Original

New approaches on the study of the psychometric properties of the STAI

Javier Ortuño-Sierra[1]
Lorena García-Velasco[2]
Félix Inchausti[3]
Martin Debbané[4,5]
Eduardo Fonseca-Pedrero[6,7]

[1]Department of Psychology, Universidad Loyola Andalucía, Spain
[2]Department of Psychology, International University of La Rioja, Spain
[3]Hospital Benito Menni of Elizondo, Spain
[4]Office Médico-Pédagogique, Research Unit, Department of Psychiatry, University of Geneva School of Medicine, Switzerland
[5]Research Department of Clinical, Educational and Health Psychology, University College London, United Kingdom
[6]Department of Educational Sciences, University of La Rioja, Spain
[7]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM)

Introduction. The main purpose of this study was to analyze the psychometric properties of the State-Trait Anxiety Inventory (STAI[1]). Previous studies have indicated different factor solutions. Nevertheless, there is still a lack of consensus about the best dimensional model of STAI scores.

Method. The sample consisted of 417 participants, composed of 387 (29.71% male) healthy participants (comparison group: *M*=35.5 years; *SD*=8.40), and 30 (36.66% male) patient (clinical group *M*=35.8 years; *SD*=12.94).

Results. The internal consistency evaluated through Ordinal Alpha was good, 0.98 and 0.94 in the non-clinical and the clinical samples, respectively. Test-retest reliability (two weeks) for Total Score was 0.81 for the non-clinical subsample, and 0.93 for the clinical subsample. Confirmatory factor analyses supported both a four factor model and bifactor model. Also, STAI scores showed statistically significant correlations with *Burns Anxiety Inventory* (Burns-A) scores. Furthermore, results showed statistically significant differences in the mean scores of the STAI between the clinical and the non-clinical subsamples.

Conclusions. The psychometric properties of the STAI were adequate. The present study contributes to better understand the STAI structure through the comparison of new approaches in the study of the STAI internal structure. The results found may contribute in the efforts to improve the evaluation and identification of anxiety symptoms and disorders.

Keywords: Anxiety, STAI, Validation, Psychometric properties, Self-report, Validity

Actas Esp Psiquiatr 2016;44(3):83-92

Correspondence:
Javier Ortuño-Sierra
Department of Psychology
Universidad Loyola Andalucía
C/ Energía Solar 1, Edificio G
41014 Sevilla, Spain
E-mail: jortuno@uloyola.es

Nuevas aproximaciones en el estudio de las propiedades psicométricas del STAI

Introducción. El objetivo principal de este estudio fue analizar las propiedades psicométricas del *State-Trait Anxiety Inventory* (STAI[1]) (Inventario de Ansiedad Estado Rasgo). Estudios previos han encontrados diferentes soluciones factoriales; sin embargo, existe una falta de consenso acerca de cuál es la mejor solución factorial que subyace a las puntuaciones del STAI.

Método. La muestra consistió en 417 participantes de los cuales 387 (29,71% hombres) eran no-clínicos (*M*=35,5 años; *DT*=8,40) y 30 (36,66% hombres) pacientes (grupo clínico *M*=35,8 años; *DT*=12,94).

Resultados. La consistencia interna estimada mediante el alfa ordinal para la puntuación total fue de 0,98 y 0,94 en las muestras no-clínica y clínica, respectivamente. De igual forma, la fiabilidad test-retest (2 semanas) para la puntuación Total fue 0,81 para la muestra no-clínica y 0,93 para la muestra clínica. Los análisis factoriales confirmatorios realizados revelaron que la estructura de cuatro factores y un modelo bifactor mostraron adecuados índices de bondad de ajuste. Asimismo, las puntuaciones del STAI mostraron correlaciones significativas con las puntuaciones del Inventario de Ansiedad de Burns (*Burns Anxiety Inventory* Burns-A). Los resultados mostraron, de igual forma, diferencias estadísticamente significativas entre las puntuaciones medias del STAI entre la muestra no-clínica y la muestra clínica.

Conclusiones. Las propiedades psicométricas del STAI fueron adecuadas. El presente estudio contribuye a una mejor comprensión de la estructura factorial subyacente al STAI mediante el análisis de nuevas aproximaciones al estudio de su estructura interna. Los hallazgos encontrados pueden ayudar en los esfuerzos por mejorar la evaluación e identificación de los síntomas y los trastornos de ansiedad.

Palabras Clave: Ansiedad, STAI, Validación, Propiedades psicométricas, Auto-informe, Evidencias de validez

INTRODUCTION

Among emotional disorders, anxiety is the most prevalent in general population[2-5], and will, likely, become one of the leading causes of disability in XXI century in European countries[6]. For instance, an international research carried out among 2001-2003, in 14 countries of America, Europe and Asia, on a total of 60463 adults participants, revealed that anxiety disorders were the most frequent in almost all the countries, with prevalence rates ranging from 2,4% to 18,2%[7]. In another relevant study, Alonso et al.[6], with a representative sample of 21425 adults belonging to six European countries, found that vital prevalence for any anxiety disorder was 13.6%. The review of Somers et al.[8] indicated year-prevalence rates and life-prevalence rates of anxiety disorders between 10.6 and 16.6%. In the same line, the research conducted by Bloom[9] showed that up to 16% of the population reported some type of anxiety problems.

Crucial to identification and intervention efforts in anxiety disorders and symptoms is the existence of well-validated, psychometrically sound assessment tools. The State-Trait Anxiety Inventory (STAI) is a self-reported instrument widely used to evaluate anxiety trait and anxiety state in both general and clinical population[10], being one of the most used to this extent among Spanish psychologist[11]. The STAI, has been translated into more than 40 languages. The Spanish adaptation version was made through the work of Bermudez[12,13]. The STAI is composed by two subscales. Each of the subscales (trait and state) has 20 items. Some of them are positive and others are written in a negative way. There are a lot of studies in the review of the literature that have analyzed the psychometric properties of the STAI scores with regards to the internal consistency, the test-retest reliability, and different sources of validity evidence[14-20].

Different studies about the internal structure of the STAI scores have found a three dimensional or mix structure (positive state anxiety, negative state anxiety, and trait anxiety)[16,20,21]. Other studies point out that the dimensional structure of the STAI could be determined by the nature of the items[18,19]. The STAI has items formulated both in a positive and in a negative way, in order to avoid bias effects (i.e. acquiescence). As a result, the dimensional structures might respond to a statistical artifact or measure bias, questioning its empirical validity[18,19]. With this regard, some researchers suggest a model of two different factors and two different methods, where anxiety state and trait would be the constructs and positive and negative polarity the methods[19]. This factorial structure has received support in different studies[17,22,23]. It is worth noting that the majority of the studies found in the literature analyzing the internal structure of the tool were conducted with data considered as continuous and, therefore, using Maximun Likelihood Method (MLM) estimator. However, the STAI is administered with a Likert-type response format with four options, that its, ordinal data.

Recently, a bifactor model has been found to better fit the data[14] in the case of anxiety trait factor. The bifactor approach incorporates a general factor underlying all variables (e.g., general anxiety) as well as a specific factor for each variable; moreover, the bifactor model allows including uncorrelated group factors (e.g., Anxiety-Trait, Anxiety-State). The group variables in a bifactor model are not subsumed by the general factor (e.g., negative affectivity), and group factors are conceptualized as uncorrelated and distinct, given the presence of the general factor accounting for all covariance among items in the model[24]. Researchers have recently begun to apply the bifactor model to the study of psychological constructs showing that bifactor models adequately represent psychological constructs[24]. To date, few studies have studied the adequacy of the bifactor model in order to explain the STAI scores. For instance, Bados et al.[14] found that a bifactor model explained better the dimensional structure of the Anxiety-Trait dimension.

As can be seen, results are still contradictory and new studies are needed in order to capture the dimensional structure of the STAI. Due to the fact that previous studies have considered data as continuous, new studies that account the ordinal nature of the data are still needed. Furthermore, new techniques, as it is the case of the bifactor approach, have not been widely analysed and could better explain the internal structure of the STAI scores. Moreover, it is interesting to test the validity of the STAI in order to differentiate clinical and non-clinical samples. Also, it is important to analyse the relation of the STAI with other measuring instrument in order to gather new sources of validity evidence. Within this research framework, the main objective of the present work is to study the psychometric properties of the STAI in clinical and non-clinical population. We therefore study: a) the internal consistency and the test-retest of the STAI scores, b) the dimensional structure of the STAI scores using Confirmatory Factor Analysis (CFAs) and considering data both as ordinal and continuous, c) the relationship between the Burns Anxiety Inventory-A[25] and the STAI scores, and d) the discriminant validity between a clinical and a non-clinical subsample. It is hypothesised that the bifactor solution and the four-factor model will result in a better model fit. It is also hypothesised that STAI scores would show adequate levels of internal consistency and stability in both samples. It is further hypothesised that the STAI scores will be associated to other measures of Anxiety (e.g., Burns-A) and that non-clinical group will score lower than the clinical group in the STAI mean scores.

METHOD

Participants

The sample comprised a total of 417 non-clinical and clinical adults. Participants volunteered to take part in the study (convenient samples). Non-clinical sample was composed by 387 adults, 115 were male (29.72%). Participants' ages ranged from 18 to 72 years (*M*=35.47 years; *SD*=8.4). Participants belonging to several Spanish communities, with more participation from La Rioja (30.23%), followed by Catalonia (28.42%) and Madrid (13.96). Attending to the study level, a 77.26% had university studies, a 16.02% had professional studies, and 4.6% had secondary level. The initial sampling was formed by 429 participants, eliminating those participants that were taking some type of medication for anxiety (*n*=30) and presented outlier scores in the STAI (*n*=12). The clinical sample was composed of 30 participants that at the moment of the study were diagnosed with some anxiety disorder according to the DSM-IV Manual[26]. Participants in this subsample completed the questionnaires before starting intervention in the *Psychology Centre BCN*, 11 were male (36.66%). Participants' ages ranged from 18 to 61 years (*M*=35.8 years; *SD*=12.94). All participants were living in Catalonia. Attending to the study level, a 70% had university studies, a 23.33% had professional studies, and a 6.6% had secondary level studies. At the moment of the research, 14 participants (46.66%) were taking some type of medication for anxiety.

Instruments

State-Trait Anxiety Inventory (STAI)[1,22]. The STAI is a self-reported questionnaire composed by 40 items developed with the aim of evaluating two different types of anxiety: state anxiety (emotional condition transitory), whose reference frame is the "now, at this moment" (20 items), and the anxiety trait (anxiety tendency relatively stable), whose reference frame is "in general, in most of the times". The STAI has a Likert-type response format with four options (0=*almost never/nothing*; 1=*some/some times*; 2=*quite/often*; 3=*a lot/almost always).* Score in each subscale ranges from 0 to 60. The STAI is a tool widely used for the screening of state-anxiety and trait-anxiety in non-clinical and clinical population, being one of the most used among clinical psychologist[11]. In the present study we have used the X version of the STAI. The STAI Spanish version has been reported to have adequate psychometric properties with a Cronbach's Alpha of 0.93 for the Total Score[21]. In addition, evidences of its internal structure have been reported for a three and a four-dimensional structure[17,21].

Burns Anxiety Inventory (Burns-A)[25]. The Burns-A is a measuring instrument composed by 33 items that refer to anxiety symptoms. The Burns-A consists of three subscales: Anxious Feelings (6 items), Anxious Thoughts (11 items), and Physical Symptoms (16 items). Anxious Feelings are defined like "anxiety, nervousness, fear or worry". Anxious Thoughts include "difficulties to focus or fear to be alone, isolated form others or to be abandoned". Physical symptoms is composed for 16 items including "pain, oppression or thoracic constriction" among others. Participants have to respond about how they have experimented or have been worried about each symptom in the last days, in a Likert-type respond format with 4 options (0=*not at all* to 3=*a lot*). The sum of all the items forms the Total Anxiety Score. A score from 0 to 4 show minimum anxiety whereas a score from 55 to 99 indicates extreme anxiety. Spanish version of the Burns-A was used in this study. Psychometric properties of the Burns-A have been studied[27].

Procedure

Sampling method varied according to each of the subsamples. In this way, the non-clinical subsample was obtained through the use of new information and communication technologies. Collaboration in the study was requested through different media (social networks, chats and e-mail). Socio-demographic data and written consent were collected from every participant and, in addition, all of them were given a code. As inclusion criteria for the total sample, participants had to be Spanish and over 18 years. As regards to the non-clinical sample, participants had not to have been diagnosed of any anxiety disorder, whereas for the clinical sample, participant had to have a diagnosis of an anxiety disorder in the *Psychological Centre* BCN. The *Psychological Centre* (BCN) is a clinical centre focused on evaluation, diagnosis and treatment of children, adolescents, and adult population.

Data analysis

First, we calculated descriptive statistics for the subscales of the STAI. In addition, Ordinal alpha was calculated as a measured of the internal consistency of the scores in both subsamples. Ordinal alpha is conceptually equivalent to Cronbach's alpha and it is more adequate for dichotomous and ordinal data[28]. Also, we analyzed the test-retest reliability in both subsamples through Intraclass Correlation Coefficient (ICC). Participants of the total sample were asked to complete it again, 15 days after being administered the STAI. In the non-clinical sample, 186 participants completed the retest form, while all participants of the clinical sample completed for the second time the STAI.

Second, with the aim of studying the internal structure of the STAI, several competing models were tested by means of CFAs. The first model to be tested was a one-factor model (model 1). This model is the most parsimonious, and in addition expresses the hypothesis of a single dimension underlying of the STAI scores, rather than two separate dimensions. Second, a two-dimensional model with anxiety state and anxiety trait as two separate dimensions was tested (model 2). A three dimensional or mix structure (positive state anxiety, negative state anxiety and trait anxiety) (model 3)[16,20,21] was also tested. In addition, the four dimensional structure was tested (model 4)[17,22,23]. Attending to the new approaches in the research about anxiety models, we decided to study the bifactor approach. We intended to analyse two different models under this approach. On the one hand, a bifactor model that account for the content of the items, that was, therefore, composed by two different dimensions (positive and negative items), plus a general factor (model 5). On the other hand, we tested a bifactor model including an anxiety-state, an anxiety trait, and a general factor (model 6).

Due to the categorical nature of the data, we used the weighted least squares means and variance adjusted (WLSMV) for the estimation of parameters[29]. With the aim to compare the adequacy of different estimators, we also estimated parameters with data considered as continuous, by means of the MLM estimator. The following goodness-of-fit indices were used: Chi-square (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA) (and 90% confidence interval), and Weighted Root Mean Square Residual (WRMR) in the case of the WLSMV and Standardized Root Mean Square Residual (SRMR) for MLM. To achieve a good fit of the data to the model, the values of CFI and TLI should be over 0.95 and the RMSEA values should be under 0.08 for a reasonable fit and under 0.05 for a good fit[30,31]. For the WRMR values <1.0 have been suggested as indicative of adequate model fit whereas for the SRMR a cut-off value close to 0.08 or below is recommended[31,32]. In order to compare model fit between competing models, the DIFFTEST option of the Mplus can be use when the models to be compared are nested. However, in this case the models cannot be considered as nested, so we declined to use this option.

Third, we gathered sources of validity evidences with other external variables. We analysed the association between the STAI and the Burns-A through Pearson correlations. Also, in order to gather new sources of validity evidence, we carried out an ANOVA between the non-clinical and clinical subsamples, taking into account the subscales and Total Score. Due the disparity of the subsamples, we selected a randomly number of n=45 from the non-clinical subsample. To this purpose we use the SPSS program in order to generate a random sample from the non-clinical sample. Neither age (t=0.557; p=0.421) nor sex rates (χ^2=0.013; p=0.901) differed across subsamples. The statistical analyses were carried out using the programs SPSS 15.0[33] and Mplus 7.0[34].

RESULTS

Descriptive statistics and reliability of the STAI scores

Descriptive statistics of the subscales and the Total Score of the STAI in the non-clinical and the clinical subsamples were calculated (see Table 1). The internal consistency levels calculated by means of Ordinal alpha were adequate in both subsamples. Results for the test-retest ICC displayed a significant relation between the means of the Total Scores in the non-clinical subsample with a coefficient of 0.81 (F=9.3, $p\leq$ 0.001), and in the clinical subsample with a coefficient of 0.93 (F=32.4, $p\leq$0.001). The ICC for the STAI subscales and Total score are shown in Table 2.

Validity evidence based on the internal structure

The goodness-of-fit indices for the STAI models estimated by means of CFAs are presented in Table 3. As it can be seen, the one dimensional model showed a poor fit to the data. In addition, the two-factor, the three-factor, and the bifactor (positive and negative items) models did not reach to 0.95 in CFI value and also showed a RMSEA over 0.08. The bifactor model (anxiety and trait dimensions) showed a CFI below 0.95 (0.94) but the RMSEA was under 0.08. The four factor model showed a reasonable adequacy in RMSEA (0.07), and in addition CFI was 0.95. Moreover, as shown in Table 3, we studied the models fit when data were considered as continuous and MLM estimator was used. In all the cases goodness-of-fit indices were lower compared to the case in which WLSMV estimator was used. Therefore, we decided to continue the analysis with WLSMV estimator.

In order to determine the best model between the bifactor (anxiety and trait dimensions) and the four-factor solution, the DIFFTEST option from Mplus was not possible as

Table 1 Descriptive statistics of the STAI for the non-clinical and the clinical subsamples

	Non-Clinical sample (n=387)					Clinical Sample (n=30)				
	Mean	SD	Skewness	Kurtosis	Ordinal Alpha	Mean	SD	Skewness	Kurtosis	Ordinal Alpha
STAI-State	16.93	10.79	0.97	0.69	0.97	35.03	11.77	-0.23	-0.13	0.93
STAI-Trait	19.03	19.03	0.68	0.17	0.95	35.67	12.90	-0.44	-0.42	0.92
STAI-Total score	35.96	35.96	0.81	0.37	0.98	70.70	20.40	-0.40	-0.22	0.94

Table 2 Test-retest reliability of the STAI scores in the two subsamples

	Non-Clinical sample (n=387)		Clinical Sample (n=30)	
	ICC	F*	ICC	F*
STAI-State	0.73	7.0	0.86	13.0
STAI-Trait	0.84	10.7	0.91	31.6
STAI-Total score	0.81	9.3	0.93	32.4

*All ICC were statistically significant p ≤ 0.01

Table 3 Goodness-of-fit indices of the measurement models tested

Models	χ^2	*df*	CFI	TLI	RMSEA (CI 90%)	WRMR	SRMR
WLSMV (Ordinal Data)							
One-factor	3789.30	740	0.90	0.89	0.10 (0.09-0.10)	2.27	
Two-factor	3369.27	739	0.91	0.90	0.09 (0.09-0.10)	2.06	
Three-factor	2617.22	737	0.94	0.93	0.08 (0.07-0.08)	1.73	
Four-factor	2089.24	733	0.95	0.95	0.07 (0.06-0.07)	1.48	
Bifactor (Anxiety State + Anxiety Trait)	2443.34	699	0.94	0.93	0.07 (0.07-0.08)	1.54	
Bifactor (Positive + Negative Items)	2569.53	700	0.94	0.93	0.08 (0.08-0.09)	1.58	
MLM (Continuous Data)							
One-factor	3423.99	740	0.72	0.70	0.09 (0.09-0.10)	1.90	0.08
Two-factor	2971.79	739	0.76	0.75	0.08 (0.08-0.09)	1.81	0.07
Three-factor	2257.20	737	0.84	0.83	0.07 (0.07-0.08)	1.60	0.06
Four-factor	1953.88	733	0.87	0.86	0.06 (0.06-0.07)	1.39	0.06
Bifactor (Anxiety State + Anxiety Trait)	1982.48	699	0.87	0.84	0.07 (0.06-0.07)	2.15	0.08
Bifactor (Positive + Negative Items)	1944.89	700	0.87	0.85	0.07 (0.06-0.07)	2.31	0.09

χ^2: Chi square; df: degrees of freedom; WLSMV: Weighted Least Square Mean and Variance; MLM: Mean adjusted Maximum Likelihood; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; WRMR: Weighted Root Mean Square Residual; SRMR: Standardized Root Mean Square Residual the models cannot be considered as nested. We then decided to study the factor loadings for each model. In the case of the bifactor model, some factor loadings were not statistically significant. Also, the lower weight of the factor loadings in the case of the bifactor model lead us to concluded that the four-factor model better fit the data. Factor loadings for the four-factor model are shown in Table 4. As it can be seen, all factor loadings were statically significant ($p \leq 0.01$), ranging from 0.50 (item 25 of Trait Negative) to 0.94 (item 18 of State Negative).

Sources of validity evidence: Relations to external variables and mean scores comparison

Table 5 shows the Pearson's correlation between the Anxiety-Trait, Anxiety-State subscales, and Total score of the STAI, and the Burns-A subscales, and Total score. As shown in Table 5, all of the associations between scores were statistically significant. In addition, we compared clinical and non-clinical subsamples through analysis of the variance. Results showed statistically significant differences in the means scores between the subsamples in the STAI-

Table 4 Standardized factor loadings for the STAI four-factor model

Items	Loadings	R^2	Items	Loadings	R^2
State Positive			Trait Positive		
1	0.80	0.65	21	0.87	0.76
2	0.83	0.69	23	0.77	0.36
5	0.80	0.66	26	0.74	0.59
8	0.67	0.68	27	0.69	0.52
10	0.84	0.59	30	0.87	0.25
11	0.83	0.72	32	0.75	0.55
15	0.86	0.45	33	0.80	0.48
16	0.87	0.84	34	0.67	0.57
19	0.84	0.71	36	0.87	0.51
20	0.87	0.69	39	0.78	0.76
State Negative			Trait Negative		
3	0.82	0.78	22	0.60	0.26
4	0.83	0.41	24	0.72	0.56
6	0.85	0.78	25	0.50	0.64
7	0.67	0.74	28	0.76	0.78
9	0.92	0.77	29	0.72	0.44
12	0.88	0.77	31	0.51	0.75
13	0.64	0.89	35	0.89	0.59
14	0.88	0.71	37	0.77	0.47
17	0.88	0.76	38	0.68	0.61
18	0.94	0.76	40	0.78	0.61

All standardized factor loadings were statistically significant ($p \leq 0.01$)

Table 5 Pearson's correlations between STAI and Burns-A

	Anxious Feelings	Anxious Thoughts	Physical Symptoms	Burns-A Total	STAI-State	STAI-Trait
STAI-State	0.63*	0.67*	0.65*	0.71*		
STAI-Trait	0.71*	0.74*	0.69*	0.76*	0.79*	
STAI-Total	0.74*	0.78*	0.74*	0.86*	0.95*	0.94*

* *$p \leq 0.01$*

Table 6 Mean scores comparisons between non-clinical and clinical subsamples

STAI	Non-Clinical M (SD)	Clinical M (SD)	F	p	Parcial η^2
State	16.93 (10.79)	35.03 (11.77)	77.35	≤ 0.001	0.16
Trait	19.03 (10.27)	35.67 (12.90)	20.20	≤ 0.001	0.15
Total Score	35.96 (19.94)	70.70 (20.40)	84.22	≤ 0.001	0.17

η^2: eta square

State (F(417)=77.35; $p \leq 0.001$; partial η^2=0.16), STAI-Trait (F(415)=24.5; $p \leq 0.001$; partial η^2=0.25), and in the Total Score (F(415)=84.22; $p \leq 0.001$; partial η^2=0.17) (see Table 6).

DISCUSSION AND CONCLUSIONS

The main purpose of the study was to analyze the psychometric quality of the Spanish version of the STAI[22]. We thus examined the internal structure of the STAI, estimated the reliability of the scores, and gathered different sources of validity evidence. This goal provided new information on the psychometric properties of the STAI scores in order to use it as a screening instrument in Spanish-speaking adult populations. Results found in the present study indicate that the STAI has adequate psychometric properties and, therefore, it is a useful instrument that could be used to evaluate trait and state anxiety symptoms.

In order to study the internal consistency we decided to compute the Ordinal alpha. Previous studies[14-20] have used the Cronbach's alpha, obtaining, in some cases, lower values than in the present work. In this sense, the fact that Ordinal alpha was used might be a relevant variable that explains these differences. Ordinal alpha, has been shown to estimate reliability more accurately than Cronbach's alpha for ordinal response scales[28]. In this sense, this is an added value, as the Ordinal alpha is preferred for categorical data as is the case in the STAI. Results showed good values in all the subscales both in the non-clinical and in the clinical subsamples. Evidence obtained from the test-retest study also ratified the good stability of the scores, confirming the reliability of the STAI for its use both with non-clinical and clinical population.

Analysis of the internal structure underlying the STAI scores allowed to support a four-factor model and a bifactor model (anxiety state, anxiety trait, and a general factor). The one-factor model and the two-factor model showed the poorer goodness-of-fit indices (e.g., RMSEA=0.10, and 0.09, respectively). The three-factor and the other bifactor model (positive items, negative items, and general factor) showed acceptable CFI indices but RMSEA was over 0.08. As the four-factor model and the bifactor model (anxiety state, anxiety trait, and a general factor comprising both) displayed the better fit, we decided to analyse the factor loadings. The study of the factor loadings showed that the bifactor models displayed some factor loadings no significant. Specifically four items in the Anxiety dimension and two in the Trait were no significant. The bifactor model in which positive and negative items were considered, showed the same problems with factor loadings. For this reason we concluded that the four-factor model was the most adequate model. Nevertheless, the bifactor model showed adequate goodness-of-fit indices and should be considered for future research as it has a valuable psychological meaning. In addition, the four factor model has been critized because is based on a statistical artifact[19]. The results found are in line with other studies indicating that a four-factor model was the most appropriate[17,19,22,23]. However, goodness-of-fit indices in some of these studies did not reach acceptable values, questioning the adequacy of the model[19].

It is worth noting that differences with our study that may account for the discrepant findings pertaining to the fit of these models could be that we treated our Likert-scale data as categorical, whereas the other studies treated their Likert-scale data as continuous. Technically, a Likert-scale data are categorical (ordinal) in nature and, as consequence, should be treated as such (e.g., base matrices on polychoric correlations, use appropriate estimators) given that treating Likert-scale data as continuous may lead to bias in parameter estimation[30]. With regard to this issue, given that we treated our data as categorical, it is not exactly clear to affirm that we obtained better fit in the four-factor model than previous studies. In this sense, more studies with data treated as ordinal are advisable in order to a better comparison of hypothetical model explaining the STAI dimensional structure.

Results from the analysis of the sources of validity evidence with external variables yielded a significant association between the STAI and the BURNS-A scores[25] in all the subscales and in the Total Score. These results support the validity of the STAI with other external variables and are consistent with previous studies that showed evidence for the validity of the STAI with other external sources[14-19]. Anxiety traits and states have been related, in different studies with psychological problems such as depression[35]. In this sense, screening of anxiety traits in each of the dimensions of the STAI could be relevant for the early detection in order to avoid psychological disorders that may become permanent. In addition, results of the present study related to the sources of discriminant validity, permit to asure, according to previous studies[36], the goodness of the STAI to distinguish between clinical and non-clinical samples.

The results of the present study should be interpreted in the light of the following limitations. First of all, there is an inherent issue in the administration of every type of self-reported instrument, with the very well-known effect of stigmatization, the possibility of misunderstanding of some items or the lack of introspection of some participants, and the social desirability. For these reasons, it would have been relevant to use external sources of information via hetero-informs or structured interviews. Second, we used the X version of the instrument. In this sense, the more recent Y version is considered to be most advanced[37], and studies about this version are also needed. Third, we did not make any statistical analysis to study the response distribution of the items attending different relevant variables like gender or age. Finally, the clinical sample was smaller than the non-clinical sample and as a consequence comparisons might benefit from a larger clinical sample in order to establish statistically sound results. Future research should keep on studying the internal structure of the STAI with new approaches and techniques such as Exploratory Structural Equation Modelling (ESEM). In addition, the study of the measurement invariance of the STAI across the culture would contribute to better understand the STAI structure.

Despite the noted limitation and areas that would benefit of future research, the present study identified the fruitfulness of the STAI in order to study anxiety in non-clinical and clinical populations. It also has an added value as is the first to the best of our knowledge that compares this model with new approximations such as the bifactor approach. In this sense, the study contributes with valuable information with regard to the fact that the study of the internal structure should be conducted considering the ordinal nature of the data. Furthermore, results found in the present study have clinical implications. The comprehension of the underlying structure of anxiety manifestations might benefit the clinical evaluation and intervention for professionals. The results reveal that the STAI is a useful tool that allows identifying clinical and non-clinical populations.

ACKNOWLEDGEMENTS

This research was founded by the Spanish Ministry of Science and Innovation (MICINN) (reference PSI2014-56114-P), the Institute Carlos III, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), and by the 2015 BBVA grant for Cultural Creators and Researchers.

REFERENCES

1. Spielberger CD, Gorsuch R, Lushene R. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologist; 1970.
2. Copeland WE, Shanahan L, Costello EJ, Angold A. Childhood and Adolescent Psychiatric Disorders as Predictors of Young Adult Disorders. Arch Gen Psychiatry. 2009;66(7):764-72.
3. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? J Child Psychol Psychiatry. 2011;52(10):1015-25.
4. Costello EJ, Egger H, Angold A. 10-Year Research Update Review: The Epidemiology of Child and Adolescent Psychiatric Disorders: I. Methods and Public Health Burden. J Am Acad Child Adolesc Psychiatry. 2005;44(10):972-86.
5. Kessler RC, Berglund P, Demler O, Jin R, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62:593-602.
6. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand. Suppl. 2004;420:21-7.
7. Kessler RC, Avenevoli S, Costello EJ, Georgiades K, Green JG, Gruber MJ, et al. Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. Arch Gen Psychiatry. 2012;69(4):372-80.
8. Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence and incidence studies of anxiety disorders: Review of the literature. Can J Psychiatry. 2006;51:100-13.
9. Bloom BL. Brief interventions for anxiety disorders: Clinical outcome studies. Brief Treat Crisis Interv. 2002;2:325-39.
10. Beutel ME, Bleichner F, von Heymann F, Tritt K, Hardt J. Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. Int J Clin Health Psychol. 2011;11:443-57.
11. Muñiz J, Fernández-Hermida JR. La opinión de los psicólogos españoles sobre el uso de los tests. Papeles del Psicólogo. 2010;31:108-21.
12. Bermúdez J. Anxiety and performance. Revista de Psicología General y Aplicada. 1978;151:183-207.
13. Bermúdez J. Functional analysis of anxiety. Revista de Psicología General y Aplicada. 1978;153:617-34.
14. Bados A, Gómez-Benito J, Balaguera G. The State-Trait Anxiety Inventory, Trait Version: Does it really measures anxiety? J Pers Assess. 2010;92:560-7.
15. Mystakidou K, Tsilika E, Parpa E, Sakkas P, Vlahos L. The psychometric properties of the Greek version of the State-Trait Anxiety Inventory in cancer patients receiving palliative care. Psychol Health. 2009;24:1215-28.
16. Suzuki T, Tsukamoto K, Abe K. Characteristics factor structures of the Japanese version of the State-Trait Anxiety Inventory: coexistence of positive-negative and state-trait factor structures. J Pers Assess. 2000;74:447-58.
17. Guillén-Riquelme A, Buela-Casal G. Actualización psicométrica y funcionamiento diferencial del ítem en el State Trait Anxiety Inventory (STAI). Psicothema. 2011;23:510-5.
18. Vautier S, Pohl S. Do balanced scales assess bipolar constructs? The case of the STAI scales. Psychol Assessment. 2009;21:187-93.
19. Vigneau F, Cormier S. The factor structure of the State-Trait Anxiety Inventory: an alternative view. J Pers Assess. 2008;90:280-5.
20. Virella B, Arbona C, Novy DM. Psychometric properties and factor structure of the Spanish version of the State-Trait Anxiety Inventory. J Pers Assess. 1994;63:401-12.
21. Fonseca-Pedrero E, Paino Piñeiro M, Sierra Baigrie S, Lemos Giráldez S, Muñiz Fernández J. Propiedades psicométricas "Cuestionario de Ansiedad Estado/Rasgo" (STAI) en universitarios [Psychometric properties "Anxiety State/Trait Questionnaire" in college students]. Psicol Conductual. 2012;20(3):547-62.
22. Spielberger CD, Gorsuch RL, Lushene RE. STAI. Cuestionario de ansiedad estado-rasgo (7ª ed. rev.). Madrid: TEA; 2008.
23. Hishinuma ES, Miyamoto RH, Nishimura ST, Nahulu LB. Differences in State-Trait Anxiety Inventory scores for ethnically diverse adolescents in Hawaii. Cultur Divers Ethnic Minor Psychol. 2000;6:73-83.
24. Reise SP, Moore TM, Haviland MG. Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. J Pers Assess. 2010;92:544-59.
25. Burns DD. Ten Days to Self-Esteem. New York: Quill; 1993.
26. American Psychiatric Association. Diagnostic and statistical manual of mental disorders 4ª ed. rev. ed. Washington, DC: American Psychiatric Association Publishing; 2000.
27. Ortuño-Sierra J, Rodríguez L, Debbané M, Fonseca-Pedrero E. Anxiety Assessment: Psychometric properties of the Spanish version of the Burns Anxiety Inventory. Span J Psychol. 2015 Jun 22;18:E44.
28. Zumbo BM, Gadermann AM, Zeisser C. Ordinal versions of coefficients alpha and theta for Likert rating scales. J Mod Appl Stat Meth. 2007;6:21-9.
29. Muthen LK, Muthen BO. Mplus User's Guide. (5th ed.). Los Angeles, CA: Muthen & Muthen; 1998-2007.
30. Brown TA. Confirmatory Factor Analysis for Applied Research. New York: Guilford Press; 2006.
31. Hu LT, Bentler PM. Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6(1):1-55.
32. Yu CY. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Los Angeles: University of California, Doctoral dissertation; 2002.
33. Statistical Package for the Social Sciences. SPSS Base 15.0 User's Guide. Chicago, IL: SPSS Inc; 2006.
34. Muthén LK, Muthén BO. Mplus User's Guide. Seventh Edition. Los Angeles: Muthén & Muthén; 1998-2012.
35. López-Ibor JJ. Ansiedad y depresión, reacciones emocionales frente a la enfermedad [Evolution of anxiety and depression detected during hospitalization in an Internal Medicine service]. An med interna. 2007;24(5):209-11.
36. Kokanalia MK, Cavkaytara S, Guzela AI, Topçua HO, Eroğlub E, Aksakala O, et al. Impact of preprocedural anxiety levels on pain perception in patients undergoing office hysteroscopy. J Chin Med Assoc. 2014;77(9):477-81.

37. González-Fraile E, Domínguez-Panchón A, Fernández-Catalina P, Gonçalves-Pereira M. Las diferentes versiones del inventario de ansiedad estado-rasgo. Rev Psiquiatr Salud Ment. 2014;7(3):151-3.