26. Hoffman, JR, Ratamess, NA, and Kang, J. Performance Changes During a College Playing Career in NCAA Division III Football Athletes. J of Strength & Cond Res 2011; 25: 2351. <10.1519/JSC.0b013e31821743df>.

# 27. Hoffman, JR, Ratamess, NA, Klatt, M, et al. Comparison Between Different Off-Season Resistance Training Programs in Division III American College Football Players. J of Strength & Cond Res 2009; 25: 2351-2357. <10.1519/JSC.0b013e3181876a78>.

# 28. Johnson, MB and Thiese, SM. A Review of Overtraining Syndrome—Recognizing the Signs and Symptoms. J Athl Train 1992; 27: 352–354. [PMID:16558192](https://pubmed.ncbi.nlm.nih.gov/16558192/).

# 29. Lohman, TG, Roche, AF, and Martorell, R. Anthropometric standardization reference manual. 1 st ed. Chicago: Human Kinetics, 1988.

# 30. Lonsdale, C, Hodge, K, and Rose, EA. The Behavioral Regulation in Sport Questionnaire (BRSQ): Instrument Development and Initial Validity Evidence. J of Sport & Exer Psychol 2008; 30: 323-355. [10.1123/jsep.30.3.323](https://doi.org/10.1123/jsep.30.3.323).

# 31. Matos, NF, Winsley, RJ, and Williams, CA. Prevalence of nonfunctional overreaching/overtraining in young English athletes. Med Sci Sports Exerc 2011; 43: 1287–1294. <10.1249/MSS.0b013e318207f87b>.

# 32. Matveev, LP. Treino desportivo: metodologia e planejamento. 1st ed. Londrina: FMU FIAM FAAM, 1997.

# 33. McKay, BA, Delaney, JA, Simpkin, A, et al. The Association Between Alterations in Redox Homeostasis, Cortisol, and Commonly Used Objective and Subjective Markers of Fatigue in American Collegiate Football. Int J Sports Physiol Perform 2021; 16: 1851–1857. [10.1123/ijspp.2020-0933](https://doi.org/10.1123/ijspp.2020-0933).

# 34. McKay, BA, Delaney, JA, Simpkin, A, et al. Objective Measures of Strain and Subjective Muscle Soreness Differ Between Positional Groups and Season Phases in American College Football. Int J Sports Physiol Perform 2023; 18: 625–633. [10.1123/ijspp.2022-0347](https://doi.org/10.1123/ijspp.2022-0347).

# 35. Meeusen, R, Duclos, M, Foster, C, et al. Prevention, diagnosis, and treatment of the overtraining syndrome: joint consensus statement of the European College of Sport Science and the American College of Sports Medicine. Med Sci Sports Exerc 2013; 45: 186–205. [10.1249/mss.0b013e318279a10a](https://doi.org/10.1249/mss.0b013e318279a10a).

# 36. Moreira, A and Cavazzoni, PB. Monitorando o treinamento através do Wisconsin upper respiratory symptom survey -21 e daily analysis of life demands in athletes nas versões em língua Portuguesa. R Educ Fís UEM 2009; 20: 109–119. [10.4025/reveducfis.v20i1.5289](https://doi.org/10.4025/reveducfis.v20i1.5289).

# 37. Morrow Jr, JRM, Jackson, AW, Disch, JG, and Mood, DP. Medida e Avaliação do Desempenho Humano. 4 th ed. São Paulo: Artmed Editora, 2014.

# 38. Ng, JYY, Lonsdale, C, and Hodge, K. The Basic Needs Satisfaction in Sport Scale (BNSSS): Instrument development and initial validity evidence. Psychology of Sport and Exercise 2011: 12: 257–264. [10.1016/j.psychsport.2010.10.006](https://doi.org/10.1016/j.psychsport.2010.10.006).

# 39. Pasquali, L. Psicometria: Teoria dos testes na psicologia e na educação. 1 st ed. Petrópolis: Editora Vozes Limitada, 2017.

# 40. Platonov, VN. Tratado Geral de Treinamento Desportivo. 1 st ed. São Paulo: Phorte, 2008.

# 41. Robbins, DW and Young, WB. Positional relationships between various sprint and jump abilities in elite American football players. J Strength Cond Res 2012: 26: 388–397. <10.1519/JSC.0b013e318225b5fa>.

# 42. Rushall, BS. A tool for measuring stress tolerance in elite athletes. J of App Sport Psychol 1990: 2: 51–66. [10.1080/10413209008406420](https://doi.org/10.1080/10413209008406420).

# 43. Sampson, JA, Murray, A, Williams, S, Sullivan, A, and Fullagar, HHK. Subjective Wellness, Acute: Chronic Workloads, and Injury Risk in College Football. J of Strength & Cond Res 2019; 33: 3367-3373. <10.1519/JSC.0000000000003000>.

# 44. Savioli, FP, Medeiros, TM, Camara Jr, SL, Biruel, EP, and Andreoli, CV. Diagnosis of overtraining syndrome. Rev Bras Med Esporte 2018; 24: 391–394. [10.1590/1517-869220182405185927](https://doi.org/10.1590/1517-869220182405185927).

# 45. Selye, H. Stress and disease. Science 1995; 122: 625–631.

# 46. Sireci, S and Padilla, J-L. Validating assessments: Introduction to the Special Section. Psicothema 2014; 26: 97–99. <10.7334/psicothema2013.255>.

# 47. Urbina, S, Dornelles, C, and Nunes, MLT. Fundamentos da Testagem Psicológica. 1 st ed. São Paulo: Artmed, 2007.

# 48. Urhausen, A and Kindermann, W. Diagnosis of overtraining: what tools do we have? Sports Med 2002; 32: 95–102. [10.2165/00007256-200232020-00002](https://doi.org/10.2165/00007256-200232020-00002).

# 49. Viru, A. Early contributions of Russian stress and exercise physiologists. J of Applied Physiology 2002; 92: 1378–1382. [10.1152/japplphysiol.00435.2001](https://doi.org/10.1152/japplphysiol.00435.2001).

# 50. Wardenaar, FC, Vento, KA, Ortega-Santos, CP, Connolly, J, and Vanos, JK. The impact of different playing surfaces on physiological parameters in collegiate DI American football athletes. Int J of Sports Sci & Coaching 2023; 18: 781–792. [10.1177/1747954122108974](https://doi.org/10.1177/17479541221089748).

# 51. Wellman, AD, Coad, SC, Flynn, PJ, Siam, TK, and McLellan, CP. Perceived Wellness Associated With Practice and Competition in National Collegiate Athletic Association Division I Football Players. J Strength Cond Res 2019; 33: 112–124. <10.1519/JSC.0000000000002169>.

# 52. Wellman, AD, Coad, SC, Goulet, GC, and McLellan, CP. Quantification of Competitive Game Demands of NCAA Division I College Football Players Using Global Positioning Systems. J Strength Cond Res 2016; 30: 11–19, 2016. <0.1519/JSC.0000000000001206>.

# 53. Wellman, AD, Coad, SC, Goulet, GC, and McLellan, CP. Quantification of Accelerometer Derived Impacts Associated With Competitive Games in National Collegiate Athletic Association Division I College Football Players. J Strength Cond Res 2017; 31: 330–338. <10.1519/JSC.0000000000001506>.

54. Zumbo, BD and Chan, EKH (eds.). Validity and Validation in Social, Behavioral, and Health Sciences. 1st ed. New York: Springer, 2016.