Psychometric properties of the Three-Dimensional Anxiety Scale for Sport (ETApE) through CFA and IRT approaches

Karen Cristine Teixeira[1], António Fernando B. Rosado[2], Carlos Henrique S. S. Nunes[3]

Abstract

The present study aimed to improve and examine the reliability, evidence regarding internal structure and invariance of the ETApE, an anxiety trait scale for sport. The study sample was formed by 500 athletes from the south of Brazil, being 64% males, from different sports and competitive levels, ranging from international to local. To examine the psychometric properties of the measurement model, a confirmatory factor analysis (CFA) with maximum likelihood method was conducted and the composite reliability was assessed. Item parameters were analysed in light of the Item Response Theory (IRT). Moreover, the model's invariance was tested and a latent means comparison according to gender was performed. The instrument presented good psychometric qualities that point to its validity regarding internal structure. Good indicators of reliability and factorial invariance were also achieved. A statistically significant difference was found in the three dimensions of the instrument according to gender. Women scored higher in cognitive and physiological anxiety and had lower score in the perceived control dimension.

Keywords: Anxiety; Sport; Validity; Reliability; Invariance

One of the most studied emotions in sport context is anxiety, since it is the one that most frequently mobilizes athletes, regardless of their competitive level (Fernandes et al., 2014; Marcel and Paquet, 2010). The construct can be defined as an unpleasant psychological state in response to perceived threats concerning the sport performance under pressure (Cheng, Hardy and Markland, 2009). Can also be considered as a psychological trait that represents a stable aspect of personality, a predisposition to present anxiety behaviours front of situations interpreted as threats (Weinberg and Gould, 2018).

The athlete is constantly exposed to social and environmental stressors, high training loads and internal and external pressures to achieve better results (Hamlin, Wilkes, Elliot, Lizamore, and Kathiravel, 2019). Therefore, the development of psychological instrumentation and studies of variables that may interfere and / or enhance performance are relevant to obtain different levels of preparation. In addition, have knowledge about the anxiety latent trait of athletes can contribute to the identification of behaviors that can lead to high psychic suffering that requires an intervention of the psychologist (Molina, Sandín, and Charot, 2014).

The three-dimensional model of anxiety is structured by physiological anxiety, cognitive anxiety and perceived control, the latter being a regulatory dimension that represents the adaptive aspects of the phenomenon (Cheng et al., 2009). The physiological dimension aggregates the manifestations related to changes in organism activation, in order to act on the detected threat. The first facet (autonomic hyperactivity) concerns the symptoms expressed by the involuntary musculature, organs and glands (e. g. changes in heart and respiratory rate, cold sweat, etc.). The second facet (somatic tension) comprises the physiological manifestations expressed by the voluntary musculature, such as muscular tensions and fatigue (Cheng, Hardy and Markland, 2011).

The worry facet of cognitive dimension concerns the anticipation of negative consequences related to sport performance that lead the athlete to feel apprehensive (Cheng et al., 2009). The self-focus facet is defined as a self-evaluation state in which the athlete presents an exaggerated perception and criticism about his/her weaknesses (Teixeira, 2016). The perceived control dimension, on the other hand, is defined as the athlete's perception of his ability to achieve goals and a good performance even in the face of adverse situations (Cheng and Hardy, 2016; Teixeira, 2016).

A CFA with maximum likelihood minimization estimation method in a sample of 327 athletes from United Kingdom was conducted by Cheng et al. (2009) to verify the factorial validity of the Three Factor Anxiety Inventory (TFAI), instrument that operationalizes the three-dimensional model. The parcelling procedure was utilized due to the small sample size. All factorial loads were significant and between .79 and .35, with 80% of the

1 Federal University of Santa Catarina, Florianópolis, SC, Brazil. Correspondence: Psychological Assessment Research Laboratory, Psychology Department – Federal University of Santa Catarina. Rua Eng. Agronômico Andrei Cristian Ferreira, s/n - Trindade, Florianópolis - SC, 88040-900 – Brasil. E-mail: kkmclean@gmail.com. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

2 University of Lisbon, Faculty of Human Kinetics, Lisbon, Portugal

3 Universidade Federal de Santa Catarina (Brazil)

items with a load greater than .50. The parceled model with five first-order factors didn't obtain good fit, fact explained by the authors as a result of the high correlations between the facets of each dimension (above .90).

The model with only first-order factors revealed an adequate fit, both for the parceled (robust χ^2 (32) = 47.9, p = 0.01, RMSEA = 0.04, CFI = 0.99 and SRMR = 0.05), and non-parceled version (robust χ^2 (272) = 477.6, p <0.001, RMSEA = 0.05, CFI = 0.97 and SRMR = 0.07). The measure presented good indicators of reliability for the physiological, cognitive and control dimensions, with α = 0.75, α = 0.86 and α = 0.85, respectively.

TFAI was also cross-culturally adapted in a Chinese athletes sample, and their factorial validity and reliability were investigated (Cheng et al., 2011). Indicators from confirmatory analysis showed good fit of the model for the parceled (robust χ^2 = 24, 41.4, p <0.001, RMSEA = 0.056, CFI = 0.99 and SRMR = 0.047) and non-parceled version (robust χ^2 (186) = 303.4, p <0.001, RMSEA = 0.05, CFI = 0.98 and SRMR = 0.076). The factorial loads were significant, ranging from .42 to .80, with 19 items with loads above .50. The factors obtained Cronbach's alphas between .80 and .87.

There's a lack of research on measurements specific to Brazilian sports context (Peixoto, Nakano and Balbinotti, 2016) – information aligned with those provided by the Brazilian Federal Psychological Council, which regulates the utilization of psychological tests (Satepsi, 2019). Studies also point out that measures validated for use with clinical populations or in reference groups outside the sports context have been used in psychological assessment of athletes, compromising the quality of the results and the validity of the inferences made from these (Garcia and Borsa, 2016; Rubio, 2011; Silva et al., 2014).

However, efforts have been made to revert this situation and studies have focused on creating, validating and/or improving instruments in sporting context for Brazilian samples (e. g. Fernandes, et al., 2014; Fernandes et al., 2013; Fernandes, Vasconcelos-Raposo and Fernandes, 2012; Teixeira, 2016). Teixeira (2016) conducted the only Brazilian study to create a measure based on the Three-dimensional model of anxiety: The Three-Dimensional Anxiety Scale for Sport (*Escala Tridimensional de Ansiedade para o Esporte – ETApE* in Portuguese). An exploratory factor analysis with promax oblique rotation showed that the three first-order model best describes the data, a result that was in agreement with the results of Cheng et al. (2009, 2011).

In order to improve the instrument and verify the validity of the framework, the present study aimed to examine the reliability, evidence regarding internal structure and invariance of the ETApE through CFA and also partial credit model (IRT). The use of both aims at giving greater robustness to results, bringing the best of each theoretical contribution to improve the the test structure.

Method

Ethics

This research was submitted to the Ethics in Human Research Committee of the Federal University of Santa Catarina - Brazil, according to the norms of the Brazilian National Health Council (Resolution 466/12 of the National Health Council), and approved under the Certificate of Presentation for Ethical Appraisal 42882815.6.0000.0121.

Participants and procedure

The sample was formed by 500 athletes from the south of Brazil, being 64% males, with an average age of 26.5 years (SD = 9.7). A convenience sampling was adopted for data collection. The athletes were from 27 different sports, being 10 of them team sports. Athletics (29,12%), football (11,61%) and volleyball (10,57%) were the sports with more expression in the sample. The atheletes were of different competitive levels, ranging from international to local. The inclusion criteria were to present at least the primary school completed; to be 18 years of age or above; and have participated, at least, in a local competition in the last season.

The measure was administered individually or in small groups, at the university's premises or next to the training sites of athletes in a quiet and secluded location. The participants were informed about the research procedures and objectives and were given instructions to answer the instrument. The participants were asked to focus on how they felt in the most important competitions they had participated in the last or actual season that could be clearly remembered.

Measures

The ETApE (Teixeira, 2016), based on the theoretical model of Cheng et al. (2009), have 74 self-report items with a five-point Likert scale anchored at the extremes, ranging from "describes me very poorly" to "describes me very well", to access trait anxiety. Initial studies reported evidence of validity and Cronbach's alphas of α = 0.94 for physiological anxiety, α = 0.94 for cognitive anxiety and α = 0.89 for perceived control (Teixeira, 2016). A sociodemographic and profile questionnaire was also used to characterize the sample.

Data analysis

The dataset factorability was investigated through the statistical sample adequacy test of Kaiser-Meyer-Olkin (KMO) and Barlett test for sphericity (Hair, Black, Babin, Anderson and Tatham, 2009; Tabachnick and Fidell, 2007). The univariate normality was inspected by the items skewness and kurtosis, and an index for multivariate normality was generated (Kline, 2016). The presence of multivariate outliers was verified with Mahalanobis square distance (Arbuckle, 2016). A confirmatory factor analysis (CFA) with maximum likelihood method was conducted in AMOS 24.0 (SPSS Inc. Chicago, IL). Items with standardized regression weights below 0.50 were deleted from the initial item pool.

The refinement of the model was based on modification index values by Lagrange multipliers (LM), with trajectories and correlations above 14 ($p < 0.001$) as indicators of significant variation in model quality. In order to improve the model, items with LMs significatively high that weren't able to be corrected (e.g. errors of items belonging to different factors, or relations between errors and latent variables) were excluded.

After this step, item parameters were analysed separately for each dimension in light of the IRT Partial Credit Model with Winsteps (Linacre, 2014) and those with inappropriate psychometric properties were excluded. The following were assessed:

a. item fit statistics (infit and outfit), of which values of the mean square residual statistic up to 1.5 are considered indicators of a good fit (Linacre, 2014);
b. response category ordering (García-Pérez, 2017);
c. Rasch principal components analysis of the residuals. Contrasts with eigenvalues greater than 2.0, which are greater than expected with random data, show that items may be measuring more than one dimension; and
d. Person-item map, for inspection of the relative position of the item difficulty to person ability. The map also assists in checking for overlap between items, which do not add information about the level of the latent trait evaluated.

After the treatment of data by IRT, the model was reconfigured in AMOS and the CFA was executed again. The model with only first-order factors were compared with another containing two layers and second-order factors to verify which one best describes the data set. To estimate the adequacy of the overall model some goodness of fit indexes were used (Table 1).

Table 1.
Goodness-of-fit indexes

Index	Criteria	Interpretation	Reference
X^2/gl	< 5.0	acceptable model fit	Arbuckle, 2016; Kline, 2016
	< 3.0	good model fit	
SRMR	< 0.1	good model fit	Marôco, 2014
RMSEA	< 0.05	very good model fit	
CFI* and IFI	> 0.9	good model fit	Hair et al., 2009, Hu and Bentler, 1999
PCFI	> 0.8	good model fit	Marôco, 2014

*Note. *Considering the number of observed variables above 30 and a sample size greater than 250.*

Internal consistency was assessed by composite reliability with values greater than 0.70 demonstrating a good reliability of the instrument (Hair et al., 2009). The average variance extracted (AVE) wasn't estimated, because its cutoff criteria was considered too restrictive and relevant information could be lost. Malhotra and Dash (2011) argued that reliability can be established through composite reliability alone.

The model's invariance was tested with a stepwise strategy of increasing constraints. Configural, metric and scalar invariance were tested across male and female groups (n_{male}= 327; n_{female}= 165) through a multigroup confirmatory approach. For the configural invariance, it was verified if the factorial structure remains the same for the two groups. For the metric invariance, the model was constrained to be equivalent in both groups, and the scalar invariance was tested by constraining intercepts and structural covariances to be equivalent in the two groups (Milfont and Fischer, 2010).

Invariance is achieved when models do not present a statistically significant difference in the chi-square difference test (Gaskin, 2016; Putnick and Bornstein, 2016). However, according to Rudnev, Lytkina, Davidov, Schmidt and Zick (2018) the chi-square difference test tend to reject models with large sample sizes. For the situations with samples larger than 300 cases, it is recommended to look at the ΔCFI between models, with values below 0.01 indicating invariance (Chen, 2007; Cheung and Rensvold, 2002). The goodness-of-fit indexes also must be within the criteria established to indicate a good fit (Rudnev et al., 2018). After the assumption of invariance across genders was satisfied, a latent means comparison was performed, with the male group mean fixed to zero, being this group the reference one. The effect size was estimated through the Cohen's d and interpreted according to Cohen (1988) to verify the size of the difference between groups.

Results

The application of KMO to the dataset generated an indicator of 0.93, considered admirable. From the analysis of the KMO criterion, it is possible to say that there is enough shared variance to perform the factor analysis. The result of Bartlett's sphericity test for the dataset rejected H0, confirming the appropriateness of the application of the factor analysis [χ^2 (4851) = 25292.349, $p < 0.001$].

Skewness and kurtosis indicated univariate normality (Kline, 2016). However, there was a violation of multivariate normality, reported by the Mardia coefficient, which implied the use of bootstrapping techniques in order to adjust the *p* value of the chi-square statistic (Bollen and Stine, 1993). Moreover, the Mahalanobis d-squared indicates that there were no multivariate outliers (p1 and p2 ≤ 0.001).

After running the model in AMOS, eight items with regression weights below .50 and three items with LMs significatively high that weren't able to be corrected were dropped off. In the IRT analysis, two items were excluded because they presented inadequate fit statistics. Although the criterion for the exclusion of items with a regression weight below .50 was used, it was decided to maintain items 31 and 75, as the item-*theta* correlation was .55 for both items and the items are associated with levels of theta that are not represented by any other items (Figure 1). The exclusion of items 31 and 75 does not significantly impact the adjustment indicators ($\chi 2/gl$= 1.802; SRMR= 0.0532; RMSEA= 0.040, CFI= 0.941; PCFI= 0.843, IFI=0.942), and the decision to leave them has more benefits to the model, as they are at higher *theta* levels that are not well represented by any other item.

Just one item of the perceived control factor showed problems in the response category ordering in categories 1 and 2, but it was decided to keep it, since the disorder may be the result of the small sample, as well as the most expressive mean standard error for the first category (Linacre, 2014). The Rasch principal components analysis of the residuals showed that there were correlated standardized residuals between items for the three factors and contrasts with eigenvalues above the cutoff criteria. In order to avoid local dependence and respect the assumption of unidimensionality, eight items were excluded in total. After the exclusions, the contrasts showed eigenvalues below 2.0, which means that the factors presented unidimensionality.

With the assistance of the person-item map and the item parameters, overlappings between items for the three factors were found. Eleven items from the physiological factor, three from the cognitive factor, and two items from the perceived control factor were eliminated because they didn't add information to the measure and many of its contents were similar. After the exclusion of the overlapped items, new person-item maps were generated for each dimension (Figure 1). The optimized version of the instrument contains 14 items in the physiological factor (being 6 items from the somatic tension subcomponent and eight from autonomic hyperactivity), 10 in the cognitive factor (seven from worry and 3 from self-focus subcomponents) and 13 items in the perceived control factor, totalizing 37 items.

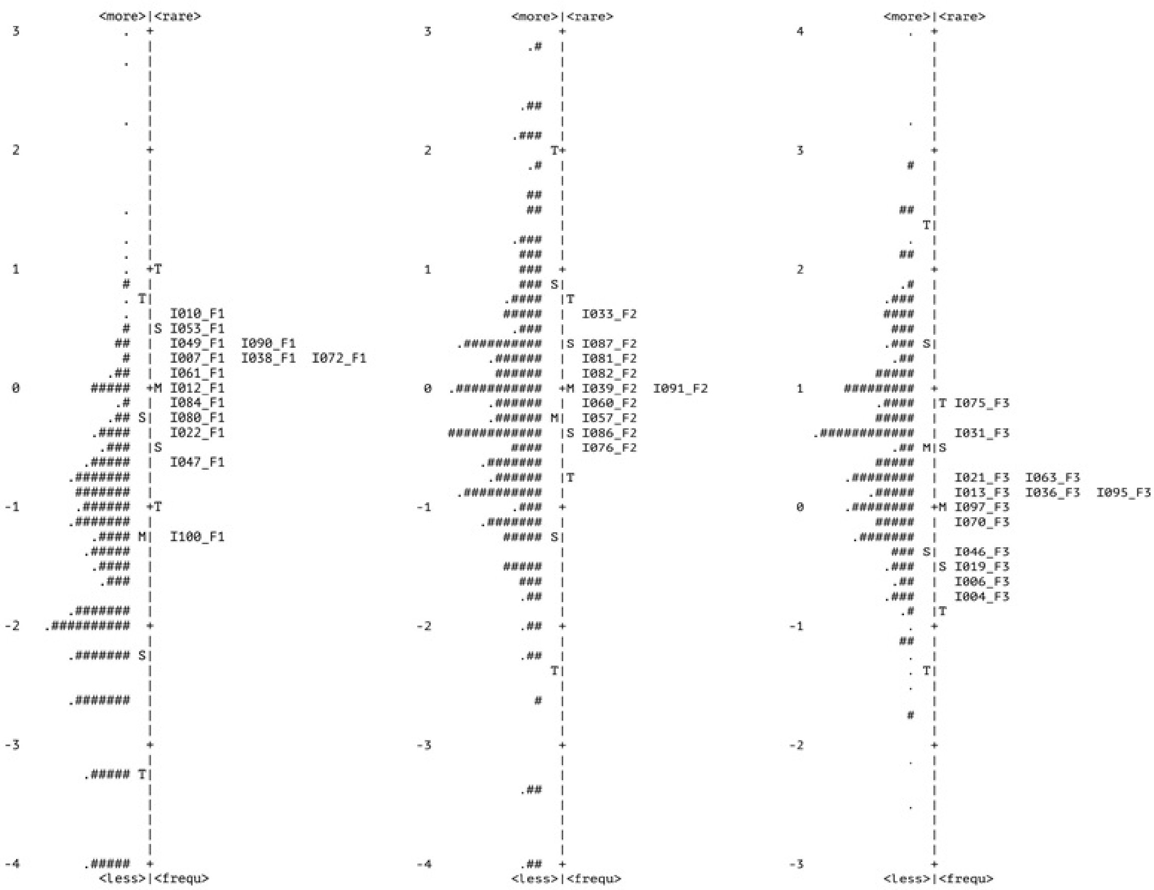

Figure 1. *Person-item-map of the dimensions Physiological anxiety, Cognitive anxiety and Perceived control, respectively*

The model with second-order factors presented good fit [χ2/gl= 1.891; SRMR= 0.054; RMSEA= 0.042, CFI= 0.929; PCFI= 0.844]. Although the model has reached good indicators, the correlations between the subcomponents are high, which suggests the union of them under a single factors. The correlation between somatic tension and autonomic hyperactivity was r = 0.97 (p < 0.001) with a MSV of 0.94, and r = 1.00 with p < 0.001 (MSV = 1.00) between worry and self-focus. The composite reliability for the first-order factors have ranged from 0.66 to 0.86.

The model with one layer and just first-order factors (Figure 2) was more promising and showed a better fit to the data [χ2/gl= 1.775; SRMR= 0.052; RMSEA= 0.039, CFI= 0.939; PCFI= 0.847; IFI= 0.939] (Marôco, 2014; Arbuckle, 2016; Hair et al., 2009; Kline, 2016). Table 2 show squared multiple correlations, standardized and unstandardized coefficients for first-order factor model.

Table 2.
Squared multiple correlations, standardized and unstandardized coefficients for first-order factor model

Observed variable	Latent construct	R^2	β	B	*p*	SE
ins_100	Physiological anxiety	0,27	0,52	1,16	***	0,12
ins_90	Physiological anxiety	0,51	0,71	1,36	***	0,12
ins_84	Physiological anxiety	0,38	0,61	1,24	***	0,12
ins_80	Physiological anxiety	0,53	0,73	1,44	***	0,12
ins_72	Physiological anxiety	0,52	0,72	1,28	***	0,11
ins_61	Physiological anxiety	0,49	0,70	1,25	***	0,10
ins_53	Physiological anxiety	0,57	0,76	1,24	***	0,10
ins_49	Physiological anxiety	0,52	0,72	1,35	***	0,12
ins_47	Physiological anxiety	0,29	0,54	1,11	***	0,11
ins_38	Physiological anxiety	0,43	0,65	1,21	***	0,11
ins_22	Physiological anxiety	0,26	0,51	1,08	***	0,12
ins_12	Physiological anxiety	0,45	0,67	1,27	***	0,11
ins_10	Physiological anxiety	0,45	0,67	1,00	***	0,07
ins_7	Physiological anxiety	0,29	0,54	1,00		
ins_91	Cognitive anxiety	0,56	0,75	1,04	***	0,07
ins_86	Cognitive anxiety	0,31	0,56	0,83	***	0,08
ins_82	Cognitive anxiety	0,41	0,64	0,96	***	0,08
ins_81	Cognitive anxiety	0,56	0,75	1,12	***	0,08
ins_76	Cognitive anxiety	0,35	0,59	0,91	***	0,08
ins_60	Cognitive anxiety	0,56	0,75	1,09	***	0,08
ins_57	Cognitive anxiety	0,30	0,55	0,83	***	0,08
ins_39	Cognitive anxiety	0,46	0,68	1,00	***	0,08
ins_33	Cognitive anxiety	0,42	0,65	1,00		
ins_95	Perceived control	0,29	0,54	1,16	***	0,13
ins_75	Perceived control	0,23	0,48	1,05	***	0,13
ins_70	Perceived control	0,46	0,68	1,37	***	0,14
ins_63	Perceived control	0,29	0,54	1,15	***	0,13
ins_46	Perceived control	0,39	0,62	1,33	***	0,14
ins_36	Perceived control	0,36	0,60	1,28	***	0,14
ins_31	Perceived control	0,22	0,47	1,08	***	0,14
ins_21	Perceived control	0,26	0,51	1,13	***	0,13
ins_19	Perceived control	0,43	0,66	1,30	***	0,13
ins_13	Perceived control	0,28	0,53	1,15	***	0,13
ins_6	Perceived control	0,43	0,66	1,26	***	0,11
ins_4	Perceived control	0,25	0,50	1,00		
ins_87	Cognitive anxiety	0,54	0,74	1,03	***	0,07
ins_97	Perceived control	0,58	0,76	1,65	***	0,16

The composite reliability achieved values above the thresholds recommended by the specialized literature, being 0.91 for Physiological anxiety, 0.88 for Cognitive anxiety and 0.86 for Perceived control (Hair et al., 2009). Since the model with only three first-order factors obtained the best indicators, its invariance was tested.

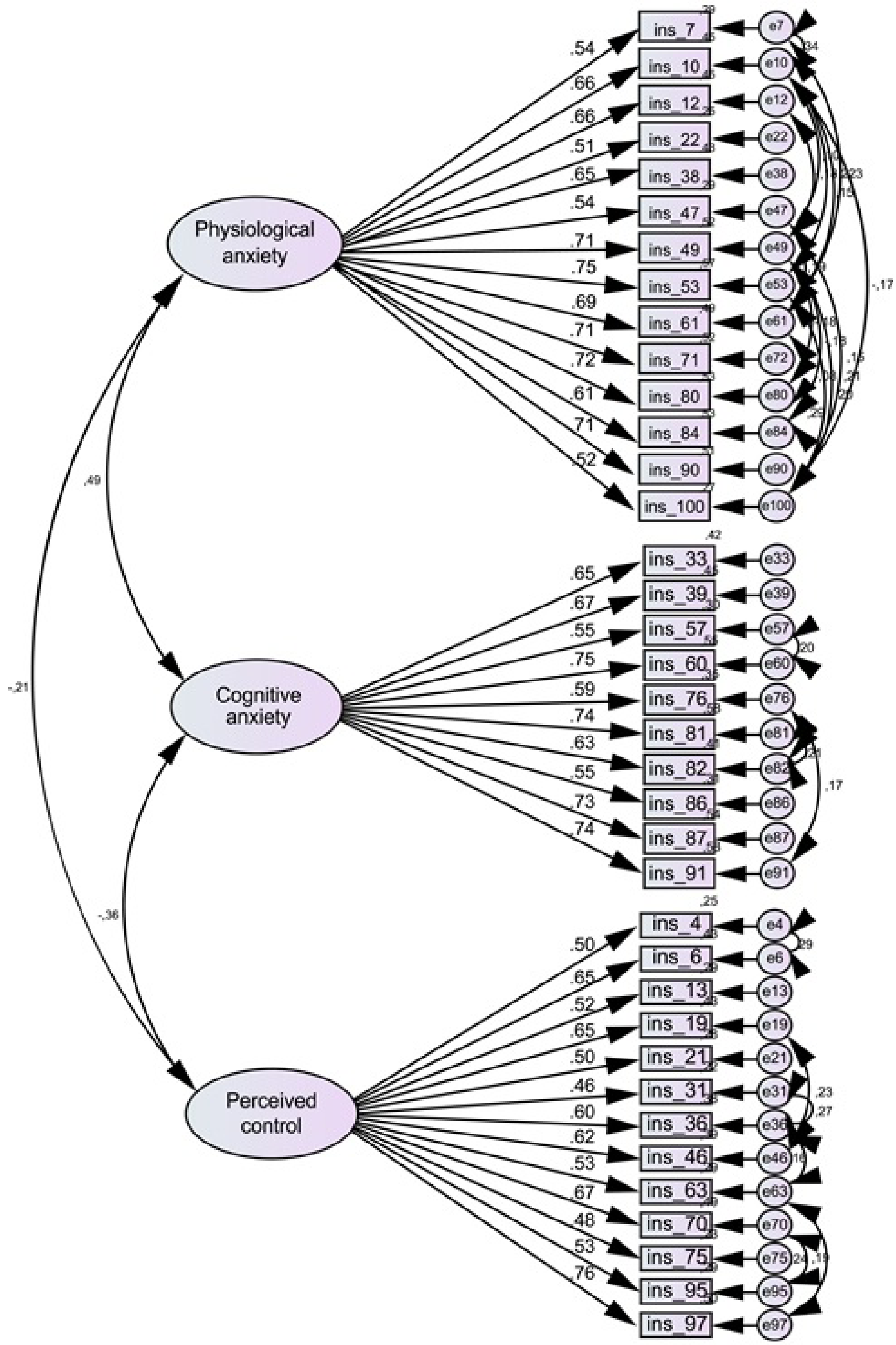

Figure 2. *Model with the most promising indicators*

Configural invariance was supported across groups, evidenced by good model fit when estimating two groups without constraints [χ2/gl= 1.419; SRMR= 0.058; RMSEA= 0.029, CFI= 0.933; PCFI= 0.842]. Metric invariance was corroborated by the goodness-of-fit measures [χ^2/gl= 1.410; SRMR= 0.060; RMSEA= 0.029, CFI= 0.932; PCFI= 0.867] and supported by a non-significant chi-square difference [χ^2 (37)= 40.977, p= 0.3] between unconstrained constrained models (Byrne, 2010). Chi-square difference test for the scalar invariance resulted in a significant p-value. However, the change in the CFI was 0.01 (Table 3), evidencing that the scalar invariance is supported in the model. The goodness-of-fit measures was also acceptable when the intercepts were constrained [χ2/gl= 1.458; SRMR= 0.062; RMSEA= 0.031, CFI= 0.922; PCFI= 0.883].

Table 3.
Goodness-of-fit indices for invariance tests

	χ^2/gl	CFI	ΔCFI	RMSEA
Configural	1.41	0.933	-	0.029
Metric	1.41	0.932	-0.001	0.029
Scalar	1.45	0.922	0.01	0.031

Regarding the multigroup analysis between genders, a latent means difference between male and female athletes was found for the three dimensions. The means of physiological and cognitive anxiety were significatively higher for the female group, having the first a medium and the second a small effect size. The latent mean for the perceived control dimension was significatively higher for the male group, with a medium effect size (Table 4).

Table 4.
Latent means comparison, according to gender

Factor	Difference of latent means	p-value	Effect size (Cohen's *d*)
Physiological	0.303	< 0.001	-0.49
Cognitive	0.255	0.005	-0.29
Perceived control	-0.249	< 0.001	0.47

Discussion

The aim of the study was examining the reliability, evidences of validity regarding the internal structure and the invariance of the ETApE with the use of CFA and Rasch analysis. In terms of the theoretical model, there is evidence of its validity also with south brazilian athletes and the results corroborate other studies about the three-dimensional model in other cultures (Cheng et al., 2011; Cheng et al., 2009). Convenience sampling may limit the generalization of results and additional studies are recommended to verify that the results are replicable. Besides that, More research should be done to consider that the model is valid in Brazilian culture as a whole, since the sample used in this study was formed mostly by athletes from the southern region of the country.

Two models were tested, but as it was found by Cheng et al. (2011; 2009), the structure that best describes the data is that of three factors, instead of the hierarchical five-dimensional. The correlations between the subcomponents were too high, which suggests that they share closely related content and therefore cannot be considered distinct. The physiological manifestations of anxiety can present an underlying mechanism that may be overlapped between the subcomponents, which explains the high correlation (Cheng et al., 2011). In the case of the worry and self-focus subcomponents, both are negatively toned and may have uncertainty as a common influence. Uncertainty is increased when the threats are not explicit and increases the vulnerability, being a trigger to worries and self-evaluation, because gives scope to interpretation and thoughts about how things could be (Chen, Yao and Qian, 2018).

The deletion of half of the items helped to achieve a parsimonious version of the instrument. The model with 37 items presented good indicators of validity regarding the internal structure and invariance, as well as composite reliability. Choosing only the best items and avoiding overlapping helps to prevent examinee fatigue with items that were not able to add new information and to increase their engagement with the task (Oswald, McAbee, Redick and Hambrick, 2015). However, as highlighted by the inspection of the person-item map, there are *theta* levels that are not well represented, which makes the instrument's reliability lower in areas where there are no items. Therefore, to improve the measure and adequately represent the extent of the construct, one solution would be to develop new items for *theta* areas that are not well covered by the current items (Bond and Fox, 2015).

As pointed out, two items presented factorial loads below .50 in the confirmatory factorial analysis and were candidates for exclusion. However, the item-*theta* correlation provided by the IRT showed values of .55 for both items. This correlation indicates the association between the item and the latent variable and, in turn, the item's relevance for construct evaluation. The item-*theta* correlation justifies the decision of keeping the items, since they have a correlation above .50 with the factor (Embretson and Reise, 2000).

With the assistance of the person-item-map based on the Rasch model, it was verified that the items were associated with levels of *theta* that are not represented by any other item. Besides that, the items 31 and 75 are those that represent the highest levels of the construct in the set of measurement items. The exclusion of both items would result in loss of amplitude of the measure in .68 logits.

This context reinforces the need to use more of an analysis strategy that evaluates both the instrument in a global way and the parameters of the items.

Regarding the multigroup analysis of the measure, the configural, metric and scalar invariance were reached, which means that the model structure and parameters are invariant across gender groups. These results point to the absence of significant bias among gender and that the differences found between them for the scores of each dimension are the result of real differences between the groups. The achievement of scalar invariance allows to say that the dimensions have the same meaning for both groups and allows the latent means comparison (Arbuckle, 2016).

A statistically significant difference was found when the mean score of men and women were compared for the three dimensions. According to the results, women presented a significantly higher physiological and cognitive anxiety mean when compared to men, with medium and small effect size respectively. The same result was found in other studies in sport (Allawy, 2013; Marcel and Paquet, 2010). In the other hand, women exhibited significantly lower perceived control mean compared to men's group. Taking into account the practical aspects, the measure could be used in different stages of the season to assess the anxiety levels of the athletes and more specifically to design interventions focused on the specific needs of each group in order to develop competencies that allow them to handle with the anxiogenic situations of the competitive context and cope (Kristjánsdóttir, Elingsdóttir, Sveinsson and Saavedra, 2018).

Overall, the instrument presented good psychometric qualities that point to its validity regarding internal structure. The indicators showed adequacy of the model to the data set, which reinforces its theoretical solidity. Good indicators of reliability and factorial invariance were also achieved. Taking into account that a posteriori respecifications have been made, studies that replicate the factorial structure are necessary to raise more evidence of its validity. It is also suggested that the model structure be tested in athletes from different regions of the country in order to verify its invariance in relation to the different cultural contexts.

Acknowledgments and Financial Support

This work was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

Propiedades psicométricas de la Escala tridimensional de ansiedad para el deporte (ETApE) mediante los enfoques de CFA e IRT

Resumen

El presente estudio tuvo como objetivo perfeccionar y examinar la confiabilidad, las evidencias sobre la estructura interna y la invariancia de la ETApE, una escala de ansiedad rasgo para el deporte. La muestra del estudio fue formada por 500 atletas del sur de Brasil, siendo 64% hombres, de diferentes deportes y niveles competitivos, que van desde lo internacional a lo local. Para examinar las propiedades psicométricas del modelo de medición, se realizó un análisis factorial confirmatorio (AFC) con el método de máxima verosimilitud y se evaluó la confiabilidad compuesta. Los parámetros de los ítems fueron analizados a la luz de la Teoría de Respuesta al Ítem (TRI). Además, se probó la invariancia del modelo y se realizó una comparación de medios latentes, según el género. El instrumento presentó buenas cualidades psicométricas que apuntan a su validez sobre la estructura interna. También se lograron buenos indicadores de confiabilidad e invarianza factorial. Se encontró una diferencia estadísticamente significativa en las tres dimensiones del instrumento según el sexo. Las mujeres obtuvieron una puntuación más alta en la ansiedad cognitiva y fisiológica y una puntuación más baja en la dimensión de control percibido.

Palabras-clave: ansiedad; deporte; validez; confiabilidad; invariancia

Propriedades psicométricas da Escala Tridimensional de Ansiedade para o Esporte (ETApE) mediante abordagens de CFA e IRT

Resumo

O presente estudo objetivou aprimorar e examinar a confiabilidade, as evidências relacionadas à estrutura interna e invariância da ETApE, uma escala de ansiedade traço para o esporte. A amostra do estudo foi formada por 500 atletas do sul do Brasil, sendo 64% do sexo masculino, de diferentes esportes e níveis competitivos, variando entre internacional e local. Para examinar as propriedades psicométricas do modelo de medida, uma análise fatorial confirmatória (AFC) com método de máxima verossimilhança foi conduzida e a confiabilidade compósita foi avaliada. Os parâmetros dos itens foram analisados à luz da Teoria de Resposta ao Item (TRI). Ademais, a invariância do modelo foi testada e uma comparação de médias latentes, em função do sexo, foi realizada. O instrumento apresentou boas propriedades psicométricas que apontam para sua validade relacionada à estrutura interna. Bons indicadores de confiabilidade e invariância também foram atingidos. Uma diferença estatisticamente significativa foi encontrada para as três dimensões do instrumento em função do sexo. As mulheres obtiveram escores mais altos em ansiedade cognitiva e fisiológica e escores mais baixos na dimensão de controle percebido.

Palavras-chave: ansiedade; esporte; validade; confiabilidade; invariância

References

Allawy, M. H. (2013). Anxiety among egyptian athletes as measured by the 'arabic precompetitive state anxiety inventory'. *Acta Kinesiologica*. *7*(2), 70-72.

Arbuckle, J. L. (2016). IBM® SPSS® Amos™ 24 user's guide. Chicago, IL: IBM.

Bollen, K. A., and Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In K. Bollen and J. Long (Eds.), Testing structural equation models (pp. 111-135). Newbury Park, CA: Sage Focus Edition.

Bond, T. G., and Fox, C. M. (2015). *Applying the Rasch Model*: Fundamental Measurement in the Human Sciences. New York: Routledge.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: basic concepts, applications, and programming* (2nd ed.). Nova Iorque: Routledge, Taylor and Francis.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling*, *14*(3), 464-504. doi: 10.1080/10705510701301834.

Chen, S., Yao, N., and Qian, M. (2018). The influence of uncertainty and intolerance of uncertainty on anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, *61*, 60-65. doi: 10.1016/j.jbtep.2018.06.005.

Cheng, W. K., and Hardy, L. (2016). Three-dimensional model of performance anxiety: tests of the adaptative potential of the regulatory dimension of anxiety. *Psychology of Sport and Exercise*, *22*, 255-263. doi: 10.1016/j.psychsport.2015.07.006.

Cheng, W. K., Hardy, L., and Markland, D. (2009). Toward a threedimensional conceptualization of performance anxiety: rationale and initial measurement development. *Psychology of Sport and Exercise*. *10*, 271-278. doi: 10.1016/j.psychsport.2008.08.001.

Cheng, W. K., Hardy, L., and Markland, D. (2011). Cross-cultural validation of a three-dimensional measurement modelo f performance anxiety in the context of Chinese sports. *International Journal of Sport Psychology*, *42*, 417-435.

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255. doi: 10.1207/S15328007SEM0902_5.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Embretson, S., and Reise, S. (2000). *Item Response Theory for Psychologists.* Mahwah, NJ: Lawrence Erlbaum.

Fernandes, M. G., Nunes, S. A. N., Vasconcelos-Raposo, J., and Fernandes, H. M. (2014). Efeitos da experiência nas dimensões de intensidade, direção e frequência da ansiedade e autoconfiança competitiva: um estudo em atletas de desportos individuais e coletivos. *Motricidade*, *10*(2), 81-89. doi: 10.6063/motricidade.10(2).2930.

Fernandes, M. G., Nunes, S., Vasconcelos-Raposo, J., Fernandes, H. M., and Brustad, R. (2013). The CSAI-2: an examination of the instrument's factorial validity and reliability of the intensity, direction and frequency dimensions with Brazilian athletes. *Journal of Applied Sport Psychology*, *25*(4), 337-391. doi: 10.1080/10413200.2012.744780.

Fernandes, M. G., Vasconcelos-Raposo, J., and Fernandes, H. M. (2012). Propriedades sicométricas do CSAI-2 em atletas brasileiros. *Psicologia: Reflexão e Crítica*, *25*(4), 679-687. doi: 10.1590/S0102-79722012000400007.

Garcia, R. P., and Borsa, J. C. (2016). A prática da avaliação psicológica em contextos esportivos. *Temas em Psicologia*, *24*(4), 1549-1560. doi: 10.9788/TP2016.4-20.

García-Perez, M. (2017). Analysis of (dis)ordered categories, thresholds, and crossing in difference and divide-by-total IRT models for ordered responses. *The Spanish Journal of Psychology*, *20*(10), 1-27. doi: 10.1017/sjp.2017.11.

Gaskin, J. (2016). Measurement model invariance. Gaskination's StatWiki. http://statwiki.kolobkreations.com.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2009). *Análise Multivariada de Dados*. Porto Alegre: Bookman.

Hamlin, M. J., Wilkes, D., Elliot, C. A., Lizamore, C. A., and Kathiravel, Y. (2019). Monitoring Training Loads and Perceived Stress in Young Elite University Athletes. *Frontiers in Physiology*, *10*(34), 1-12. doi: 10.3389/fphys.2019.00034.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Kline, R. B. (2016). Principles and practice of structural equation modeling (4th ed.). New York: The Guilford Press.

Kristjánsdóttir, H., Erlingsdóttir, A. V., Sveinsson, G., and Saavedra, J. M. (2018). Psychological skills, mental toughness and anxiety in elite handball players. *Personality and Individual Differences*, *134*, 125-130. doi: 10.1016/j.paid.2018.06.011.

Linacre, J. M. (2014). *Winsteps Rasch measurement computer program*. Beaverton: Winsteps.com.

Malhotra, N. K., and Dash, S. (2011). *Marketing Research an Applied Orientation*. London: Pearson Publishing.

Marcel, J., and Paquet, Y. (2010). Validation française de la version modifiée du 'Sport Anxiety Scale' (SAS). *L'Encéphale*, *36*, 116- 121. doi : 10.1016/j.encep.2009.05.006.

Marôco, J. (2014). Análise de Equações Estruturais: fundamentos teóricos, software and aplicações. Lisboa: ReportNumber.

Milfont, T. L., and Fischer, R., (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, *3* (1), 111-121. doi: 10.21500/20112084.857.

Molina, J., Sandín, B., and Chorot, P. (2014). Sensibilidad a la ansiedad y presión psicológica: efectos sobre el rendimiento deportivo em adolescentes. Cuadernos de Psicología del Deporte, 14(1), 45-54.

Oswald, F. L., McAbee, S. T., Redick, T. S., and Hambrick, D. Z. (2015). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343-1355. doi: 10.3758/s13428-014-0543-2.

Peixoto, E. M., Nakano, T. C., and Balbinotti, M. A. A. (2016). *Novas perspectivas para avaliação em psicologia do esporte e do exercício físico*. Curitiba: CRV.

Putnick, D. L., and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90. doi: 10.1016/j.dr.2016.06.004.

Rubio, K. (2011). A avaliação em Psicologia do Esporte e a busca de indicadores de rendimento. In: Angelo, L. F., and Rúbio, K. (Orgs.). *Instrumentos de avaliação em Psicologia do Esporte* (pp. 13-26). São Paulo: Casa do Psicólogo.

Rudnev, M., Lytkina, E., Davidov, E., Schmidt, P., and Zick, A. (2018). Testing measurement invariance for a second-order factor: a cross-national test of the Alienation Scale. *Methods, Data, Analyses*, *12*(1), p. 47-76. doi: 10.12758/mda.2017.11.

Satepsi. *Lista completa dos testes*. 2019. Disponível em: <http://satepsi.cfp.org.br/listaTeste.cfm>. Acesso em: 22 abr. 2019.

Silva, A. M. B., Foch, G. F. L., Guimarães, C. A., and Enumo, S. R. F. (2014). Instrumentos aplicados em estudos brasileiros em psicologia do esporte. *Estudos Interdisciplinares em Psicologia*, *5*(2), 77-95. doi: 10.5433/2236-6407.2014v5n2p77.

Teixeira, K. C. (2016). Construção e busca de evidências de validade e precisão de uma medida de ansiedade para atletas. (Dissertação de mestrado). Universidade Federal de Santa Catarina, Florianópolis.

Weinberg, R. S., and Gould, Gould, D. (2018). *Foundations of Sport and Exercise Psychology*. Champaign: Human Kinetics.