

UNIVERSIDAD DE GRANADA

FACULTAD DE PSICOLOGÍA

Departamento de Personalidad, Evaluación y Tratamiento Psicológico



Programa de doctorado:

Diseños de Investigación y Aplicaciones en Psicología y Salud (P33.56.1)

TESIS DOCTORAL

**Validación de la adaptación española del State-Trait Anxiety
Inventory en diferentes muestras españolas.**

Doctorando: Alejandro Guillén Riquelme

Director: Dr. Gualberto Buela Casal

Editor: Editorial de la Universidad de Granada
Autor: Alejandro Guillén Riquelme
D.L.: GR 1852-2014
ISBN: 978-84-9083-035-2

El doctorando Alejandro Guillén Riquelme y el director de la tesis “Validación de la adaptación española del State-Trait Anxiety Inventory en diferentes muestras españolas”. Garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección del director de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Y para que conste, se expide en Granada el presente a día 17 de enero de 2014



Fdo. El director de la tesis

Dr. Gualberto Buela Casal



Fdo. El doctorando

D. Alejandro Guillén Riquelme

Esta Tesis Doctoral ha sido realizada según la normativa reguladora de los Estudios de Tercer Ciclo y del Título de Doctor de la Universidad de Granada aprobada por el **Consejo de Gobierno a 26 de Septiembre de 2005 (artículo nº 27)** referida a la modalidad de *Tesis Doctoral compuesta por el reagrupamiento de trabajos de investigación publicados por el doctorando.*

-A mis abuelos Josefina y Manolo-

AGRADECIMIENTOS

Sin las siguientes personas esta tesis no hubiese sido, por ello quisiera expresar mi gratitud a todos ellos:

Al Dr. Gualberto Buela Casal, por su dirección, correcciones y consejos, y, sobretodo, por haber confiado en mí y haberme dado la oportunidad de investigar, descubriéndome así un mundo fascinante.

A la Dra. María de la Paz Bermúdez, por todo lo que me enseñó, su orientación y consejos.

A todos los clínicos que me han ayudado con la recogida de pacientes, por su implicación que ha hecho posible contar con las dos muestras clínicas: Carmen Pitti, Lilisbeth Perestelo, África Urbano, Anna Aznar, Izaskun Basterra, Isabel Pinillos, Lourdes Escribano, Francisco José Maestre, María Jesús Álava, Mar Torres, Esteban Vallejo, Beatriz Corbí, Esperanza Medina, María José Moraga, María Teresa González y Tíscar Rodríguez.

Quisiera agradecer también a todas las personas que cumplimentaron los cuadernillos formando parte de esta tesis.

Índice

Resumen	17
Abstract.....	25
Introducción	33
Objetivo 1	
<i>Comprobación de la fiabilidad y validez del State-Trait Anxiety</i>	
<i>Inventory</i>	55
 Estudio 1: Diferencias entre pacientes con ansiedad y población general mediante STAI: Meta-análisis de comparación de grupos y generalización de la fiabilidad.....	57
 Estudio 2: Actualización psicométrica y funcionamiento diferencial de los ítems en el <i>State-Trait Anxiety Inventory</i> (STAI).....	85
Objetivo 2	
<i>Creación y ajuste de una versión breve del State-Trait Anxiety</i>	
<i>Inventory en muestra española</i>	109

Estudio 3: Short form of the Spanish adaptation of the State-Trait Anxiety Inventory.....	111
Estudio 4: Versión breve del STAI en adolescentes y universitarios españoles	133
Objetivo 3	
<i>Fiabilidad y validez del State-Trait Anxiety Inventory en muestras clínicas.....</i>	155
Estudio 5: Dimensionalidad del <i>State Trait Anxiety Inventory</i> en pacientes diagnosticados con depresión	157
Estudio 6: Fiabilidad y validez del <i>State-Trait Anxiety Inventory</i> en pacientes diagnosticados con depresión	179
Discusión	203
Conclusiones.....	215
Conclusions	217
Anexos.....	227

RESUMEN

Resumen

En la presente tesis doctoral se ha fijado el objetivo de comprobar la fiabilidad y validez del STAI en varias muestras, así como establecer una versión breve de este instrumento. Para ello se calculó la consistencia interna y la factorización así como el funcionamiento diferencial de los ítems en una muestra de adultos españoles. Tras la comprobación de la fiabilidad se pasó a extraer un conjunto de ítems que permitiesen evaluar la ansiedad estado y rasgo. Tras ello, se comparó el ajuste de dicha selección de ítems con otras versiones breves anteriormente validadas. Tanto la versión obtenida como las versiones breves propuestas por otros autores fueron comprobadas en una muestra de adolescentes y otra de universitarios. Por último, se emplearon dos muestras de pacientes: la primera con pacientes con alguno de los diagnósticos de depresión recogidos en el DSM-IV-TR y la segunda con pacientes con alguno de los diagnósticos de ansiedad recogidos en dicho manual. En estas muestras se comprobó la fiabilidad y validez. Para lograr la consecución de los objetivos propuestos y realizar estos pasos descritos se ha realizado seis estudios, agrupados en tres objetivos generales. A continuación se resume cada uno de los estudios de forma independiente:

Objetivo 1: Comprobación de la fiabilidad y validez del State-Trait Anxiety Inventory

Estudio 1: Diferencias entre pacientes con ansiedad y población general mediante STAI: Meta-análisis de comparación de grupos y generalización de la fiabilidad

Fundamentos: Desde su creación el STAI se ha citado en más de 14.000 documentos, contando con más de 60 adaptaciones en diversos países. En algunas de ellas este instrumento no cuenta con puntuaciones clínicas. El objetivo de este trabajo es determinar si el STAI tiene puntuaciones superiores en pacientes diagnosticados con

ansiedad respecto a población general. Además, se pretende analizar si la consistencia interna es adecuada en personas con ansiedad.

Método: Se realizó una búsqueda bibliográfica en *Tripsdatabase*, *Cochrane*, *Web of Knowledge*, *Scopus*, *PyscINFO* and *Scholar Google*, para documentos publicados entre 2008 y 2012. Se seleccionaron 131 artículos para la comparación entre pacientes diagnosticados con ansiedad respecto a población general y 25 para la generalización de la fiabilidad. Para los análisis se utilizó la *d* de Cohen para la comparación de medias (método de efectos aleatorios) y el alfa de Cronbach para la generalización de la fiabilidad (método de efectos fijos).

Resultados: En la comparación entre grupos en la ansiedad estado las diferencias fueron estadísticamente significativas ($d = 1,39$; IC 95%: 1,22-1,56). En la ansiedad rasgo fue $d = 1,74$ (IC 95%: 1,56-1,91). Además de ello, la fiabilidad para pacientes con algún trastorno de ansiedad abarcó de 0,87 a 0,93.

Conclusiones: el STAI es sensible al nivel de ansiedad de la persona y fiable para muestras de pacientes diagnosticados con crisis de angustia, fobia específica, fobia social, fobia social generalizada, trastorno de ansiedad generalizada, trastorno de estrés post-traumáticos, trastorno obsesivo compulsivo o trastorno por estrés agudo.

Estudio 2: Actualización psicométrica y funcionamiento diferencial de los ítems en el State-Trait Anxiety Inventory (STAI)

La ansiedad es uno de los problemas psicológicos con mayor prevalencia. Entre los instrumentos para medirla se encuentra el State Trait Anxiety Inventory. Este cuestionario mide ansiedad rasgo (factor de personalidad que predispone a sufrir o no ansiedad) y ansiedad estado (los factores ambientales que protegen o generan ansiedad). La adaptación española del cuestionario se realizó en 1982, por ello el objetivo del estudio es realizar una revisión del STAI. Para ello, se reunió una muestra de 1.036

adultos. Se realizó un análisis de fiabilidad mediante alfa de Cronbach (0,90 para ansiedad rasgo y 0,94 para ansiedad estado). También se realizó una reducción factorial con unos resultados similares a los de la adaptación original. Además, se comprobó si existía funcionamiento diferencial de ítem por sexo y únicamente una de las 40 preguntas mostró problemas. Por último se realizó comparaciones t-Student con los valores de la adaptación; mientras el rasgo varía en 1 punto, el estado tiene diferencias de hasta 6 puntos. En estos resultados de forma general se observa que el STAI mantiene unas adecuadas propiedades métricas y que, además, ha sido sensible al aumento de estímulos ambientales que producen estrés.

Objetivo 2: Creación y ajuste de una versión breve del State-Trait Anxiety Inventory en muestra española

Estudio 3: Versión breve de la adaptación española del State-Trait Anxiety Inventory

El State-Trait Anxiety Inventory (STAI) es uno de los instrumentos de evaluación más empleados por psicólogos. Desde su creación, se han llevado a cabo varias versiones breves del mismo; pese a ello, no existe ninguna para población general, en muestra española. El objetivo del presente estudio es establecer y validar una versión breve. Para ello, se aplicó el STAI completo a 1.157 adultos. En función de las puntuaciones en el STAI se realizaron grupos de alta y baja ansiedad. Mediante dichos grupos, se realizó un análisis discriminante. Para la ansiedad estado y rasgo, respectivamente, se obtuvieron reducciones de cinco ítems. Seguidamente, se comparó esta reducción de ítems con otras versiones breves, mediante análisis factorial confirmatorio. La versión establecida mediante análisis discriminante ajusta adecuadamente. No obstante, hay una selección de 12 ítems (seis por subescala) en la que se observa un mejor ajuste. Así pues, los ítems 2, 4, 11, 15, 17 y 18 (para ansiedad

estado) y los ítems 7, 14, 15, 16, 17 y 18 (para ansiedad rasgo) son los ítems más adecuados para establecer una versión breve en muestra española. En conclusión, se puede emplear estos 12 ítems para favorecer evaluaciones muy largas o de personas con problemas de salud que no puedan realizar evaluaciones muy largas.

Estudio 4: Versión breve del STAI en adolescentes y universitarios españoles

El STAI es uno de los instrumentos de evaluación psicológica más empleados, generándose diversas versiones breves del mismo. No se han hallado versiones breves en las que se haya comprobado el ajuste para muestras de adolescentes ni universitarios españoles. El objetivo de este estudio es comprobar cuál de las versiones cortas propuestas en la literatura tiene un mejor ajuste en adolescentes y universitarios. Para ello, se aplicó el STAI a 482 adolescentes escolarizados y 510 estudiantes universitarios, de diversas ciudades españolas. Mediante AFC se evaluó el ajuste de cinco versiones breves del STAI. En los universitarios hay dos modelos en los que se observa un buen ajuste. En los adolescentes no ajusta ningún modelo. Puede que el proceso de diferenciación emocional con la edad explique que en los adolescentes las versiones breves halladas (siempre con adultos) no serían válidas, sí pudiendo emplearse con universitarios.

Objetivo 3: Fiabilidad y validez del State-Trait Anxiety Inventory en muestras clínicas

Estudio 5: Estructura factorial del State-Trait Anxiety Inventory en pacientes diagnosticados con depresión

Introducción: Pese a ser uno de los cuestionarios para evaluar ansiedad más empleados por profesionales de la salud mental, el STAI ha sido objeto de críticas, entre las que destaca la posible existencia de un conjunto de ítems que por evaluar depresión

conformarían un factor independiente. El objetivo de este trabajo es evaluar la factorización del STAI en una muestra de pacientes con diagnóstico de depresión.

Material y métodos: Para ello se aplicó la adaptación española del STAI a 266 pacientes españoles, diagnosticados con diferentes trastornos depresivos.

Resultados: Mediante un análisis factorial exploratorio se determinaron tres factores subyacentes: ansiedad estado, ansiedad rasgo positiva y ansiedad rasgo negativa.

Conclusiones: Así pues, la factorización realizada no permite confirmar la presencia de conjuntos de ítems específicos para la depresión, señalada anteriormente como la principal crítica a este cuestionario. Además, los elevados valores del alfa categórico, tanto en la estructura factorial obtenida, como en las subescalas teóricas, son indicios de una elevada fiabilidad para el empleo del STAI en pacientes con diagnóstico de depresión.

Estudio 6: Fiabilidad y validez del State-Trait Anxiety Inventory en pacientes diagnosticados con ansiedad

Resumen: Introducción: La adaptación española del STAI ha demostrado ser fiable y válida en población general española, no obstante no ha sido validada para pacientes con diagnóstico de ansiedad. El objetivo es analizar los ítems la fiabilidad y la validez de constructos en pacientes con diagnóstico de ansiedad. Método: Para lograr el objetivo se aplicó el STAI, así como varios cuestionarios de ansiedad y depresión a una muestra de 251 pacientes con diagnóstico principal de ansiedad. Resultados: La media de la ansiedad rasgo y estado son elevadas, siendo significativamente diferentes del valor de la adaptación española para población normal. La fiabilidad es superior a 0,9 en ambas escalas. Hay relaciones moderadas entre el STAI y el BAI, así como con el IDER y el CBD. Discusión: El STAI es fiable en pacientes con ansiedad, no obstante las

correlaciones con instrumentos de depresión hace pensar que es posible que evalúe depresión. La correlación con el BAI es elevada por lo que en pacientes con ansiedad la evaluación de los niveles de ansiedad son válidas.

Abstract

In this thesis the objective is to check the reliability and validity of the STAI in several samples, and to establish a short version of this instrument. The internal consistency, factorization and differential item functioning in a sample of Spanish adults was calculated. After to check the reliability were extracted a set of items that would allow to assess state and trait anxiety. Thereafter, the selection of items was compared with other previously validated short versions. Both, the version obtained as brief versions proposed by other authors, were tested in a sample of adolescents and other of high school students. Finally, two patient samples were used: patients with depression diagnoses in the DSM -IV- TR and patients with any of the anxiety diagnoses listed in the manual. In these samples the reliability and validity were verified. To ensure the attainment of the objectives and follow these steps has been six studies, grouped in three objectives. The following summarizes each of the studies separately:

Objective 1: Testing the reliability and validity of the State-Trait Anxiety Inventory

Study 1: Differences between patients with anxiety and general population using the STAI: Meta-analysis of group comparison and reliability generalization.

Background: Since its creation the STAI has been cited in more than 14,000 documents, with more than 60 adaptations in different countries. In some adaptations this instrument has no clinical scores. The aim of this work is to determine if the STAI has higher scores in patients diagnosed with anxiety than in general population. In addition, we want to examine if the internal consistency is adequate in anxious patient samples.

Methods: We performed a literature search in *Tripdatabase*, *Cochrane*, *Web of Knowledge*, *Scopus*, *PyscINFO* and *Scholar Google*, for documents published between 2008 y 2012. We selected 131 scientific articles to compare between patients diagnosed with anxiety and general population, and 25 for the generalization of reliability.

Results: In the groups comparison the differences in state anxiety was significant ($d = 1,39$; CI 95%: 1,22-1,56). In the trait anxiety was $d = 1,74$ (CI 95%: 1,56-1,91). 0.87 and 0.93. For the analysis we used Cohen's d for means comparisons (random-effects method) and Cronbach's alpha for the reliability generalization (fixed-effects method).

Discussion: So it seems that the STAI is sensitive to the level of anxiety of the individual and reliable for patients with diagnosis of panic attack, specific phobia, social phobia, generalized social phobia, generalized anxiety disorder, post-traumatic stress disorder, obsessive compulsive disorder o acute Stress disorder.

Study 2: Psychometric revision and differential item functioning in the State Trait Anxiety Inventory (STAI).

One of the psychological problems with highest prevalence is anxiety. The State Trait Anxiety Inventory is one of the instruments to measure it. This questionnaire assesses Trait Anxiety (understood as a personality factor that predisposes one to suffer from anxiety) and State Anxiety (refers to environment factors that protect from or generate anxiety). The questionnaire was adapted in Spain in 1982. Therefore, the goal of the study is to review the current psychometric properties of the STAI. A total of 1,036 adults took part in the study. Cronbach's alpha reliability was .90 for Trait and .94 for State Anxiety. Factor analysis showed similar results compared with the original data. Moreover, differential item functioning (DIF) was carried out to explore sex bias.

Only one of the 40 items showed DIF problems. Lastly, a t-Test was run, comparing the original and current values; whereas Trait Anxiety varied in 1 point, State Anxiety had differences of up to 6 points. In general, this result shows that the STAI has maintained adequate psychometric properties and has also been sensitive to increased environmental stimuli that produce stress.

Objective 2: Creating and adjusting a short version of the State-Trait Anxiety Inventory in Spanish shows

Study 3: Short form of the Spanish adaptation of the State-Trait Anxiety Inventory

The State-Trait Anxiety Inventory (STAI) is one of the assessment instruments that are most widely used by psychologists. Although several short forms of the STAI have been developed since its creation, none are available for the Spanish general population. The aim of the present study was to develop and validate a short form of the STAI. To achieve this, we applied the full STAI to 1,157 adults. We created high and low anxiety groups based on participants' scores on the STAI and conducted a discriminant analysis using such groups. We obtained a selection of five items for state anxiety and five items for trait anxiety and compared it to other short forms through a confirmatory factor analysis. The short form obtained with the discriminant analysis showed good fit. However, a selection of 12 items (six for each subscale) showed better fit. More specifically, items 2, 4, 11, 15, 17 and 18 (for state anxiety) and items 7, 14, 15, 16, 17 and 18 (for trait anxiety) were found to be the most adequate items to develop a short form for Spanish samples. In conclusion, these 12 items can be used to facilitate the application of very long assessments or to assess individuals with health problems who cannot perform very long assessments.

Study 4: Short version of STAI in adolescents and college Spanish

The STAI is an instrument very used, generating different short versions. We found no short versions in which it is tests the fit for teenagers and college samples Spanish. The aim of this study is to test which of the short versions proposed in the literature have a better fit in adolescents and college students. To do this, we applied the STAI to 482 adolescent students and 510 university students from different Spanish cities. Through AFC we evaluated five short versions of the STAI. At the university there are two models in which there is a good fit. In adolescents not fit any model. Maybe the emotional differentiation process explains that the found short versions (always with adults) are not valid in teenagers, and can be used with university students.

Objective 3: Reliability and validity of the State-Trait Anxiety Inventory in clinical samples

Study 5: State Trait Anxiety Inventory's factorial structure for patients diagnosed with depression.

Introduction: The STAI, despite to be one of the questionnaires to assess anxiety more employees for mental health professionals, has been criticized, the most notably critic is the possible existence of a set of items to assess depression would form an independent factor. The aim of this work is to evaluate the STAI factoring in a sample of patients diagnosed with depression.

Material and methods: We applied the Spanish adaptation of the STAI to 266 Spanish patients diagnosed with various depressive disorders.

Results: In the exploratory factor analysis were identified three underlying factors: state anxiety, trait anxiety positive and negative trait anxiety.

Conclusions: Thus, factorization does not confirm the presence of specific item sets for depression, noted above as the main criticism of this questionnaire. Furthermore, the high values of alpha categorical, in the factor structure obtained, as in the theoretical subscales are highly reliable indications for the use of the STAI in patients diagnosed with depression.

Study 6: Reliability and validity of the State-Trait Anxiety Inventory for patients diagnosed with anxiety.

Introduction: The Spanish adaptation of the STAI has proven reliable and valid in Spanish general population, however has not been validated for patients diagnosed with anxiety. The objective is to analyze the items the reliability and validity of constructs in patients diagnosed with anxiety. Method: To achieve the objective the STAI and several questionnaires of anxiety and depression were applied in a sample of 251 patients with a diagnosis of anxiety. Results: Means to state and trait anxiety are high, being significantly different from the value of the Spanish adaptation to normal population. Reliability is above 0.9 on both scales. There are moderate relationships between the STAI and BAI, IDER and CBD. Discussion: In anxiety patients sample, the correlation between STAI and BAI is an indication that the validity is adequate, however the correlations with depression instruments imply that may be evaluating different constructs. In addition, the STAI is reliable in patients with anxiety.

INTRODUCCIÓN

Introducción

La ansiedad es uno de los trastornos psicológicos que tiene una mayor prevalencia entre la población general (Beutel, Bleichner, von Heymann, Tritt y Hardt, 2011). De la misma forma en el grupo de universitarios es uno de los trastornos psicológicos más frecuentes (Micin y Bagladi, 2011). En muestras de Estados Unidos las tasas de prevalencia para “fobia específica” son superiores al 15% (Kessler, Petukhova, Sampson, Zaslavsky y Wittchen, 2012; Kessler, Ruscio, Shear y Wittchen, 2009). En el caso de Europa las tasas de prevalencia, pese a ser menores, siguen siendo muy elevadas. Por ejemplo, Alonso et al. (2004b) observan tasa de prevalencia en seis países Europeos de un 13,6% y Andlin-Sobocki y Wittchen (2005) un 12%. De hecho, más del 26% de las consultas psiquiátricas son para tratar trastornos de ansiedad (Alonso et al., 2004c). En el caso de España desciende el porcentaje hasta un 6,8% en el último año desde el momento de la consulta (Wang et al., 2007). Pese a las diferencias entre países, la tasa de diagnósticos de ansiedad es alta de forma general (Somers, Goldner, Waraich y Hsu, 2006; Wang et al., 2007), a lo que hay que añadir para un 28% de los pacientes la presencia de otros trastornos (Alonso et al., 2004a).

Para la evaluación de los niveles de ansiedad generales existen multitud de instrumentos. Entre ellos, uno de los más empleados es el *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch y Lushene, 1970). Desde su creación el STAI ha sido citado en más de 14.000 documentos de archivo y adaptado a más de 60 culturas (Spielberger y Reheiser, 2009).

Ansiedad estado y ansiedad rasgo

El STAI evalúa la ansiedad desde la teoría del estado y el rasgo, lo que supone la existencia de dos componentes independientes, que conforman el constructo. El primero de ellos sería la ansiedad rasgo entendida como un factor de personalidad; sería la

tendencia, relativamente estable, de la persona para evaluar los estímulos como más o menos ansiosos, provocando con ello un mayor aumento de la ansiedad. Por el contrario, la ansiedad estado sería la cantidad de estímulos ansiosos de la persona en el momento de la evaluación (Spielberger et al., 1970). La ansiedad estado fluctúa notablemente y, en teoría, es independiente de la ansiedad rasgo, no así en la práctica, pues por lo general el sujeto con altos niveles de ansiedad rasgo que percibe en una situación concreta los estímulos como más amenazantes, tenderá a puntuar más en ansiedad estado, con lo cual la pretendida independencia de los factores queda comprometida. De hecho, para aplicar el STAI se debe seguir el orden de ansiedad estado y rasgo, ya que de no hacerlo en este sentido, las puntuaciones del evaluado en la ansiedad rasgo influirían sobre la ansiedad estado, provocando que esta última se equipare con la primera (Buela-Casal, Guillén-Riquelme y Seisdedos Cubero, 2011).

Primera versión del STAI: forma X

Las primeras versiones del STAI, anteriores a 1970 compartían los mismos ítems para evaluar ambas escalas con la sola variación del inicio de los enunciados: para ansiedad rasgo la persona puntuaba cómo se sentía “en general” y en ansiedad estado “en los días previos” a la evaluación para cada uno de los ítems. Esta forma se descartó, y se estableció la primera versión comercial del STAI en 1970 (Spielberger et al., 1970). Esta versión está compuesta por 40 ítems (20 para cada una de las subescalas). En ambas escalas existen ítems que evalúan presencia de ansiedad, mientras que otros miden ausencia de ansiedad. Sin embargo, el número de ítems negativos es diferente para cada una de las subescalas: en ansiedad estado la mitad de los ítems son positivos y la otra mitad están invertidos; en el caso de la ansiedad rasgo 13 de los ítems miden presencia de ansiedad, mientras que los otros 7 evalúan ausencia de ansiedad. Aunque la disparidad observada al comparar la proporción ítems positivos/negativos en ambas escalas pueda

considerarse poco transcendental hay que señalarla como el germén de una de la principales críticas al STAI: el número de factores observados.

Esta primera versión está diseñada para la evaluación de la ansiedad, tanto en adultos como en adolescentes a partir de los 14 años (Spielberger et al., 1970). Posteriormente, se desarrollaría una versión para niños desde los seis años (Spielberger, 1973). En el STAI para adultos se observa una adecuada fiabilidad, superior a 0,9 en el estudio original y de 0,9 para ambas subescalas, a partir de los datos de un meta-análisis (Barnes, Harp y Jung, 2002). En el mismo trabajo se analiza la media de las correlaciones test-retest, obteniendo una correlación media en la subescala de estado de 0,7 y en la subescala de ansiedad rasgo de 0,88.

A raíz del diferente número de ítems invertidos en cada una de las subescalas, se obtienen factorizaciones de tres y cuatro factores en lugar de los dos teóricos. La estructura de tres factores estaría conformada por ansiedad estado negativa y positiva y ansiedad rasgo (Hishinuma, Miyamoto, Nishimura y Nahulu, 2000). Otros trabajos incorporan la ansiedad rasgo negativa a estos tres (Vera-Villarroel, Celis-Atenas, Córdova-Rubio, Buela-Casal y Spielberger, 2007), siendo esta factorización la obtenida en la adaptación española (Seisdedos Cubero, 1982), en la que se incidirá con mayor detalle.

Críticas

Conforme se señaló en párrafos anteriores la primera versión comercial del STAI no se vio libre de críticas junto a revisiones favorables y un notable apoyo empírico (Katkin, 1975). La primera de ellas, ya indicada, se produce por el diferente número de ítems invertidos en cada una de las escalas. Así pues diversos autores señalan que la factorización del STAI es inestable, habiendo confusión en la pertenencia de alguno de los ítems. Por ejemplo, Vautier (2004) señala que diversos ítems negativos del STAI

evalúan no sólo ausencia de ansiedad, también presencia de la misma, provocando los problemas en las factorizaciones de tres y cuatro subescalas.

La segunda crítica es que la presencia de palabras complejas contenidas en algunos ítems producen numerosas omisiones y respuestas anómalas, que provoca un aumento en el sesgo, y se hace más evidente en las evaluaciones en adolescentes, colectivos con bajo nivel socio-educativo y en personas con baja inteligencia. De esta forma la evaluación en estos colectivos estaba sesgada (Spielberger et al., 1983).

Entre las críticas que el STAI ha recibido, es que puede estar evaluando depresión o constructos similares. Esta crítica se fundamenta en diversos trabajos con metodologías muy diferentes. En primer lugar, estudios en los que se observan correlaciones significativas del STAI con medidas de depresión, malestar general o afecto negativo. Un ejemplo es el trabajo de Grös, Antony, Simms y McCabe (2007) donde se realizan correlaciones entre el STAI y varios cuestionarios de depresión, obteniendo valores entre 0,45 y 0,65. En segundo lugar se analizan las puntuaciones en el STAI para pacientes con diagnóstico de depresión, obteniendo valores muy elevados (Kennedy, Schwab, Morris y Beldia, 2001; Spielberger y Reheiser, 2009).

Para comprobar la posible relación entre el contenido de los ítems del STAI y la depresión también se empleó el análisis factorial exploratorio de los ítems junto con escalas de depresión. Así, Endler, Cox, Parker y Bagby (1992), realizaron un análisis factorial exploratorio en el que se incluyen 61 ítems: 1) 20 ítems del STAI (la subescala estado); 2) 20 ítems de la subescala de ansiedad estado de la escala de ansiedad multidimensional de Endler (Endler, Edwards y Vitelli, 1991); 3) 21 ítems del cuestionario de depresión de Beck (BDI; Beck, Ward, Mendelson, Mock y Erbaugh, 1961). Con este total de 61 ítems se realiza dicho análisis factorial exploratorio. En los resultados se puede observar que hay cuatro factores: el primero con los 20 ítems de la

escala de Endler y del STAI; el segundo incluye 20 de los 21 ítems del BDI y los factores tercero y cuarto incluyen los ítems del STAI-estado. Este análisis se repitió empleando 101 ítems: 1) 20 ítems del STAI (en este caso de la subescala rasgo); 2) 60 ítems de la subescala de ansiedad estado de la escala de ansiedad multidimensional de Endler (Endler et al., 1991); 3) al igual que en el caso anterior, 21 ítems del cuestionario de depresión de Beck (BDI; Beck et al., 1961). De nuevo, los ítems del STAI no saturan en el factor que contiene los ítems de depresión.

En el estudio de la factorización del STAI junto con los ítems del BDI hay estudios que señalan problemas en las saturaciones de los ítems. Empleando la subescala de ansiedad rasgo se observó que los ítems 1, 10, 15 y 16 saturaban en el factor formado por los ítems de depresión (Andrade et al., 2001). Por ello, Andrade et al. (2001) defienden que dichos ítems del STAI en realidad evalúan depresión o malestar general. En ambas investigaciones las muestras empleadas para realizar las pruebas son procedentes de la población general.

La última aproximación para el estudio de la validez de contenido del STAI consiste en el uso de modelos de ecuaciones estructurales para la realización de análisis factoriales confirmatorios en los que se determine si hay conjuntos de ítems del STAI que formen factores propios y cuyo contenido no sea claro. Dentro de esta aproximación, cabe destacar el estudio de Bieling, Antony y Swinson (1998), en el que se comprueban varios modelos, concluyendo que el bifactorial es en el que se observa un mejor ajuste: el primer factor incluía 14 ítems y el segundo estaba formado por los otros 6 ítems de la subescala de ansiedad rasgo (ítems 1, 2, 3, 13, 15 y 16) que medirían teóricamente depresión. Seguidamente, se realizó una correlación entre la puntuación de este factor de “STAI-depresión” con el BDI, obteniendo un valor de 0,65. Pese a que la relación del factor de “STAI-ansiedad” con el BDI es muy similar a la realizada con el

“STAI-depresión”, estos autores concluyen que los seis ítems del factor dos evalúan depresión, pero en ningún momento se justifican las elevadas correlaciones del primer factor y el BDI. Además de plantear la estructura de un factor de depresión, concluyen que los ítems 2, 8, 9, 11, 17, 18 y 20 miden ansiedad rasgo de forma pura (Bieling et al., 1998).

El trabajo de Bieling supuso un punto de partida, siendo replicado completa o parcialmente en estudios similares. La mayoría de estos trabajos concluyen que algunos de los ítems de la subescala de ansiedad rasgo evalúan depresión (Bados, Gómez-Benito y Balaguer, 2010; Caci, Baylé, Dossios, Robert y Boyer, 2003). Otros autores defienden que, en realidad, dichos ítems evalúan afecto negativo general (Hill, Musso, Jones, Pella y Gouvier, 2012), llegando al extremo de plantear que el STAI-rasgo completo evaluaría afecto general negativo (Balsamo et al., 2013; Reiss, 1997), esto permite explicar las elevadas puntuaciones en ansiedad depresión, ya que el afecto negativo es propio tanto de los trastornos depresivos como de los ansiosos (Watson, Clark y Stasik, 2011). Según Caci et al. (2003), diversos ítems del STAI no evalúan ansiedad; más bien cunifican “preocupación”, “inquietud”, “auto-confianza”, “tendencia al fracaso” y “anhedonia”. Partiendo del contenido de las preguntas, seleccionan los ítems 2, 20, 9, 11, 17, 18, 7, 12, 14 y 15, ya que consideran que estos ítems medirían tres de dichos factores, siendo la ansiedad rasgo un factor de segundo orden.

En resumen, el STAI ha sido tachado de tres defectos principales: 1) diferentes factorizaciones; 2) vocabulario complejo; 3) problemas de validez de constructo. En lo que respecta a este último punto es necesario destacar que, en la mayoría de ocasiones, únicamente se emplea la subescala de ansiedad rasgo. Además, salvo excepciones, solamente utilizan muestras de población general, sin incluir personas diagnosticadas con

trastornos depresivos o ansiosos, lo que permitiría detectar los trastornos de la muestra, a partir de análisis factorial.

Segunda versión del STAI: forma Y

Tras las críticas recibidas por la primera versión del STAI (Spielberger et al. 1970), 13 años después se realizó una segunda versión, que incorporaba dos novedades sustanciales: 1) cambio en el contenido de algunos de los ítems y 2) se equiparó el número de ítems inversos (50% en ambas subescalas) para que el factor de ansiedad rasgo aumentase su relación con la ausencia de ansiedad (Spielberger, 1983; 1989). Desde su creación en 1983, se conoció como "forma Y", y, desde entonces, la primera versión comercial del STAI (Spielberger et al., 1970) se denominó "forma X", que sigue siendo utilizada, en España y otros países donde ya se había adaptado y no se validó la nueva forma. De hecho, pese a que hay notables diferencias entre ambas versiones y que la forma Y presenta unas propiedades psicométricas superiores, se puede emplear el cuestionario original ya que la correlación entre ambas formas oscila desde 0,96 a 0,98 (Spielberger, 1989).

Versión breve

El STAI se ha utilizado para la evaluación de los niveles de ansiedad en numerosas poblaciones, generándose además una notable cantidad de publicaciones en las que se analizan aspectos psicométricos del cuestionario. Dentro de la investigación psicométrica realizada sobre el STAI, cabe destacar el estudio de versiones breves del mismo. Las investigaciones al respecto comienzan tras la creación de la forma Y, teniendo en común el no contar con adecuada metodología y/o análisis estadísticos. El primer estudio que abordó la posibilidad de un conjunto de ítems que permita evaluar adecuadamente la ansiedad estado y rasgo fue el realizado por van Knippenberg, Duivenvoorden, Bonke y Passchier (1990). En este trabajo el objetivo principal era

abordar la posibilidad de que determinadas combinaciones de ítems midan ansiedad estado y rasgo de forma fiable. Para demostrar su hipótesis, realizaron un análisis de correlación ítem total corregida, seleccionando los tres ítems positivos y los tres ítems negativos en los que mayores puntuaciones se observaban. Estos autores realizaron nueve versiones breves, a partir de las correlaciones ítem-total corregidas, con variación en el número de ítems (combinaciones de 4, 6 y 8 ítems), siendo diferentes para cada muestra utilizada (dos muestras de pacientes médicos y una de estudiantes). Tras la selección de los ítems, analizaban las reducciones obtenidas, concluyendo que todas ellas eran válidas.

A partir de la forma Y, el primer trabajo realizado cuyo objetivo fuese establecer una reducción de ítems del STAI fue el realizado por Marteau y Bekker (1992). En este caso, siguiendo una metodología similar a la propuesta por van Knippenberg et al. (1990), se concluye que la selección de los ítems 1, 3, 6, 15, 16 y 17 es la más adecuada para la evaluación de la ansiedad. No obstante, en este trabajo la muestra estaba formada exclusivamente por mujeres embarazadas. En un intento de ampliar la muestra con el fin de generalizar su uso, van de Bij, de Weerd, Cikot, Steegers y Braspenning (2003) aplicaron esta versión en una muestra compuesta por mujeres embarazadas y sus parejas. La conclusión fue que la selección de esos ítems es válida también en el caso de las parejas. Otro de los estudios en los que se trató de confirmar esta versión breve es el de Davey, Barrant, Butow y Deeks (2007). En este caso se comparó la versión breve propuesta por Mateau y Bekker (1992) con una de las versiones de ocho ítems propuestas por van Knippenberg et al. (1990). Según los resultados obtenidos, Davey et al. (2007) afirman que ambas versiones breves presentan buenas propiedades psicométricas y pueden ser utilizadas indistintamente.

Chlan, Savik y Weinert (2003), validaron una versión de seis ítems (5, 9, 10, 12, 17 y 20) de la escala de estado del STAI, cuya originalidad es el empleo en pacientes con

dificultades respiratorias (limitados por intubación). En esta versión se comprobó que dicha versión tenía adecuadas propiedades en dicha muestra. El estudio fue replicado posteriormente, obteniendo resultados similares (Chen, 2006).

De las cinco versiones breves presentadas, salvo la de van Knippenberg et al., (1990), se realizaron a partir de la forma Y del STAI. Al revisar la literatura en busca de versiones breves desarrolladas a partir de la forma X, el número de propuestas y estudios es mucho menor. Además de la propuesta ya presentada, de van Knippenberg et al. (1990), se encuentran dos versiones breves de la adaptación brasileña (realizando la reducción de ítems para ambas subescalas). La primera de ellas es la de Fioravanti-Bastos, Cheniaux y Landeira-Fernandez (2011) en la que se estableció una versión de seis ítems para cada una de las escalas. En segundo lugar, Kaipper, Chachamovich, Hidalgo, Torres y Caumo (2010) propusieron una selección de 13 ítems para la ansiedad estado y 12 para la de rasgo, que no muestra a pesar de su mayor longitud, ni mejores índices de consistencia interna ni correlaciones superiores con la escala completa.

Pese a la utilidad de las versiones breves al evaluar ansiedad, es preciso considerar que generan mayores tasas de error de medida. Otra consideración es señalar que la consistencia interna media de varios estudios es de 0,88, algo menor que la fiabilidad media de la escala completa (Kruyen, Emons y Sijtsma, 2013).

STAI en España

En España el STAI se comercializó en 1982 (Seisdedos Cubero, 1982) a partir de la de los trabajos de Bermúdez (1978a; 1978b), en los que se observa una consistencia interna superior a 0,9, correlaciones test re-test entre 0,3 y 0,77 a los cien días, correlaciones con medidas de ansiedad elevada y una factorización de cuatro factores que encaja con la propuesta de ansiedad estado/rasgo positiva y negativa. El buen

funcionamiento psicométrico del STAI hace que sea el séptimo cuestionario más utilizado por psicólogos clínicos en España (Muñiz y Fernández-Hermida, 2010).

No se ha encontrado ninguna versión breve comprobada en muestra general española, aunque existen dos trabajos realizados en pacientes con intubación respiratoria (Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez y Martínez-Durá, 2011), derivados de la versión breve propuesta por Chlan et al. (2003). Esta adaptación española muestra una adecuada fiabilidad (Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011) así como pruebas de validez de contenido (Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez et al., 2011). Si bien esta investigación supuso un hito en España, cabe indicar dos limitaciones: 1) el contar únicamente con pacientes hospitalarios que están sometidos a un proceso de respiración asistida con intubación y 2) la selección de ítems utilizada procede de la forma Y del STAI (Chlan et al., 2003), mientras que el cuestionario empleado es el STAI en su forma X (Seisdedos Cubero, 1982), por lo que el contenido de los ítems no es el mismo, confundiendo los resultados.

Objetivos

Esta Tesis Doctoral tiene tres objetivos principales: el primero de ellos, comprobar la fiabilidad del STAI, el segundo estudiar si es posible establecer una versión breve de la adaptación española del STAI y el tercero tratar de establecer la fiabilidad y validez del STAI en muestras clínicas. Para ello es necesario cumplir los siguientes objetivos específicos:

- Estudiar la fiabilidad del STAI y la detección de los niveles de ansiedad en pacientes diagnosticados con ansiedad.
- Analizar la fiabilidad y el funcionamiento diferencial de los ítems del STAI en

población general española.

- Establecer una versión breve del STAI que garantice adecuados niveles de fiabilidad.

- Estudiar el ajuste de la versión breve (si la hubiera) en adolescentes y universitarios.

- Comprobar la estructura factorial en pacientes diagnosticados con trastornos depresivos.

- Comprobar la estructura factorial en pacientes diagnosticados con trastornos ansiosos.

Para culminar el triple objetivo de esta tesis es necesario lograr cada uno de los epígrafes propuestos.

Referencias

- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004a). 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 28–37. doi:10.1600-0047.2004.00328.x
- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004b). Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 21–27. doi:10.1111/j.1600-0047.2004.00326
- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004c). Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 47–54. doi:10.1111/j.1600-0047.2004.00330.x
- Andlin-Sobocki, P. y Wittchen, H. U. (2005). Cost of anxiety disorders in Europe. *European Journal of Neurology*, 12(Suppl. 1), 39–44. doi:10.1111/j.1468-1331.2005.01196.x
- Andrade, L., Gorenstein, C., Vieira Filho, A. H., Tung, T. C. y Artes, R. (2001). Psychometric properties of the Portuguese version of the State-Trait Anxiety Inventory applied to college students: factor analysis and relation to the Beck Depression Inventory. *Brazilian Journal of Medical and Biological Research*, 34, 367-374. doi:10.1590/S0100-879X2001000300011

- Bados, A., Gómez-Benito, J. y Balaguer, J. (2010). The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety? *Journal of Personality Assessment*, 92, 560-567. doi: 10.1080/00223891.2010.513295
- Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L. y Saggino, A. (2013). The State-Trait Anxiety Inventory: Shadows and Lights on its Construct Validity. *Journal of Psychopathology and Behavioral Assessment*, descargado el 13 de agosto de 2013 de <http://link.springer.com/article/10.1007/s10862-013-9354-5>. doi:10.1007/s10862-013-9354-5
- Barnes, L. L. B., Harp, D. y Jung W. S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618. doi: 10.1177/0013164402062004005
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. y Erbaugh, J. (1961). An Inventory for Measuring Depression. *Archives of General Psychiatry*, 4, 561-571. doi:10.1001/archpsyc.1961.01710120031004
- Bermúdez, J. (1978a). Anxiety and performance. *Revista de Psicología General y Aplicada*, 151, 183-207.
- Bermúdez, J. (1978b). Functional analysis of anxiety. *Revista de Psicología General y Aplicada*, 153, 617-634.
- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K. y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology*, 11, 443-457.

- Bieling, P. J., Antony, M. M. y Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behavior Research and Therapy*, *36*, 777-788. doi: 10.1016/S0005-7967(98)00023-0
- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos Cubero, N. (2011). *Cuestionario de ansiedad estado-rasgo* (8ª edición). Madrid: TEA ediciones.
- Caci, H., Baylé, F. J., Dossios, C., Robert, P. y Boyer, P. (2003). The Spielberger trait anxiety inventory measures more than anxiety. *European Psychiatry*, *18*, 394-400. doi:10.1016/j.eurpsy.2003.05.003
- Chen, Y. J. (2006). *Psychophysiological determinants of repeated ventilator weaning failure* (Tesis doctoral). Descargada de <http://www.nursing.arizona.edu>
- Chlan, L., Savik, K. y Weinert, C. (2003). Development of a Shortened State Anxiety Scale From the Spielberger State-Trait Anxiety Inventory (STAI) for Patients Receiving Mechanical Ventilatory Support. *Journal of Nursing Measurement*, *11*, 283-293. doi:10.1891/jnum.11.3.283.61269
- Davey, H. M., Barratt, A. L., Butow, P. N. y Deeks, J. J. (2007). A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *Journal of Clinical Epidemiology*, *60*, 356-360. doi: 10.1016/j.jclinepi.2006.07.015
- Endler, N. S., Cox, B. J., Parker, J. D. A. y Bagby, R. M. (1992). Self-Reports of Depression and State-Trait Anxiety: Evidence for Differential Assessment. *Journal of Personality and Social Psychology*, *63*, 832-838. doi:10.1037/0022-3514.63.5.832
- Endler, N. S., Edwards, J. M. y Vitelli, R. (1991). *Endler Multidimensional Anxiety Scales (EMAS): Manual*. Los Angeles, CA: Western Psychological Services.

- Fioravanti-Bastos, A. C. M., Cheniaux, E. y Landeira-Fernandez, J. (2011). Development and Validation of a Short-Form Version of the Brazilian State-Trait Anxiety Inventory. *Psicologia: Reflexão e Crítica*, 24, 485-494. doi:10.1590/S0102-79722011000300009
- Grös, D. F., Antony, M. M., Simms, L. J. y McCabe, R. E. (2007). Psychometric Properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment*, 19, 369-381. doi: 10.1037/1040-3590.19.4.369
- Hill, B. D., Musso, M., Jones, G. N., Pella, R. D. y Gouvier Wm. D. (2012). A Psychometric Evaluation of the STAI-Y, BDI-II, and PAI Using Single and Multifactorial Models in Young Adults Seeking Psychoeducational Evaluation. *Journal of Psychoeducational Assessment*, XX, 1-13. doi: 10.1177/0734282912462670
- Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T. y Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory Scores for Ethnically Diverse Adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83. doi:10.1037/1099-9809.6.1.73
- Kaipper, M. B., Chachamovich, E., Hidalgo, M. P. L., Torres, I. L. S. y Caumo, W. (2010). Evaluation of the structures of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research*, 68, 223-233. doi:10.1016/j.jpsychores.2009.09.013
- Katkin, E. S. (1975). Review of the State-Trait Anxiety Inventory. En: O. K. Buros (Ed.), *The eighth mental measurements yearbook*, (Vol. 1) (pp.1095-1099). Lincoln: University of Nebraska Press.

- Kennedy, B. L., Schwab, J. J., Morris, R. L. y Beldia, G. (2001). Assessment of the state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatric Quarterly*, 72, 263-276. doi:10.1023/A:1010305200087
- Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M. y Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, 21, 169-184. doi: 10.1002/mpr.1359.
- Kessler, R. C., Ruscio, A. M., Shear, K. y Wittchen, H.-U. (2009). Epidemiology of Anxiety Disorders. En M. B. Stein y T. Stecker (eds.), *Behavioral Neurobiology of Anxiety and Its Treatment*, [libro electrónico]. doi:10.1007/7854_2009_9.
- Kruyen, P. M., Emons, W. H. M. y Sijtsma, K. (2013). Shortening the S-STAI: Consequences for research and clinical practice. *Journal of Psychosomatic Research* 75, 167-172. doi:10.1016/j.jpsychores.2013.03.013
- Marteau, T. M. y Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31, 301-306. doi:10.1111/j.2044-8260.1992.tb00997.x
- Micin, S. y Bagladi, V. (2011). Salud Mental en Estudiantes Universitarios: Incidencia de Psicopatología y Antecedentes de Conducta Suicida en Población que Acude a un Servicio de Salud Estudiantil. *Terapia Psicológica*, 29, 53-64.
- Muñiz, J. y Fernández-Hermida, J.R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.

- Perpiñá-Galvañ, J., Richart-Martínez, M. y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189. doi:10.1016/j.arbres.2010.11.006
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J. y Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem*, 19, A04:1-A04:6.
- Reiss, S. (1997). Trait Anxiety: it's not what you think it is. *Journal of Anxiety Disorders*, 11, 201-214. doi:10.1016/S0887-6185(97)00006-6
- Seisdedos Cubero, N. (1982). *Manual del Cuestionario de Ansiedad Estado/Rasgo (STAI)*. Madrid, España: TEA Ediciones.
- Somers, J. M., Goldner, E. M., Waraich, P. y Hsu, L. (2006). Prevalence and Incidence Studies of Anxiety Disorders: A Systematic Review of the Literature. *Canadian Journal of Psychiatry*, 51, 100-113.
- Spielberger, C.D. (1973). *Manual for the State-Trait Anxiety Inventory for Children*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D. (1983). *State-Trait Anxiety Inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D. (1989). *State-Trait Anxiety Inventory: A comprehensive bibliography* (2º ed.). Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C.D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.

- Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- van der Bij, A. K., de Weerd, S., Cikot, R. J. L. M., Steegers, E. A. P. y Braspenning, J. C. C. (2003). Validation of the Dutch Short Forma of the State Scale of the Spielberger State-Trait Anxiety Inventory: Considerations for Usage in Screening Outcomes. *Community Genetics*, 6, 84-87. doi: 10.1159/000073003
- van Knippenberg, F. C. E., Duivenvoorden, H. J., Bonke, B. y Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology*, 43, 995-1000. doi:10.1016/0895-4356(90)90083-2
- Vautier, S. (2004). A Longitudinal SEM Approach to STAI Data: Two Comprehensive Multitrait-Multistate Models. *Journal of Personality Assessment*, 83, 167-179. doi:10.1207/s15327752jpa8302_11
- Vera-Villarroel, P., Celis-Atenas, K., Córdova-Rubio, N., Buela-Casal, G. y Spielberger, C. D. (2007). Preliminary Analysis and Normative Data of the State-Trait Anxiety Inventory (STAI) in Adolescent and Adults of Santiago, Chile. *Terapia psicológica*, 25, 155-162. doi:10.4067/S0718-48082007000200006
- Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E. J.,... Wells, J. E. (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet*, 370, 841-850. doi:10.1016/S0140-6736(07)61414-7
- Watson, D., Clark, L. A. y Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 429-442.

OBJETIVO 1
COMPROBACIÓN DE LA
FIABILIDAD Y VALIDEZ DEL
STATE-TRAIT ANXIETY INVENTORY

Estudio 1

**Diferencias entre pacientes con ansiedad y población general mediante STAI:
meta-análisis de comparación de grupos y meta-análisis de generalización de la
fiabilidad**

Alejandro Guillén-Riquelme y Gualberto Buela-Casal

Artículo en prensa en *Revista Española de Salud Pública*

Revista indexada en el *Journal Citation Reports*

Factor de impacto 2012 = 0,696.

Referencia:

Guillén-Riquelme, A. y Buela-Casal, G. (en prensa). Diferencias entre pacientes con ansiedad y población general mediante STAI: meta-análisis de comparación de grupos y meta-análisis de generalización de la fiabilidad. *Revista Española de Salud Pública*, 88.

**Diferencias entre pacientes con ansiedad y población general mediante STAI:
meta-análisis de comparación de grupos y meta-análisis de generalización de la
fiabilidad**

Resumen

Fundamentos: Desde su creación el STAI se ha citado en más de 14.000 documentos, contando con más de 60 adaptaciones en diversos países. En algunas de ellas este instrumento no cuenta con puntuaciones clínicas. El objetivo de este trabajo es determinar si el STAI tiene puntuaciones superiores en pacientes diagnosticados con ansiedad respecto a población general. Además, se pretende analizar si la consistencia interna es adecuada en personas con ansiedad. Método: Se realizó una búsqueda bibliográfica en *Tripdatabase*, *Cochrane*, *Web of Knowledge*, *Scopus*, *PyscINFO* and *Scholar Google*, para documentos publicados entre 2008 y 2012. Se seleccionaron 131 artículos para la comparación entre pacientes diagnosticados con ansiedad respecto a población general y 25 para la generalización de la fiabilidad. Para los análisis se utilizó la *d* de Cohen para la comparación de medias (método de efectos aleatorios) y el alfa de Cronbach para la generalización de la fiabilidad (método de efectos fijos). Resultados: En la comparación entre grupos en la ansiedad estado las diferencias fueron estadísticamente significativas ($d = 1,39$; IC 95%: 1,22-1,56). En la ansiedad rasgo fue $d = 1,74$ (IC 95%: 1,56-1,91). Además de ello, la fiabilidad para pacientes con algún trastorno de ansiedad abarcó de 0,87 a 0,93. Conclusiones: el STAI es sensible al nivel de ansiedad de la persona y fiable para muestras de pacientes diagnosticados con crisis de angustia, fobia específica, fobia social, fobia social generalizada, trastorno de ansiedad generalizada, trastorno de estrés post-traumáticos, trastorno obsesivo compulsivo o trastorno por estrés agudo.

Palabras clave: Ansiedad, meta-análisis, fiabilidad.

Abstract

Background: Since its creation the STAI has been cited in more than 14,000 documents, with more than 60 adaptations in different countries. In some adaptations this instrument has no clinical scores. The aim of this work is to determine if the STAI has higher scores in patients diagnosed with anxiety than in general population. In addition, we want to examine if the internal consistency is adequate in anxious patient samples.

Methods: We performed a literature search in *Tripedatabase*, *Cochrane*, *Web of Knowledge*, *Scopus*, *PyscINFO* and *Scholar Google*, for documents published between 2008 y 2012. We selected 131 scientific articles to compare between patients diagnosed with anxiety and general population, and 25 for the generalization of reliability. For the analysis we used Cohen's *d* for means comparisons (random-effects method) and Cronbach's alpha for the reliability generalization (fixed-effects method). **Results:** In the groups comparison the differences in state anxiety was significant ($d = 1,39$; CI 95%: 1,22-1,56). In the trait anxiety was $d = 1,74$ (CI 95%: 1,56-1,91). The reliability for patients of some anxiety disorder was between 0.87 and 0.93. **Discussion:** So it seems that the STAI is sensitive to the level of anxiety of the individual and reliable for patients with diagnosis of panic attack, specific phobia, social phobia, generalized social phobia, generalized anxiety disorder, post-traumatic stress disorder, obsessive compulsive disorder o acute Stress disorder.

Keywords: Anxiety, meta-analysis, reliability.

Los trastornos de ansiedad son uno de los problemas psicológicos con una mayor prevalencia. En una muestra de más de 9.000 estadounidenses se observaron tasas de prevalencia que alcanzaban el 15,6% en el caso de la fobia específica (Kessler, Petukhova, Sampson, Zaslavsky y Wittchen, 2012). Estos resultados coinciden con medidas epidemiológicas previas (Kessler, Ruscio, Shear y Wittchen, 2009). Además, pese a que existen diferencias entre los continentes, la tasa de diagnósticos de ansiedad es muy elevada (Somers, Goldner, Waraich y Hsu, 2006; Wang et al., 2007).

Entre los distintos instrumentos disponibles para evaluar la ansiedad general se encuentra el *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch y Lushene, 1970). Este cuestionario evalúa dos facetas de la ansiedad. En primer lugar, la ansiedad rasgo, entendida como los factores personales que influyen en la percepción del nivel de ansiedad de los estímulos que se presentan. Es decir, una persona con alta ansiedad rasgo percibiría un estímulo como más amenazante y ansiógeno que una persona con baja ansiedad rasgo. Por el contrario, la ansiedad estado se corresponde con los estímulos desencadenantes de ansiedad que hay en el entorno cercano del sujeto evaluado. Ambas escalas son teóricamente independientes (Spielberger et al., 1970). Tras su creación se modificaron algunas preguntas y la escala de respuesta del STAI original, llamando a la nueva versión forma Y y nombrando desde entonces a la versión original como forma X (Spielberger, Gorsuch, Lushene, Vagg y Jacobs, 1983). En lo que respecta a su consistencia interna el alfa de Cronbach medio para 52 y 51 estudios incluidos en un meta-análisis fueron de 0,91 en ansiedad estado y de 0,89 en ansiedad rasgo (Barnes, Harp y Jung, 2002).

El STAI es uno de los cuestionarios más empleados, contando con más de 60 adaptaciones culturales y lingüísticas y con más de 14.000 citas en artículos de sus versiones para adultos (Barnes et al., 2002). Además, el STAI ha sido utilizado para

medir el nivel de ansiedad en pacientes con diversos trastornos físicos y psicológicos (Spielberger y Reheier, 2009). Por ejemplo, el STAI se utilizó en un meta-análisis para evaluar el mejor tratamiento para el trastorno de ansiedad generalizada, empleando la subescala de rasgo como principal medida de eficacia del tratamiento (Fisher y Durham, 1999). En dicho meta-análisis se incluyen seis estudios en los que se emplea el STAI con esta muestra. Otro de los ejemplos, en los que se hace uso del STAI para evaluar los niveles de ansiedad en personas que sufren alguno de los trastornos de ansiedad es el artículo de Kvaal, Ulstein, Nordhus y Engedal (2005). En este estudio se aplica el STAI a personas ancianas con diversos trastornos de ansiedad, hallando que las personas con trastornos de ansiedad puntúan significativamente más que los que no tienen ningún diagnóstico psicológico.

Por último, en un meta-análisis en el que se incluyen casi 200 estudios se concluye que el STAI estado es una medida sensible al nivel de ansiedad en grupos clínicos, así como en personas con diversos trastornos físicos (Rossi y Pourtois, 2012). Además de ello, este instrumento es lo suficientemente preciso como para detectar diferencias en las puntuaciones de sujetos al ser sometidos a estímulos estresantes o procedimientos para aumentar sus niveles de ansiedad (Rossi y Pourtois, 2012).

Por todo lo expuesto hasta el momento, se establecen los siguientes objetivos: en primer lugar, se pretende analizar si el STAI detecta las diferencias de ansiedad de grupos clínicos (de diferentes trastornos de ansiedad en función del DSM-IV-TR; American Psychological Association, 2002) y grupos de sujetos sanos, ya que no se ha encontrado en la literatura ningún meta-análisis al respecto. En segundo lugar, determinar si la consistencia interna del STAI en muestras de pacientes ansiosos es adecuada y si está influida por la forma del STAI o el país de la muestra, así como por tipo de trastorno ansioso de la muestra. La selección del DSM-IV-TR en lugar de la

nueva versión (DSM-5) se realizó debido a que en el momento de las búsquedas todavía no estaba publicada esta nueva versión y además los trabajos incluidos emplean la cuarta edición o anteriores.

Método

Tipo de estudio

Con el fin de llevar a cabo los objetivos propuestos en este estudio, clasificado como un estudio teórico de revisión con meta-análisis (Montero y León, 2007), se siguieron las recomendaciones PRISMA, ya que este documento presenta una guía para este tipo de artículos (Urrútia y Bonfill, 2013), y para su redacción se siguió las propuestas presentadas por Hartley (2012).

Búsqueda de los estudios

El proceso de búsqueda de los artículos se realizó en varias fases. En primer lugar, se localizaron revisiones sistemáticas o meta-análisis, mediante las opciones de búsqueda avanzada de las diversas bases on-line utilizadas y empleando *Tripdatabase* y la base *Cochrane*. Tras no identificar ninguna revisión previa que emplease únicamente el STAI, se pasó a la segunda fase de la búsqueda. Para ello se buscó en diversas fuentes generales y específicas, con el fin de hallar todos los documentos relevantes (Fernández-Ríos y Buela-Casal, 2009; Perestelo-Pérez, 2013) Así pues, se emplearon la *Web of Knowledge*, *Scopus* y *PyscINFO* (a través de la plataforma OVID). Se acotó la búsqueda a los artículos publicados entre los años 2008 y 2012, ya que al ser un cuestionario muy empleado, la inclusión de cinco años permitía encontrar un número representativo de artículos que lo empleen. Los términos de búsqueda se incluyeron únicamente en inglés.

La ecuación de búsqueda se formó, por una parte, con el nombre completo del STAI así como con el acrónimo y, por otra, con cada uno de los trastornos de ansiedad, indicados en el DSM-IV-R. El término STAI se buscó en todo el artículo, mientras que el trastorno de ansiedad se buscó únicamente en título, resumen y palabras clave. Se realizó una búsqueda independiente para cada uno de los trastornos. Se incluyeron truncadores siempre que era posible (v. gr. *phobi** o *obsessi**) e igualmente se utilizó el nombre completo del trastorno y el acrónimo del mismo, en el caso de que éste fuese de uso común (v. gr. *PTSD* o *GAD*). La última fase consistió en la búsqueda en *Scholar Google*. Este buscador se seleccionó con el fin de localizar la literatura gris y evitar el sesgo de publicación. Para realizar esta búsqueda se identificó el artículo original del STAI y, a través de las citas del mismo, se filtraron los años de interés y se verificaron los artículos resultantes. Este proceso se realizó desde el 21 de julio al 30 de septiembre de 2012. Tras descartar los artículos que se podían excluir por título, en total se encontraron un total de 674 referencias.

Tras la búsqueda, se trató de localizar el texto completo de los artículos. Para ello, se empleó su versión on-line (incluyendo manuscritos publicados) y mediante el préstamo inter-bibliotecario. Para los artículos a los que no se pudo tener acceso se escribió un correo electrónico al autor, donde se informaba del objetivo del estudio y se solicitaba una copia del artículo de interés. Los artículos no obtenidos finalmente fueron descartados (3,12% respecto al total de artículos).

Criterios de selección de los estudios

Una vez seleccionados los artículos, para determinar la inclusión definitiva de los mismos se emplearon diversos criterios explícitos (Sánchez-Meca, 2010): a) los artículos debían estar escritos en inglés, español o portugués, incluyendo muestras de cualquier país y publicados igualmente sin importar la región geográfica; b) únicamente

se incluyeron artículos cuasi-experimentales y experimentales, excluyendo revisiones teóricas y estudios de caso único. Las revisiones sistemáticas y los meta-análisis también se incluyeron inicialmente, no encontrando ninguno que tuviese el mismo objetivo que en el presente artículo. En las citas de estos trabajos se buscaron artículos que reuniesen los criterios para ser incluidos en este estudio; c) en lo que respecta a la muestra sólo se seleccionaron aquellos estudios realizados con mayores de edad, cuyo diagnóstico principal fuese uno de los trastornos de ansiedad establecidos en el DSM-IV-TR (American Psychological Association, 2002). Además, dicho diagnóstico no podía estar justificado mediante las puntuaciones de STAI como único criterio y d) se descartaron los artículos donde los problemas ansiosos de las personas de la muestra fuesen debidos a causas médicas.

Respecto a la versión aplicada, en los artículos seleccionados se podía emplear la versión completa del STAI o una de las dos subescalas (estado o rasgo) indistintamente. De la misma forma, la versión del STAI aplicada podía ser tanto la original como una adaptación del mismo, siempre que hubiese sido publicada y respetase el número de ítems y la escala de respuesta se mantuviese en cuatro alternativas. En diversas adaptaciones de STAI se recodifica la escala original (de 1 a 4) a una nueva escala de 0 a 3. Tal y como se indica en la literatura (Spielberger et al., 1983), para realizar comparaciones entre diferentes adaptaciones del STAI o con la versión original, se debe sumar 20 a la puntuación obtenida, en cada una de las subescalas o a las medias. Así pues, en los casos en los que se empleaba adaptaciones con este tipo de escala de respuesta, se realizó esta transformación.

Por otra parte, también, se incluyeron algunos criterios específicos para ambos tipos de análisis. En primer lugar, para la generalización de la fiabilidad se seleccionaron artículos en los que se incluyese un grupo clínico, sin importar o no la

existencia de un grupo control. Mientras en el caso de la comparación entre grupos únicamente se seleccionaron artículos con la existencia de un grupo control. Si los grupos se establecían mediante puntuaciones de un cuestionario de algún trastorno de ansiedad estos artículos eran descartados si el grupo control estaba formado únicamente por personas con puntuaciones bajas en dicha medida. En los artículos que informaban de un alfa rango (para varios trastornos de ansiedad sin especificar un alfa para cada uno de ellos) se tomó el menor de los valores. En el caso de contar con medias pre y post-tratamiento siempre se seleccionó la media pre-tratamiento de ambos grupos. Además de todo ello, se excluyó un artículo en que se preguntaba a familiares y amigos de suicidas la puntuación que el fallecido habría puesto en el STAI; así como otra investigación cuyas puntuaciones medias eran superiores a 80 (puntuación teórica máxima de la escala). En la Figura 1 se muestra un resumen del proceso de exclusión de artículos.

En los estudios en los que se empleaba varias muestras clínicas la inclusión o no era diferente en función de la finalidad del análisis (generalización de la fiabilidad o comparación entre grupos). En la generalización de la fiabilidad se incluían los trastornos de forma independiente, ya que al tratarse de muestras independientes no implicaba un incumplimiento de los supuestos del meta-análisis. En el caso de la comparación de medias no se podía optar por esta solución, ya que al mantenerse el mismo grupo control habría que duplicarlo y el cálculo del efecto de este grupo no cumpliría la independencia de las medidas. En este caso se optó por eliminar el grupo de pacientes de ansiedad con menor número de sujetos.

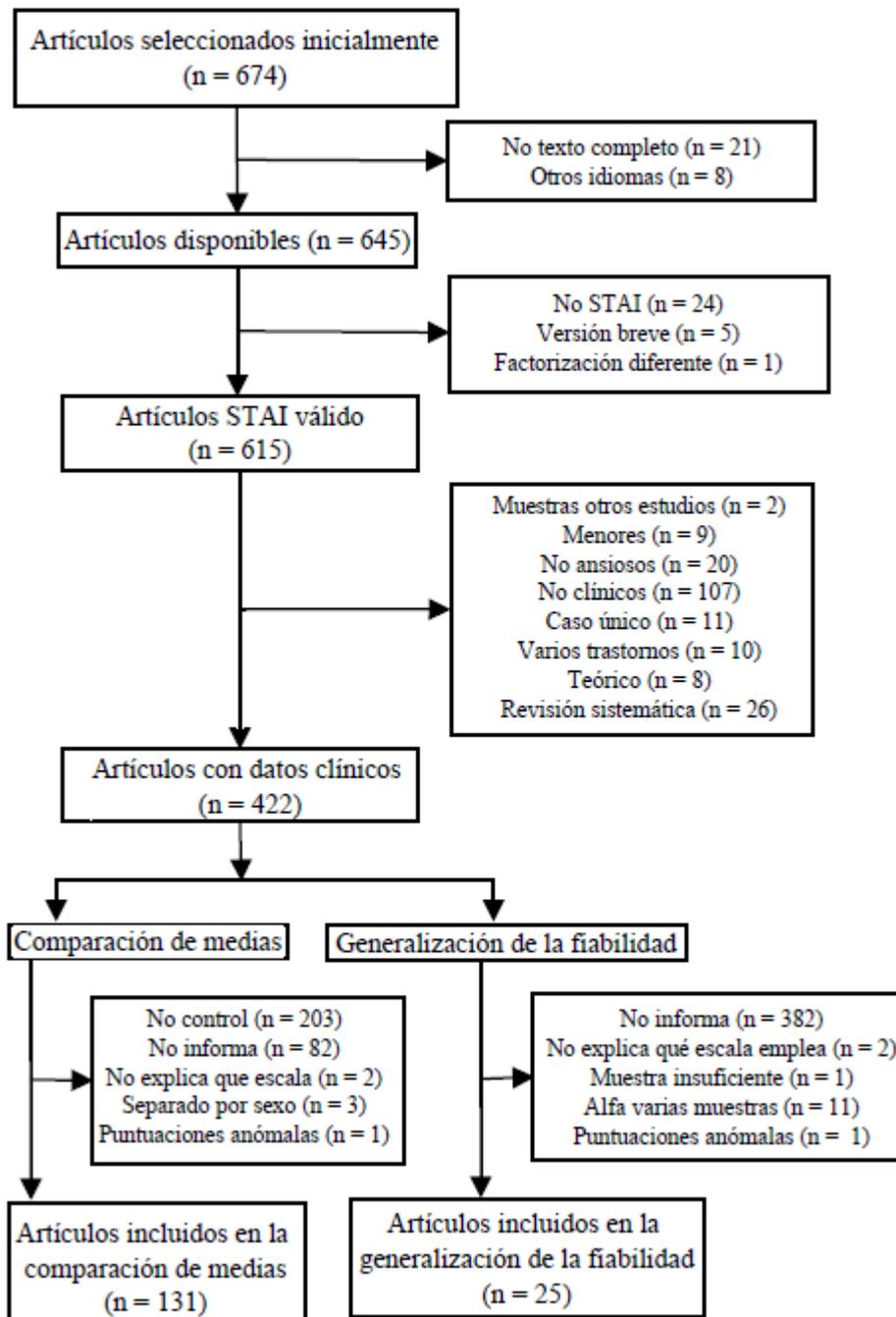


Figura 1. Diagrama del proceso de exclusión de artículos y frecuencia de artículos excluidos por cada uno de los criterios.

Índices del tamaño del efecto

Para el análisis de la comparación de medias el tamaño del efecto utilizado fue la d de Cohen y para la generalización de la fiabilidad se seleccionó el alfa de Cronbach.

Extracción de datos

De cada uno de los artículos seleccionados se extrajeron las siguientes medidas para efectuar la media en el meta-análisis:

- Medidas d de Cohen para el tamaño del efecto en las comparaciones entre grupos. En el caso de no informar del mismo, se extrajeron de los valores *t-student* o F y los grados de libertad, para poder estimarlo. En la mayoría de los casos, se emplearon las medias, las desviaciones típicas o varianzas y los tamaños de cada grupo, con el fin de estimar el tamaño del efecto para cada uno de los estudios.
- En el caso de la generalización de la fiabilidad, se extrajeron los valores del alfa de Cronbach en todos los casos, ya que no se hallaron estudios que informasen de otras medidas similares para el cálculo de la consistencia interna.

Además de las variables para evaluar el efecto, también se incluyeron diversos moderadores para ver su influencia sobre una posible homogeneidad de la generalización de la fiabilidad o del tamaño del efecto. En primer lugar, se empleó el país donde se reclutaba la muestra del artículo. En segundo lugar se codificó la forma del STAI utilizada, pudiendo ser la forma X o Y. No se incluyó el uso de adaptaciones a los diferentes países (usando sólo forma X o Y) ya que este posible efecto se analizaba en la variable del país de la muestra. Por último, se empleó el tipo de trastorno de ansiedad estudiado, empleando la codificación del DSM-IV-TR. Además de ello, se extrajo para su empleo como moderadores el tamaño muestral de los estudios, la edad media de los participantes, la distribución por género de la muestra (codificada como porcentaje de mujeres en la muestra). Las diferentes medidas fueron extraídas por dos codificadores independientes con el fin de valorar la fiabilidad del proceso de

codificación. Para ello se empleó el índice Kappa y el coeficiente de correlación intraclase, obteniendo un valor medio de todos los índices de 0,844.

Análisis estadístico

Respecto a los análisis, en el caso de la comparación de grupos, para seleccionar un método de estimación de efectos fijos o aleatorios no se empleó ningún estadístico (Sánchez-Meca, Marín Martínez y Huedo Medina, 2006). Así pues, para el presente estudio, se puede asumir variabilidad entre cada uno de los tamaños del efecto especificar a qué efecto se refiere en este caso, debido a que proceden de poblaciones diferentes (Sánchez-Meca et al., 2006). Esta asunción se basa en que los participantes provienen de diversos países (y han cumplimentado, por tanto, diferentes adaptaciones del STAI) e, igualmente, cada estudio se conforma por un grupo clínico, incluyendo diversos trastornos de ansiedad en la base final. Por ello, se optó por utilizar un método de efectos aleatorios. La medida del tamaño del efecto, seleccionada para realizar el meta-análisis, fue la d de Cohen (Cohen, 1988). El método seleccionado es el DerSimonian-Laird (DerSimonian y Laird, 1986). Este se considera adecuado para modelos de efectos aleatorios con una medida del tamaño del efecto continua como es la d de Cohen (Espallargues i Carreras y Tebé Cordoní, 2006). Se empleó el estadístico Q para comprobar la homogeneidad de los datos y se complementó con el índice I^2 (Sánchez-Meca, 2010). En el caso de hallar heterogeneidad en los resultados se empleó un ANOVA para determinar la influencia de diversos moderadores sobre la misma.

Para la generalización del alfa de Cronbach, se transformó este estadístico a valores T , mediante la fórmula: $T = (1 - \alpha)^{1/3}$, donde α es el coeficiente alfa obtenido en cada estudio. Estos valores permiten transformar el alfa de Cronbach en una medida que se distribuye normalmente, requisito necesario en los meta-análisis (Sánchez-Meca y López-Pina, 2008). Además, estas puntuaciones se ponderaron por el inverso de la

varianzas de los estudios, para calcular el tamaño medio de las puntuaciones T . Tras la obtención de un valor T medio, los valores finales se volvieron a convertir en puntuaciones alfa con el fin de lograr un índice más fácilmente interpretable. El método fue de efectos fijos, ya que al contar con un número reducido de artículos, no cabe asumir el cumplimiento de los supuestos necesarios para utilizar un modelo de efectos aleatorios. Para la realización de ambos meta-análisis se empleó el programa R , concretamente el paquete *metafor* (Viechtbauer, 2010).

Resultados

De los artículos incluidos el número de participantes en total fue de 2.476 (mínimo 6 y máximo 122) en el grupo clínico y 3.780 (mínimo 7 y máximo 1.329) en el grupo control de la comparación de medias en la ansiedad estado; de 3.104 (mínimo 6 y máximo 122) en el grupo clínico y 4.513 (mínimo 7 y máximo 1.324) en el grupo control, para la ansiedad rasgo; de 1.182 (mínimo 20 y máximo 406) en el grupo clínico de generalización de la fiabilidad estado y 1.631 (mínimo 20 y máximo 406) en la fiabilidad de la ansiedad rasgo. La media de edad osciló de 30,78 a 40,94 entre los cuatro meta-análisis. La media de ansiedad estado fue de 44,94 y de 50,26 en ansiedad rasgo. En el caso de la fiabilidad, el alfa medio sin transformar fue de 0,9 para la ansiedad estado y 0,88 en ansiedad rasgo.

Seguidamente, se comprobó si las diferencias de ansiedad entre los grupos diagnosticados con algún trastorno de ansiedad respecto a los sujetos sin trastornos eran estadísticamente significativas. Para ello en primer lugar se estudió la influencia de un posible sesgo, mediante el test de Egger, observando problemas en el caso de la ansiedad rasgo ($Z = 7,03$, $p < 0,001$), pero no en la ansiedad estado ($Z = 1,76$, $p = 0,08$). Así pues, en el caso de la ansiedad estado se observó que el tamaño medio del efecto fue de $d = 1,39$ (IC 95%: 1,22-1,56) resultando estadísticamente significativo ($p < 0,001$).

Tras ello, se analizó la homogeneidad de los tamaños del efecto, con el fin de determinar si había variabilidad entre los estudios. El test de heterogeneidad resultó significativo ($Q_{(93)} = 634,51$, $p < 0,001$). El indicador I^2 fue igual a 85,34%, lo que implica una heterogeneidad alta.

En el caso de la ansiedad rasgo el tamaño del efecto medio fue de $d = 1,74$ (IC 95%: 1,56, -1,91) y, de nuevo, siendo esta medida estadísticamente significativa ($p < 0,001$). En las pruebas de heterogeneidad se observó que hay variabilidad en las aportaciones de los diferentes estudios a la media ($Q_{(109)} = 910,52$; $p < 0,001$). Este resultado se confirmó mediante el resto de estadísticos empleados ($I^2 = 88,03\%$).

Tras comprobar que, en ambos casos, se encuentra un tamaño del efecto medio significativo y clínicamente relevante, además de observar que los resultados resultan heterogéneos, cabía estudiar qué variables explican dicha variabilidad. Para ello, se realizaron varios ANOVA utilizando el país de la muestra empleada, la forma del STAI (X o Y) aplicada y el tipo de trastorno de la muestra clínica. En lo que respecta a ansiedad estado, el país ($Q_{m(17)} = 49,02$; $p < 0,001$; $Q_{e(76)} = 419,08$; $p < 0,001$) y la forma del STAI ($Q_{m(1)} = 7,61$; $p = 0,006$; $Q_{e(91)} = 596,45$; $p < 0,001$) resultaron significativos, mientras que el tipo de trastorno no ($Q_{m(7)} = 6,48$; $p = 0,49$; $Q_{e(86)} = 602,57$; $p < 0,001$). En el caso de la ansiedad rasgo el país ($Q_{m(18)} = 22,69$; $p = 0,2$; $Q_{e(91)} = 742,25$; $p < 0,001$) y la forma del STAI ($Q_{m(1)} = 0,85$; $p = 0,356$; $Q_{e(107)} = 902,61$; $p < 0,001$) no fueron significativos, al contrario que el tipo de trastorno ($Q_{m(7)} = 26,72$; $p < 0,001$; $Q_{e(102)} = 176,39$; $p < 0,001$). Seguidamente, se analizó los niveles en los que se producía la variabilidad observada. En la Tabla 1 se recoge un resumen de los coeficientes obtenidos para cada una de las categorías de dichas variables. Además se emplearon moderadores continuos: el tamaño muestra la media, la desviación típica, la edad media y el porcentaje de mujeres en la muestra. De las variables continuas emplea-

Tabla 1.
Coefficientes de los moderadores en el meta-ANOVA de comparación de grupos para la ansiedad estado y rasgo.

Variables (<i>k</i> estado; <i>k</i> rasgo) significativo estado/significativo rasgo	Ansiedad estado					Ansiedad rasgo				
	<i>d</i> de Cohen	Error estándar	<i>Z</i>	I.C.		<i>d</i> de Cohen	Error estándar	<i>Z</i>	I.C.	
				Inferior	Superior				Inferior	Superior
País del estudio										
Alemania (17;27)	1,07	0,19	5,76***	0,71	1,44	1,51	0,18	8,55***	1,17	1,86
Australia (6;6)	1,30	0,33	3,98***	0,66	1,95	1,89	0,39	4,85***	1,13	2,66
Canadá (1;1)+/-	3,19	0,85	3,76***	1,53	4,85	0,65	0,94	0,7	-1,18	2,49
China (2;2)	1,38	0,53	2,6**	0,34	2,42	1,39	0,64	2,18*	0,14	2,63
EEUU (21;29)	1,49	0,17	8,76***	1,15	1,81	2,09	0,17	12,19***	1,75	2,42
España (7;3)	1,02	0,28	3,62***	0,47	1,57	1,14	0,52	2,2*	0,12	2,15
Francia (4;4)+/-	2,53	0,42	6,01***	1,71	3,37	2,71	0,5	5,44***	1,73	3,69
Holanda (2;1)	2,16	0,56	3,88***	1,07	3,25	2,34	0,92	2,55*	0,54	4,13
Israel (1;2)	1,14	0,72	1,6	-0,26	2,55	1,63	0,62	2,62**	0,41	2,85
Italia (4;3)	0,95	0,41	2,32*	0,15	1,74	0,87	0,54	1,61	-0,19	1,92
Japón (6;7)	1,76	0,32	5,47***	1,13	2,4	1,93	0,35	5,45***	1,24	2,63
Korea (3;3)	2,34	0,47	5,01***	1,43	3,26	1,51	0,53	2,85**	0,47	2,56
Nueva Zelanda (0;1)	---	---	---	---	---	3,22	0,94	3,43***	1,38	5,06
Países Bajos (2;2)	1,62	0,55	2,96**	0,55	2,69	1,16	0,64	1,8	-0,1	2,42
Polonia (1;1)	1,57	0,77	2,05*	0,07	3,07	1,52	0,91	1,66	-0,27	3,3
Suecia (1;2)	1,16	0,76	1,54	-0,32	2,64	1,91	0,66	2,89**	0,62	3,2
Suiza (1;1)+/+	5,01	0,82	6,13***	3,41	6,61	4,57	0,94	4,84***	2,72	6,41
Turquía (3;4)	0,54	0,44	1,23	-0,32	1,41	2,08	0,47	4,41***	1,15	3
UK (11;13)	1,6	0,23	6,87***	1,14	2,05	2,01	0,26	7,72***	1,5	2,52
Forma del STAI										
Forma Y (39;50)**/-	1,78	0,14	12,56***	1,51	2,06	2,01	0,14	14,46***	1,74	2,28
Forma X (53;61)	1,24	0,12	10,28***	1,01	1,48	1,71	0,13	13,66***	1,47	1,96
Trastorno psicológico										
Crisis de angustia (16;16)	1,54	0,23	6,67***	1,09	1,99	1,62	0,21	7,66***	1,2	2,03
Fobia específica (11;17) -/+	1,18	0,28	4,25***	0,64	1,73	1,05	0,21	5,06***	0,64	1,46
Fobia social (16;19) -/+	1,89	0,23	8,07***	1,43	2,35	2,33	0,2	11,56***	1,93	2,72
Fobia social generalizada (4;4)	1,32	0,46	2,9**	0,43	2,21	1,97	0,44	4,49***	1,11	2,83
TAG (2;12) -/+	0,79	0,63	1,25	-0,44	2,02	2,76	0,26	10,71***	2,26	3,27
TEPT (16;19)	1,59	0,23	6,99***	1,15	2,04	1,79	0,2	9,08***	1,4	2,17
TOC (26;23)	1,33	0,18	7,48***	0,98	1,68	1,84	0,18	10,22***	1,49	2,19
TEA (2;2)	1,25	0,63	2*	0,03	2,48	0,61	0,58	1,05	-0,53	1,75

Nota. *Z* = puntuación *Z* del coeficiente; I.C. = Intervalo de confianza al 95%; TAG = Trastorno de ansiedad generalizada; TEPT = Trastorno de estrés post-traumático; TOC = trastorno obsesivo-compulsivo; TEA = Trastorno por estrés agudo. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$. + la *d* de Cohen media para dicha categoría es estadísticamente diferente de la *d* de Cohen media total; - la *d* de Cohen media para dicha categoría no es estadísticamente diferente de la *d* de Cohen media total.

das únicamente se observó capacidad predictiva de la variabilidad en la media para la ansiedad estado ($b = 0,08$ [IC 95%: 0,06-0,11]; $Q_m = 44,01$, $p < 0,001$;

$Q_e = 507,65$, $p < 0,001$; $R^2 = 0,22$) y de la media ($b = 0,09$ [IC 95%: 0,08-0,11]; $Q_m = 106,54$, $p < 0,001$; $Q_e = 510,46$, $p < 0,001$; $R^2 = 0,5$) y la desviación típica en el caso de la ansiedad rasgo ($b = -0,10$ [IC 95%: -0,16--0,14]; $Q_m = 9,89$, $p = 0,002$; $Q_e = 827,42$, $p < 0,001$; $R^2 = 0,09$). El segundo tipo de meta-análisis incluido fue el de generalización de la fiabilidad. En este caso se obtuvieron 13 artículos para la escala de ansiedad estado y 19 artículos para la ansiedad rasgo, en los que se informa del alfa de Cronbach. De nuevo se comenzó por el análisis de la influencia de un posible sesgo de selección, observando problemas tanto en estado ($Z = 4,93$, $p < 0,001$) como en rasgo ($Z = 6,09$, $p < 0,001$). En los artículos de la ansiedad estado el alfa medio, calculado sin ningún tipo de ponderación fue de 0,87 ($DT = 0,08$); de 0,89 ($DT = 0,05$) para la ansiedad rasgo.

Seguidamente, se calculó la media de las puntuaciones T ponderadas, observando un valor de 0,43 en ansiedad estado ($EE = 0,006$) y de 0,45 ($EE = 0,005$) para la ansiedad rasgo. Al transformar esta puntuación de nuevo a valores alfa de Cronbach, se obtuvo un $\alpha = 0,92$ en ansiedad estado (IC 95%: 0,91-0,93) y de 0,91 en ansiedad rasgo (IC 95%: 0,90-0,92). En ambos casos, este valor resulta estadísticamente significativo (ansiedad estado: $Z = 71,24$, $p < 0,001$; ansiedad rasgo: $Z = 83,76$, $p < 0,001$). Tras ello, se analizó la homogeneidad de los valores alfa en los diferentes artículos. En ambos casos se observa heterogeneidad: ansiedad estado: $Q_{(12)} = 106,19$ ($p < 0,001$) y ansiedad rasgo: $Q_{(18)} = 70,86$ ($p < 0,001$).

El siguiente paso fue analizar las fuentes de variabilidad. En la ansiedad estado todas las variables analizadas resultaron significativas: país ($Q_{m(4)} = 65,09$; $p < 0,001$; $Q_{e(8)} = 41,1$; $p < 0,001$), forma del STAI ($Q_{m(1)} = 6,88$; $p = 0,009$; $Q_{e(11)} = 99,31$; $p <$

Tabla 2.
 Coeficientes de los moderadores en el meta-ANOVA de generalización de la fiabilidad para la ansiedad estado y rasgo.

Variables (<i>k</i> estado; <i>k</i> rasgo) significativo estado/significativo rasgo	Ansiedad estado					Ansiedad rasgo				
	α^1	Error estándar	Z	I.C. ¹		α^1	Error estándar	Z	I.C. ¹	
				Inferior	Superior				Inferior	Superior
País del estudio										
Alemania (-;3)	---	---	---	---	---	0,91	0,02	21,65***	0,89	0,94
Canadá (2;1)+/-	0,89	0,01	44,72***	0,87	0,9	0,94	0,01	41,59***	0,93	0,95
EEUU (7;10)+/-	0,94	0,01	47,5***	0,93	0,95	0,91	0,01	55,87***	0,9	0,92
Francia (-;2)	---	---	---	---	---	0,87	0,02	24,08***	0,83	0,9
Holanda (1;1) +/+	0,93	0,02	22,89***	0,91	0,95	0,81	0,03	22,89***	0,76	0,85
Italia (1;1)	0,83	0,04	13,17***	0,74	0,9	0,92	0,03	13,17***	0,88	0,95
Turquía (-;1)	---	---	---	---	---	0,92	0,03	14,7***	0,88	0,95
Forma del STAI										
Forma Y (7;13) +/-	0,9	0,03	15,33***	0,87	0,94	0,9	0,02	24,89***	0,88	0,93
Forma X (4;6)	0,93	0,04	11,24***	0,88	0,96	0,89	0,03	17,32***	0,85	0,92
Trastorno psicológico										
Fobia específica (0;3)	---	---	---	---	---	0,93	0,04	11,35***	0,88	0,96
Fobia social (0;2) -/+	---	---	---	---	---	0,9	0,05	10,49***	0,82	0,95
Fobia social generalizada (0;1)-/+	---	---	---	---	---	0,84	0,06	8,65***	0,7	0,93
TOC (2;5)	0,95	0,05	7,12***	0,9	0,98	0,9	0,03	16,1***	0,86	0,93
TEPT (6;5)+	0,9	0,03	15,05***	0,84	0,93	0,91	0,03	15,71***	0,88	0,95
TAG (2;1) -/+	0,91	0,05	8,3***	0,83	0,96	0,84	0,06	8,62***	0,7	0,93
Varios trastornos (1;2)-/+ ²	0,93	0,07	5,72***	0,83	0,98	0,88	0,04	11,91***	0,8	0,93

Nota. ¹ Los coeficientes han sido retransformados mediante $\alpha = 1 - T^3$. Z = puntuación Z del coeficiente; I.C. = Intervalo de confianza al 95%;
² Incluye artículos en los que se emplea varios trastornos juntos siempre que éstos sean exclusivamente trastornos de ansiedad; TOC =
 trastorno obsesivo-compulsivo; TEPT = Trastorno de estrés post-traumático; TAG = Trastorno de ansiedad generalizada. + el α medio para
 dicha categoría es estadísticamente diferente del α medio total; - el α medio para dicha categoría no es estadísticamente diferente del α medio
 total.

0,001) y tipo de trastorno ($Qm_{(3)} = 19,26$; $p = 0,001$; $Qe_{(9)} = 86,93$; $p < 0,001$). Mientras,
 en la ansiedad rasgo, el país ($Qm_{(6)} = 40,11$; $p < 0,001$; $Qe_{(12)} = 30,76$; $p = 0,002$) y el
 tipo de trastorno ($Qm_{(5)} = 33,68$; $p < 0,001$; $Qe_{(13)} = 37,18$; $p < 0,001$) resultan
 significativas pero no la forma del STAI ($Qm_{(1)} = 0,1$; $p = 0,751$; $Qe_{(17)} = 70,76$; $p <$

0,001). Un resumen de los estadísticos por cada nivel de las variables se recoge en la Tabla 2. Para los moderadores continuos, debido al bajo número de estudios los resultados resultaban significativos pese a lo cuál las betas eran inferiores a 0,01, salvo en el caso de la desviación típica como moderador de la variabilidad en la fiabilidad de la ansiedad rasgo ($b = -0,03$ [IC 95%: -0,4- -0,02]; $Q_m = 31,62$, $p < 0,001$; $Q_e = 38,55$, $p < 0,001$).

Discusión

Tras el análisis de datos se ha observado que, en ambos casos, ansiedad estado y rasgo, los tamaños del efecto medio son elevados. Así pues, el STAI es sensible a los niveles de ansiedad clínicos, detectando puntuaciones significativamente más elevadas que un grupo control. El segundo objetivo era el de establecer la fiabilidad media de la escala para pacientes diagnosticados de ansiedad. Nuevamente, los valores son elevados, observando una fiabilidad media superior a 0,9 en ambas subescalas. Por ello, se puede concluir que el STAI es un instrumento fiable a la hora de evaluar los niveles de ansiedad en pacientes con ansiedad y que es sensible a la hora de detectar los niveles superiores de ansiedad (estado y rasgo) de los mismos, a través de sus diferentes formas y en diversos países. Además de ello, el STAI es sensible a la ansiedad provocada por diversos trastornos ansiosos.

Al realizar el meta-análisis sobre el efecto medio se ha observado variabilidad de las medidas a lo largo de los estudios. A la hora de explicar la variación de los tamaños del efecto hay diferencias entre países. No obstante, los tamaños del efecto son adecuados (abarcando desde 0,44 a 5,01 en ansiedad estado y de 0,54 a 4,57 en ansiedad rasgo). Además, salvo en los artículos de Alemania, Italia y Turquía, todos los tamaños de efecto medios son superiores a 1. Así pues el STAI detecta diferencias entre ansiosos y control tal y cómo cabría esperar (Somers et al., 2006; Spielberger et al., 1970). En lo

que respecta a la forma del STAI, en ambos casos los tamaños del efecto son similares, pese a que la forma X tiene un tamaño del efecto medio menor que la forma Y, tanto en estado como en rasgo. Lo que indica que ambas formas son adecuadas para detectar los niveles de ansiedad, pudiendo utilizar el STAI para la evaluación de la ansiedad (Fisher y Durham, 1999; Spielberger y Reheier, 2009). El último moderador de la variabilidad empleado fue el tipo de trastorno, observando unos tamaños del efecto medios de 0,75 a 1,77, en el caso de la ansiedad estado y de 0,59 a 2,36 en el caso de la ansiedad rasgo. Así pues, el STAI permite detectar ansiedad general de forma adecuada, mostrando puntuaciones estadísticamente mayores entre diversos trastornos, respecto a población general.

La generalización de la fiabilidad permite inferir si un test es fiable a lo largo de muestras procedentes de diversas poblaciones (Espallargues i Carreras y Tebé Cordoní, 2006). En este caso, se observa que la fiabilidad media es elevada. Además, al analizar por países la fiabilidad media es adecuada. Al estimar la fiabilidad media por cada uno de los trastornos los valores superan el 0,86 en todos los casos. Estos resultados son coherentes ya que se encuentra un notable número de artículos en los que se afirma que el STAI tiene una elevada fiabilidad (Rossi y Pourtois, 2012) así como un meta-análisis en el que se establecen valores medios de alfa de Cronbach de 0,91 en ansiedad estado y de 0,89 en ansiedad rasgo (Barnes et al., 2007). Así pues en población general la fiabilidad es similar la obtenida, pese a que, en trastorno de ansiedad generaliza en el caso de la subescala de ansiedad rasgo el alfa sea ligeramente menor. Por último, al analizar la forma del STAI se observa que la forma Y tiene 0,91 para ambas escalas, mientras que la forma X tiene una fiabilidad de 0,93 para ansiedad estado y de 0,91 para ansiedad rasgo. Generalmente, se afirma que la versión Y del STAI tiene una mayor fiabilidad (Spielberger et al., 1983). El hecho de que en este meta-análisis se observe lo

contrario, puede deberse a que el tamaño de las muestras de sujetos con diagnóstico de ansiedad suele ser menor que el de las muestras de población general, lo que puede afectar a la fiabilidad.

En lo que respecta a España se observa que hay diversos artículos en los que se emplea muestra clínica, obteniendo un tamaño del efecto medio elevado. Así pues, la adaptación española del STAI detecta puntuaciones superiores en muestra ansiosa. Pese a ello, dicha adaptación no cuenta con una muestra de población ansiosa ni de baremos específicos para ellos (Buela-Casal, Guillén-Riquelme y Seisdedos Cubero, 2011). De la misma forma no hay fiabilidad para muestra ansiosa en la adaptación española. De hecho, en la revisión realizada, no se ha encontrado ningún artículo con muestra ansiosa española que informe de la fiabilidad. Pese a ello, en población general la fiabilidad del STAI es adecuada (Guillén-Riquelme y Buela-Casal., 2011).

En resumen, el STAI parece ser un instrumento fiable en muestras de sujetos ansiosos detectando mayores niveles de ansiedad en éstas respecto a muestras control. Por ello, sería conveniente que las diversas adaptaciones del STAI tuviesen baremos con muestra ansiosa. Como limitación del estudio cabe destacar la influencia de otros moderadores que expliquen la variabilidad hallada en los tamaños del efecto y los coeficientes de fiabilidad medios. Otra limitación es el hecho de que, pese a haber incluido búsquedas que permitiesen localizar la literatura gris (por ejemplo *Scholar Google*, en los resultados se observa sesgo de selección, por lo que sería interesante replicar este meta-análisis incluyendo nuevas bases y búsquedas. Pese a ello, el estudio resulta útil ya que permite conocer la fiabilidad media de un instrumento que es uno de los más utilizados en la práctica psicológica.

Referencias

- American Psychological Association (2002). *Manual diagnóstico y estadístico de los trastornos mentales. DSM-IV-TR*. Barcelona: Masson.
- Barnes, L. L. B., Harp, D. y Jung, W.S. (2002). Reliability generalization of scores on the Spielberger State-Trait anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618. doi: 10.1177/0013164402062004005
- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos Cubero, N. (2011). *Cuestionario de Ansiedad Estado-Rasgo: Adaptación española* (8ª ed.). Madrid: TEA Ediciones.
- Cohen, J. (1988). *Statistical power and analysis for the social sciences*. Hillsdale, NJ: Erlbaum.
- DerSimonian, R. y Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177-188. doi:10.1016/0197-2456(86)90046-2
- Espallargues i Carreras, M. y Tebé Cordoní, C. (2006). Qué resultados combinar y cómo combinarlos. En J. L. R. Martín, A. T. Garcés y T. S. Pillado (Coords.), *Revisiones sistemáticas en las ciencias de la vida* (pp. 173-187). Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-La Mancha.
- Fernández-Ríos, L. y Buela-Casal, G. (2009). Standars for the preparation and writing of Psychology review articles. *International Journal of Clinical and Health Psychology*, 9, 329-344.
- Fisher, P. L. y Durham, R. C. (1999). Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological Medicine*, 29, 1425-1434. doi:10.1017/S0033291799001336

- Guillén-Riquelme, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory. *Psicothema*, 23, 510-515.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M. y Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, 21, 169-184. doi: 10.1002/mpr.1359
- Kessler, R. C., Ruscio, A. M., Shear, K. y Wittchen, H.-U. (2009). Epidemiology of Anxiety Disorders. En M. B. Stein y T. Stecker (eds.), *Behavioral Neurobiology of Anxiety and Its Treatment*, [libro electrónico]. doi:10.1007/7854_2009_9
- Kvaal, K., Ulstein, I., Nordhus, I. H. y Engedal, K. (2005). The Spielberger State-Trait Anxiety Inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *International Journal of Geriatric Psychiatry*, 20, 629-634. doi:10.1002/gps.1330
- Montero, I. y León, O. G. (2007) A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862.
- Perestelo-Pérez, L. (2013). Standards on how to develop and report systematic reviews in Psychology and Health. *International Journal of Clinical and Health Psychology*, 13, 49-57.
- Rossi, V. y Pourtois, G. (2012). Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety, Stress, & Coping*, 25, 603-645. doi: 10.1080/10615806.2011.582948

- Sánchez-Meca, J. (2010). Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta*, 38, 53-64.
- Sánchez-Meca, J. y López-Pina, J. A. (2008). El enfoque meta-analítico de generalización de la fiabilidad. *Acción Psicológica*, 31, 7-17.
- Sánchez-Meca, J., Marín Martínez, F. y Huedo Medina, T. B. (2006). Modelos de efectos fijos y modelo de efectos aleatorios. En J. L. R. Martín, A. T. Garcés y T. S. Pillado (Coords.), *Revisiones sistemáticas en las ciencias de la vida* (pp. 189-204). Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-LaMancha.
- Somers, J. M., Goldner, E. M., Waraich, P. y Hsu, L. (2006). Prevalence and Incidence Studies of Anxiety Disorders: A Systematic Review of the Literature. *Canadian Journal of Psychiatry*, 51, 100-113.
- Spielberger, C. D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R. y Jacobs, G. A. (1983). *State-Trait Anxiety Inventory (Form Y)*. Palo Alto: Consulting Psychologists Press.
- Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- Urrútia, G. y Bonfill, X. (2013). La Declaración PRISMA: un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Revista Española de Salud Pública*, 87, 99-102. doi:10.4321/S1135-57272013000200001

Viechtbauer, W. (2010). metafor: Meta-Analysis Package for R. (1.4-0). Descargado de <http://cran.r-project.org/web/packages/metafor/index.html>.

Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E.J.,... Wells, J. E. (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet*, 370, 841-850. doi:10.1016/S0140-6736(07)61414-7

Estudio 2

Actualización psicométrica y funcionamiento diferencial de los ítems en el *State*

Trait Anxiety Inventory (STAI)

Alejandro Guillén-Riquelme y Gualberto Buela-Casal

Artículo publicado en *Psicothema*

Revista indexada en el *Journal Citation Reports*

Factor de impacto 2011 = 1,016.

Referencia:

Guillén-Riquelme, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory.

Psicothema, 23, 510-515.

Actualización psicométrica y funcionamiento diferencial de los ítems en el *State Trait Anxiety Inventory (STAI)*

Resumen

La ansiedad es uno de los problemas psicológicos con mayor prevalencia. Entre los instrumentos para medirla se encuentra el State Trait Anxiety Inventory. Este cuestionario mide ansiedad rasgo (factor de personalidad que predispone a sufrir o no ansiedad) y ansiedad estado (los factores ambientales que protegen o generan ansiedad). La adaptación española del cuestionario se realizó en 1982, por ello el objetivo del estudio es realizar una revisión del STAI. Para ello se reunió una muestra de 1.036 adultos. Se realizó un análisis de fiabilidad mediante alfa de Cronbach (0,90 para ansiedad rasgo y 0,94 para ansiedad estado). También se realizó una reducción factorial con unos resultados similares a los de la adaptación original. Además, se comprobó si existía funcionamiento diferencial de ítem por sexo y únicamente una de las 40 preguntas mostró problemas. Por último se realizó comparaciones t-student con los valores de la adaptación; mientras el rasgo varía en 1 punto, el estado tiene diferencias de hasta 6 puntos. En estos resultados de forma general se observa que el STAI mantiene unas adecuadas propiedades métricas y que, además, ha sido sensible al aumento de estímulos ambientales que producen estrés.

Abstract

One of the psychological problems with highest prevalence is anxiety. State Trait Anxiety Inventory is one of the instruments to measure it. This questionnaire assesses Trait Anxiety (understood as a personality factor that predisposes or not to suffer from

anxiety) and State Anxiety (refers to environment factors that protect or generate anxiety). The questionnaire was adapted in Spain in 1982 therefore the study's objective is to review the current psychometric properties of the STAI. A total of 1,036 adults took part in the study. Alpha's Cronbach reliability was .90 for Trait and .94 for State. The factorial analysis showed similar results compared with the original data. Moreover, Differential Item Functioning (DIF) was carried out exploring a sex bias. Only one of the 40 items showed DIF problems. Finally, a T-Test was run comparing the original and current values; while the Trait varied in 1 point, the State had differences of up to 6 points. In general, this result shows that the STAI has maintained adequate metric properties and has also been sensitive to increased environmental stimuli that produce stress.

En la actualidad, la ansiedad es uno de los problemas psicológicos con mayor prevalencia. Alonso et al. (2004b) indican que, en una muestra representativa de seis países europeos, un 13,6% había tenido algún problema de ansiedad a lo largo de su vida y un 6,4% en el último año; resultado similar al obtenido por Andlin-Sobocki y Wittchen (2005), nuevamente en europeos, donde un 12% presentaba trastornos de ansiedad. De la misma forma Wang et al. (2007) hallaron que un 6,8% de los encuestados españoles había tenido problemas por ansiedad y trastornos emocionales en el último año. Además, un 26,1% de las consultas psiquiátricas formales se deben a trastornos de ansiedad (Alonso et al., 2004c) y un 28,3% del total de casos de ansiedad muestran comorbilidad con otros trastornos (Alonso et al., 2004a).

Existen numerosos cuestionarios para medir la ansiedad (Inventario de Ansiedad de Beck, Inventario de Situación y Respuestas Ansiosas, Escala de Ansiedad Manifiesta de Taylor...). Entre los cuestionarios para evaluar este trastorno se encuentra el State-Trait Anxiety Inventory (STAI), desarrollado por Spielberger, Gorsuch y Lushene (1970). Este instrumento fue creado para evaluar la ansiedad según el modelo del propio Spielberger, que postulaba que dicho trastorno está constituido por dos componentes: un factor de personalidad que comprendería las diferencias individuales, relativamente estables, para responder ante situaciones percibidas como amenazantes con una subida en la ansiedad. También se define como una tendencia a percibir las situaciones como más amenazantes (ansiedad rasgo). El segundo factor (ansiedad estado) hace referencia a un periodo transitorio caracterizado por un sentimiento de tensión, aprensión y un aumento de la actividad del sistema nervioso autónomo, pudiendo variar tanto en el tiempo como en la intensidad (Spielberger et al., 1970). La elección de este cuestionario sobre el resto de pruebas para evaluar ansiedad se debe a que el STAI es uno de los más

utilizados.; de hecho Muñiz y Fernández-Hermida (2010) encontraron que es el 7º cuestionario más utilizado en España.

Cada una de las escalas del STAI (ansiedad rasgo y ansiedad estado) está compuesta por 20 ítems; una parte de los mismos redactada de forma positiva y otra de forma negativa. Diversos estudios mostraron reducciones factoriales muy robustas con cuatro factores. Estos factores estaban basados en la inversión de parte de las preguntas y se definían como ansiedad rasgo afirmativa y negativa y ansiedad estado afirmativa y negativa (Hishinuma, Miyamoto, Nishimura y Nahulu, 2000; Suzuki, Tsukamoto y Abe, 2000; Vigneau y Cormier, 2008). En la adaptación española (Spielberger, Gorsuch y Lushene, 1982) se encuentran resultados muy similares. El número de factores que subyacen en el STAI han sido objeto de numerosos análisis, existiendo defensores del modelo bifactorial, tetrafactorial y, en menor medida, un modelo pentafactorial, compuesto por los factores presentados anteriormente más un factor, compuesto de dos ítems, llamado happiness (Suzuki et al., 2000).

Este instrumento ha sido traducido a más de 40 lenguas (Spielberger y Díaz-Guerrero, 1976, 1983, 1986; Spielberger, Díaz-Guerrero y Strelau, 1990) y se han realizado más de 3.000 investigaciones que lo utilizan (Spielberger, 1989). En el 2002 Barnes, Harp y Jung, realizaron una revisión del STAI y encontraron 909 artículos (publicados desde el año 1990 hasta el 2000) que lo citaban. Así pues hasta el año 2000, se realizaron, al menos, 3.909 investigaciones con el cuestionario.

El STAI se adaptó al español a partir de los trabajos de Bermúdez (1978a; 1978b) y su versión comercial se realizó por la sección de estudios de TEA Ediciones (Spielberger et al., 1982). En estos artículos se comprueba que las medias y la fiabilidad (alfa de Cronbach y correlación test-retest) resultan similares a la versión

original. Respecto a la dimensionalidad de la escala, se mantiene el modelo de cuatro factores presentado anteriormente.

El objetivo del presente estudio es analizar las propiedades psicométricas del STAI, puesto que su adaptación española se realizó hace casi treinta años y comprobar su fiabilidad. Entre los análisis se incluirá una prueba del funcionamiento diferencial del ítem para analizar posibles sesgos, producidos por el efecto del sexo. Este análisis es pertinente puesto que no se ha realizado anteriormente en la adaptación española. Se espera que los índices psicométricos del cuestionario sean adecuados; pese a ello también es posible los valores normativos hayan variado respecto al original por el paso del tiempo.

Método

Participantes

La muestra está compuesta por 1.036 participantes (352 hombres y 667 mujeres). Las edades oscilan desde 18 hasta los 69 años. La edad media es de 29,64 (DT = 10,51). Las mujeres tienen una edad media de 29,83 años (DT = 10,31) y los hombres de 29,30 años (DT = 10,91).

Instrumentos

Para el estudio se utilizó el State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1982). La adaptación española fue realizada por la sección de estudios de TEA.

Este inventario se compone de 20 ítems para cada una de subescalas (ansiedad estado y ansiedad rasgo). La escala de respuesta es de tipo Likert, puntuando desde 0 (nada) hasta 3 (mucho). Los totales se obtienen sumando los valores de los ítems (tras las inversiones de las puntuaciones en los ítems negativos). Por ello, los totales de ansiedad rasgo y de ansiedad estado abarcan desde 0 hasta 60, correspondiéndose una mayor

puntuación con mayor ansiedad detectada. No cuenta con un criterio clínico en su versión española, aunque sí lo tiene en la versión original, así como en muchas adaptaciones.

Procedimiento

La muestra se recogió en Granada, Córdoba, Jaén, Madrid, Barcelona, Valencia, Ourense, Santiago de Compostela y Bilbao. El tipo de muestreo fue no probabilístico por cuotas (ser mayor de edad y no estar cursando una carrera universitaria). Para el muestreo se contó con la colaboración de organizaciones públicas y privadas. También se administró el cuestionario en lugares con una gran afluencia de personas como aeropuertos, estaciones, centros comerciales, etc. El cuestionario se administró por un único investigador. Se daba el test a los sujetos y era completado en presencia del administrador para posibles dudas o problemas. En todos los casos se informó a los participantes de la confidencialidad del estudio.

Diseño

Según la clasificación propuesta por Montero y León (2007) el presente estudio sería instrumental, puesto que evalúa las propiedades psicométricas de un cuestionario. Se siguieron las recomendaciones metodológicas de Carretero-Dios y Pérez (2007).

Análisis estadísticos

El análisis de la fiabilidad se empleó el *alfa* de Cronbach. La comprobación de la estructura de la escala se realizó mediante análisis factorial. El método de extracción ha sido *Máxima Verosimilitud* y la rotación *Varimax*. Para establecer el número de factores se realizó un análisis paralelo al 95% de confianza.

Para comparar los datos del estudio con los de la adaptación española se utilizó la prueba *t-student* para una sola muestra, complementados con el tamaño del efecto mediante la *d* de Cohen.

Se comprobó si había Funcionamiento Diferencial del Ítem (DIF en adelante). El DIF se define como la diferencia en la probabilidad que tienen los participantes de dos o más grupos, en este caso sexo, de responder de forma similar, tras haber sido igualados en habilidad (Millsap y Meredith, 1992). Según Mellenbergh (1995) se distingue entre DIF uniforme (la probabilidad de responder correctamente al ítem es mayor para uno de los grupos uniformemente para todos los niveles de la variable) y DIF no uniforme (la diferencia de probabilidad para responder correctamente al ítem no es igual lo largo de los diferentes niveles del atributo).

Para la detección del DIF existen varios procedimientos. En el presente artículo se usará el análisis de regresión logística, propuesto por Miller y Spray (1993). Esta técnica permite detectar el DIF uniforme y no uniforme en ítems politómicos (Hidalgo y López-Pina, 2004). Los análisis tal como se presentan en Hidalgo y Gómez (2006) consisten en un modelo de regresión logística binaria, donde la variable dependiente es el grupo (en este caso sexo) y las variables independientes son la puntuación total en el test, la respuesta al ítem y la interacción entre ambas. La estrategia para detectar DIF, siguiendo las indicaciones de Hidalgo, Gómez y Padilla (2005), consiste en ajustar tres modelos en distintas etapas. En la primera etapa se introduce en la ecuación la puntuación total del sujeto en el test, ajustando así el Modelo 1 (M.1) en base a la ausencia de DIF. En la segunda etapa, se añade la variable de agrupamiento (Modelo 2, M.2). Si la explicación de este modelo respecto al M1 fuese significativa sería indicio de DIF uniforme. En la etapa 3 se introduce la interacción entre el grupo y la puntuación

total (Modelo 3, M.3). Si la explicación que añade este modelo, respecto a los anteriores, fuese significativa habría DIF no uniforme.

Una vez detectados los ítems con DIF se elimina sus puntuaciones del total del cuestionario. Tras la purificación se repite los análisis en todos los ítems con el total corregido. Si aparecen nuevos ítems con DIF se puede proceder a una segunda depuración del total.

Por último, en los ítems con DIF cabe averiguar si se mantiene en todas las escalas de respuesta. Para ello se utiliza el análisis de categorías acumuladas propuesto Mellenbergh (1995). El análisis consiste en varias dicotomizaciones de las categorías de respuesta. Se repite el análisis de DIF con las diferentes agrupaciones realizadas. En este caso se compara el 0 con el 1, 2 y 3 agrupados, el 0 y 1 contra el 2 y 3, y el 0, 1 y 2 contra el 3.

Para la realización de todos los análisis descritos se utilizó el programa estadístico SPSS 15.0, excepto para el cálculo del tamaño del efecto mediante la *d* de Cohen, realizado con el G*Power 3.1.2.

Resultados

Primeramente se realizó un análisis de las puntuaciones medias y desviación típica de los ítems, para detectar puntuaciones extremas. Ninguna de las medias supera 1,78, siendo la menor de 0,45. Este valor, pese a ser bajo, no es de extrañar puesto que no se trata de una muestra clínica, no esperando por ello puntuaciones elevadas. Las desviaciones típicas se sitúan entre 0,71 y 1,11. Esto indica que los sujetos no tienen un patrón de respuesta con altas frecuencias en los valores centrales o en los extremos.

Seguidamente se realizó un análisis de fiabilidad, mediante el alfa de Cronbach, obteniendo, para los ítems de ansiedad rasgo, un alfa de 0,90. Solo uno de los 20 ítems

de la escala mejoraba el alfa (en una milésima) si era eliminado. Por su parte la ansiedad estado alcanzó un alfa de 0,94, con un solo ítem cuya eliminación mejoraba el alfa en cuatro milésimas.

Una vez establecido el funcionamiento adecuado de los ítems se pasó a la reducción factorial de los mismos para ver la estructura factorial subyacente. En primer lugar la prueba *Kaiser-Meyer-Olkin* ($KMO = 0,96$) así como la prueba de esfericidad de Bartlett ($\chi^2; 780 = 19,37; p < 0,001$) mostraron la adecuación de la muestra para realizar el análisis factorial. Para la elección del número de factores se utilizó un análisis paralelo. Se extrajeron cuatro factores. Sus autovalores son 14,40, 2,60, 2,48 y 1,33 y la varianza total explicada conjuntamente es de 52,04%. Las saturaciones de cada ítem en los diferentes factores se presentan en la tabla 1.

En las saturaciones se observa que el factor 3 corresponde a los ítems de la ansiedad rasgo redactados de forma positiva, mientras que los ítems que saturan en el factor 2 pertenecen a la ansiedad estado redactados de forma positiva. Los otros dos factores, especialmente el primero, aglutinan ítems de ambas subescalas cuya redacción original es negativa y sus puntuaciones se han invertido. Esta estructura es similar a hallada en la adaptación original de cuatro factores (ansiedad estado y rasgo, positiva y negativa respectivamente). El hecho de establecer cuatro factores en función de la presencia o ausencia de ansiedad estado y rasgo se puede considerar un artefacto estadístico por lo que se interpretará el resto de análisis en función de ansiedad estado y rasgo.

En siguiente lugar se analizaron las medias del total de las subescalas. Se calcularon pruebas t-student sobre una media, utilizando para el contraste los valores del manual de la adaptación española. Este análisis se realizó para ambos sexos, ya que son los valores que ofrece la adaptación.

Tabla 1.
Saturación de los ítems en los factores.

	Factor 1	Factor 2	Factor 3	Factor 4
Rasgo 1	0,31			0,66
Rasgo 2	0,32			
Rasgo 3	0,47			
Rasgo 4	0,42			0,32
Rasgo 5	0,4			
Rasgo 6				0,44
Rasgo 7				
Rasgo 8	0,45			0,3
Rasgo 9	0,64			
Rasgo 10				0,7
Rasgo 11	0,46			
Rasgo 12	0,59	0,51		
Rasgo 13	0,38	0,61		0,37
Rasgo 14	0,34			
Rasgo 15	0,51			0,4
Rasgo 16				0,68
Rasgo 17	0,59			
Rasgo 18	0,6			
Rasgo 19				0,39
Rasgo 20	0,55			
Estado 1			0,59	0,4
Estado 2		0,56	0,35	0,54
Estado 3			0,75	
Estado 4			0,6	
Estado 5			0,45	0,63
Estado 6			0,79	
Estado 7	0,37		0,38	
Estado 8			0,37	0,48
Estado 9			0,6	0,3
Estado 10			0,33	0,63
Estado 11		0,65		0,43
Estado 12			0,74	
Estado 13			0,33	
Estado 14			0,49	
Estado 15			0,54	0,52
Estado 16				0,72
Estado 17			0,58	0,34
Estado 18			0,7	
Estado 19				0,74
Estado 20			0,37	0,67

Nota. Se han eliminado las saturaciones menores a 0,3.

En la tabla 2 se puede observar, además de que todos los valores son significativos, que en las puntuaciones rasgo las diferencias son cercanas a 1,5 puntos, mientras que en la ansiedad estado las diferencias están en torno a los 5 puntos. Los tamaños del efecto son muy bajos en rasgo y elevados en la ansiedad estado. Parece ser que el rasgo no ha cambiado apenas respecto a los valores originales, mientras que la ansiedad estado sí ha sufrido fuertes modificaciones.

Tabla 2.
Diferencias de medias respecto a la adaptación original.

Sexo	Factor	Valor de prueba ¹	Media	Diferencia	<i>t</i>	<i>g.l.</i>	<i>p</i>
Hombres	Rasgo	20,19	18,96	1,23	-2,24	329	0,026
	Estado	20,54	15,87	6,67	-8,48	322	0,000
Mujeres	Rasgo	24,99	23,35	1,64	-3,85	617	0,000
	Estado	23,30	18,20	5,10	-10,58	604	0,000

¹ Media de cada grupo. Obtenidos del manual de la adaptación española. *t* = t-student; *g.l.* = grados de libertad; *p* = probabilidad.

En la adaptación española se encontraron diferencias por sexo (Spielberger et al., 1982). Para comprobar si estas diferencias se mantienen en ambas subescalas se utilizó la prueba t-student. En la ansiedad rasgo (asumiendo varianzas iguales; $F = 2,08$, $p = 0,150$) hay diferencias estadísticamente significativas ($t(946) = -6,20$; $p < 0,001$). Para la ansiedad estado (asumiendo varianzas distintas; $F = 17,60$, $p < 0,001$) también se encuentran diferencias por sexo ($t(764,21) = -3,19$; $p = 0,001$).

Para aportar evidencias de que las diferencias encontradas anteriormente no se deben a un problema de DIF se realizó un análisis del mismo. Los resultados se muestran en la tabla 3.

En la tabla 3 se observa que no se detecta DIF en ninguno de los ítems de ansiedad estado (significación menor a 0,01). Sin embargo, en la escala de ansiedad rasgo hay cuatro ítems (3, 5, 7 y 19) cuyo modelo 2 es significativo al 0,01, por lo que es un primer indicador de la presencia de DIF uniforme. Ningún ítem de ambas subescalas muestra DIF no uniforme. El siguiente paso es comprobar el tamaño del efecto mediante las diferencias en el R^2 del modelo 2 respecto al 1. Siguiendo los criterios propuestos por Jodoin y Gierl (2001), un valor inferior a 0,035 informaría de un DIF inapreciable; de 0,035 a 0,07 un DIF moderado, y mayor a 0,07 correspondería a un DIF elevado. En este caso el ítem 19 mostraría DIF moderado, mientras que, en los otros tres ítems, sería inapreciable. Seguidamente se depuró el total, eliminando estos cuatro ítems para repetir los análisis.

En la segunda fase se realizó de nuevo el procedimiento con el total depurado. Nuevamente los ítems 3, 5, 7 y 19 fueron significativos al 0,01, sin embargo los tamaños del efecto (excluyendo el del ítem 3, con un 0,061) son muy bajos, siendo el mayor de ellos de 0,026. Además de estos ítems, el 13, 14 y el 20 también fueron significativos, no obstante los tamaños del efecto son de 0,012, 0,010 y 0,010 respectivamente (inapreciables). Por todo ello se puede concluir que únicamente el ítem 3 muestra un DIF medio.

Tabla 3.

Probabilidades asociadas a los modelos (M) de la regresión y tamaño del efecto.

	Ítem	χ^2		χ^2		R^2	
		(M2-M1)	<i>p</i>	(M3-M2)	<i>p</i>	Nagelkerke (M3-M2)	Nagelkerke (M2-M1)
Rasgo	1	4,07	0,131	0,09	0,955	0	0,006
	2	0,57	0,753	1,24	0,538	0,002	0,001
	3	45,1	0	6,3	0,043	0,009	0,061
	4	0,04	0,982	0,9	0,638	0,001	0
	5	20,95	0	0,07	0,967	0	0,029
	6	0	1	1,79	0,408	0,003	0
	7	14,14	0,001	0,05	0,977	0	0,02
	8	1,7	0,428	0,29	0,867	0,001	0,002
	9	2,42	0,298	0,03	0,985	0	0,003
	10	5,46	0,065	0,42	0,809	0	0,008
	11	3,01	0,222	0,74	0,692	0,001	0,004
	12	1,76	0,415	4,69	0,096	0,007	0,002
	13	8,29	0,016	0,05	0,978	0	0,012
	14	6,87	0,032	5,14	0,077	0,007	0,01
	15	4,85	0,089	2,58	0,275	0,003	0,007
	16	5,8	0,055	0,02	0,993	0	0,008
	17	0,16	0,925	1,19	0,552	0,002	0
	18	1,07	0,587	2,51	0,285	0	0,001
	19	12,4	0,002	0,05	0,974	0	0,017
	20	6,86	0,032	1,78	0,411	0,002	0,01
Estado	1	1,14	0,565	1,74	0,42	0,003	0,001
	2	1,08	0,582	4,94	0,085	0,008	0,001
	3	0,18	0,914	5,08	0,079	0,007	0
	4	6,57	0,037	2	0,368	0,003	0,009
	5	8,88	0,012	1,68	0,433	0,002	0,013
	6	1,4	0,497	2	0,368	0,003	0,002
	7	0,05	0,973	1,35	0,51	0,002	0
	8	2,88	0,237	0,83	0,661	0,001	0,004
	9	0,39	0,823	2,23	0,327	0,004	0
	10	0,02	0,988	1,23	0,541	0,002	0
	11	7,53	0,023	2,1	0,35	0,003	0,011
	12	3,77	0,152	3,9	0,142	0,006	0,005
	13	1,82	0,404	3,14	0,208	0,005	0,002
	14	0,18	0,912	2,54	0,282	0,004	0
	15	4,29	0,117	2,47	0,29	0,004	0,006
	16	1,38	0,501	0	1	0	0,002
	17	0,85	0,652	0,76	0,685	0,001	0,001
	18	1,14	0,565	4,1	0,129	0,006	0,001
	19	0,19	0,911	0,7	0,703	0,001	0
	20	3,1	0,212	4,24	0,12	0,007	0,004

Nota. *p* = probabilidad.

Una vez se detecta el DIF, es necesario analizar si se encuentra en todas las categorías de respuesta o solo en algunas. Para ello, la escala del ítem 3 (único con DIF) se dicotomizó mediante de las agrupaciones de Mellenbergh (1995). Con las categorías dicotomizadas se realiza los mismos análisis que anteriormente con los ítems. En ninguna de las alternativas de respuesta se detecta DIF no uniforme. Sin embargo, se observa DIF uniforme al comparar la opción cero contra la uno, dos y tres de forma conjunta. Por tanto, en el ítem 3 únicamente existe DIF al comparar la primera opción (“casi nunca”) con las otras tres. Es decir que hombres y mujeres, igualados en nivel de ansiedad, responden con diferente frecuencia en esta opción.

Discusión

En los resultados se ha comprobado que las medias y desviaciones típicas de los ítems son adecuadas (Carretero-Dios y Pérez, 2007). Si bien algunas de las medias son bajas, esto se debe a que es un cuestionario de ansiedad (en su versión original y en muchas adaptaciones, tiene baremos clínicos) administrado en población normal. Por ello las medias algo bajas son esperables. Respecto a la fiabilidad los valores alfa de Cronbach son adecuados dando indicios de una adecuada consistencia interna.

El análisis factorial muestra un porcentaje de varianza explicada superior al 50%. Además, el análisis paralelo indica que se debe extraer cuatro factores, al igual que en la adaptación española (Spielberger et al., 1982). Independientemente, Vera-Villarroel, Buela-Casal y Spielberger (2007) encuentran cuatro factores al realizar dos extracciones separadamente (una para rasgo y otra para estado), con una versión en español del cuestionario. Cabe indicar que estos factores no replican exactamente los esperados teóricamente, encontrando problemas en los factores que incluyen los ítems invertidos de ambas subescalas. Si bien estos datos muestran cierta confusión son muy similares a los obtenidos por Bermúdez (1978a, 1978b), así como a los del manual de la

adaptación española (Spielberger et al., 1982). Pese a que los datos no son tan claros como en otras revisiones de la versión en inglés, muestran una estructura tetrafactorial similar a las anteriormente halladas, aportando un nuevo indicio de que el test conserva sus propiedades métricas.

Respecto a la comparación de medias respecto al original los resultados son muy interesantes. En primer lugar cabe indicar que los baremos y medias de la adaptación española se realizó con 318 varones y 365 mujeres trabajadores. El manual no ofrece más datos al respecto. En base a ello, se observa que los tamaños muestrales son similares y no hay características diferentes en la presente muestra. Al comparar puntuaciones entre ambas muestras las medias difieren un máximo de 1,64 puntos en la ansiedad rasgo, mientras que alcanzan entre 5 y 6 puntos de diferencia en la ansiedad estado. Este hecho puede ser indicio de que el cuestionario conserva unas buenas propiedades métricas pese al paso del tiempo, ya que la escala de ansiedad rasgo mantiene las puntuaciones (los factores de personalidad son relativamente estables). Al contrario, en los factores ambientales, la variación ha sido notable ya que éstos cambian con más facilidad a lo largo del tiempo (Eysenck, 1995).

En el análisis de DIF se encuentra un nuevo indicio sobre la validez del cuestionario. Se observa que únicamente un ítem muestra funcionamiento diferencial por sexo. Para comprender mejor la causa del sesgo en la respuesta de hombres y mujeres cabe realizar un análisis del contenido del ítem. La cuestión planteada es “Siento ganas de llorar”, en este caso parece lógico pensar que, pese a que se haya igualado hombres y mujeres en nivel de ansiedad, los hombres serán más propensos a puntuar que casi nunca sienten ganas de llorar. Mientras, las mujeres tendrán una respuesta más equilibrada en las diferentes alternativas, lo que explica el DIF hallado.

En conclusión, que la escala se mantenga estable en la medida de la ansiedad rasgo pero haya variado en la ansiedad estado, además de mantener las medidas psicométricas y no verse afectado por DIF son buenos indicios para considerar adecuada la utilización de la escala. No obstante, cabe indicar que el estudio cuenta con la limitación de ser un muestreo incidental lo que puede influir en las medidas. Para investigaciones futuras se debería replicar el experimento en varias muestras e incluir medidas de la validez de contenido, criterio y predictiva. Además, habría que rehacer los baremos, especialmente para la subescala de ansiedad estado.

Referencias

- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004a). 12-Month comorbidity patterns and associated factors in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 28–37. doi:10.1600-0047.2004.00328.x
- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004b). Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 21–27. doi:10.1111/j.1600-0047.2004.00326
- Alonso, J., Angermeyer, M. C., Bernert, S., Bruffaerts, R., Brugha, T. S., Bryson, H.,...
Vollebergh, W. A. M. (2004c). Use of mental health services in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109(Suppl. 420), 47–54. doi:10.1111/j.1600-0047.2004.00330.x
- Andlin-Sobocki, P. y Wittchen, H. U. (2005). Cost of anxiety disorders in Europe. *European Journal of Neurology*, 12(Suppl. 1), 39–44. doi: 10.1111/j.1468-1331.2005.01196.x
- Barnes, L., Harp, D. y Jung, W. S. (2002). Reliability Generalization of Scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618. doi: 10.1177/0013164402062004005
- Bermúdez, J. (1978a). Anxiety and performance. *Revista de Psicología General y Aplicada*, 151, 183-207.

- Bermúdez, J. (1978b). Functional analysis of anxiety. *Revista de Psicología General y Aplicada*, 153, 617-634.
- Carretero-Dios, H. y Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology*, 7, 863-882.
- Eysenk, H. J. (1995). Un modelo de personalidad: rasgos generales. En M. D. Avia y M. L. Sánchez-Bernados (Comps.), *Personalidad: aspectos cognitivos y sociales*. Madrid: Pirámide.
- Hidalgo, M. D. y Gómez, J. (2006). Nonuniform DIF detection using discriminant logistic analysis and multinomial logistic regression: A comparison for polytomous items. *Quality & Quantity*, 40, 805-823. doi:10.1007/s11135-005-3964-2
- Hidalgo, M. D., Gómez, J. y Padilla, J. L. (2005). Regresión logística: alternativas de análisis en la detección del funcionamiento diferencial del ítem. *Psicothema*, 17, 509-515.
- Hidalgo, M. D. y López-Pina, J. A. (2004). Differential Item Functioning Detection and Effect Size: a Comparison between Logistic Regression and Mantel-Haenszel Procedures. *Educational and Psychological Measurement*, 64, 903-915. doi:10.1177/0013164403261769
- Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T. y Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory Scores for Ethnically Diverse Adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83. doi:10.1037/1099-9809.6.1.73

Jodoin, M. G. y Gierl, M. J. (2001). Evaluating Type I Error and Power Rates Using an Effect Size Measure with the Logistic Regression Procedure for DIF Detection. *Applied Measurement in Education, 14*, 329–349.

doi:10.1207/S15324818AME1404_2

Mellenbergh, G. J. (1995). Conceptual Notes on Models for Discrete Polytomous Item Responses. *Applied Psychological Measurement, 19*, 91-100.

doi:10.1177/014662169501900110

Miller, T. R. y Spray, J. A. (1993). Logistic Discriminant Function Analysis for DIF Identification of Polytomously Scored Items. *Journal of Educational Measurement, 30*, 107-122. doi:10.1111/j.1745-3984.1993.tb01069.x

Millsap, R. E. y Meredith, W. (1992). Inferential Conditions in the Statistical Detection of Measurement Bias. *Applied Psychological Measurement, 16*, 389-402.

doi:10.1177/014662169201600411

Montero, I. y León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology, 7*, 847-862.

Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo, 31*, 108-121.

Spielberger, C. D. (1989). *State-Trait Anxiety Inventory: A comprehensive bibliography* (2ºed.). Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D. y Díaz-Guerrero, R. (Eds.). (1976). *Cross-cultural anxiety*. Washington, DC: Hemisphere/Wiley.

Spielberger, C. D. y Díaz-Guerrero, R. (Eds.). (1983). *Cross-cultural anxiety* (Vol. 2). Washington, DC: Hemisphere/Wiley.

- Spielberger, C. D. y Díaz-Guerrero, R. (Eds.). (1986). *Cross-cultural anxiety* (Vol. 3). Washington, DC: Hemisphere/Wiley.
- Spielberger, C. D., Díaz-Guerrero, R. y Strelau, J. (Eds.). (1990). *Cross-cultural anxiety* (Vol. 4). Washington, DC: Hemisphere/Wiley.
- Spielberger, C. D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., Gorsuch, R. L. y Lushene, R. (1982). *Manual del Cuestionario de Ansiedad Estado/Rasgo (STAI)*. Madrid, España: TEA Ediciones S.A.
- Suzuki, T., Tsukamoto, K. y Abe, K. (2000). Characteristic Factor Structures of the Japanese Version of the State-Trait Anxiety Inventory: Coexistence of Positive-Negative and State-Trait Factor Structures. *Journal of Personality Assessment*, 74, 447-458. doi:10.1207/S15327752JPA7403_8
- Vera-Villarroel, P., Buela-Casal, G. y Spielberger, C. D. (2007). Preliminary Analysis and Normative Data of the State-Trait Anxiety Inventory (STAI) in Adolescent and Adults of Santiago, Chile. *Terapia Psicológica*, 25, 155-162. doi:10.4067/S0718-48082007000200006
- Vigneau, F. y Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment*, 90, 280-285. doi:10.1080/00223890701885027
- Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E. J.,... Wells, J. E. (2007). Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet*, 370, 841-850. doi:10.1016/S0140-6736(07)61414-7

OBJETIVO 2
CREACIÓN Y AJUSTE DE UNA
VERSIÓN BREVE DEL STATE-TRAIT
ANXIETY INVENTORY EN
MUESTRA ESPAÑOLA

Estudio 3

Short form of the Spanish adaptation of the State-Trait Anxiety Inventory

Gualberto Buela-Casal y Alejandro Guillén-Riquelme

Manuscrito en proceso de revisión en *Assessment*

Short form of the Spanish adaptation of the State-Trait Anxiety Inventory

Abstract

The State-Trait Anxiety Inventory (STAI) is one of the assessment instruments that are most widely used by psychologists. Although several short forms of the STAI have been developed since its creation, none are available for the Spanish general population. The aim of the present study was to develop and validate a short form of the STAI. To achieve this, we applied the full STAI to 1,157 adults. We created high and low anxiety groups based on participants' scores on the STAI and conducted a discriminant analysis using such groups. We obtained a selection of five items for state anxiety and five items for trait anxiety and compared it to other short forms through a confirmatory factor analysis. The short form obtained with the discriminant analysis showed good fit. However, a selection of 12 items (six for each subscale) showed better fit. More specifically, items 2, 4, 11, 15, 17 and 18 (for state anxiety) and items 7, 14, 15, 16, 17 and 18 (for trait anxiety) were found to be the most adequate items to develop a short form for Spanish samples. In conclusion, these 12 items can be used to facilitate the application of very long assessments or to assess individuals with health problems who cannot perform very long assessments.

Key words: STAI, State-Trait Anxiety Inventory, anxiety, short version

The *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch & Lushene, 1970) is a questionnaire aimed at assessing the level of anxiety of individuals, as anxiety is one of the most prevalent psychological disorders in the general population (Beutel, Bleichner, von Heumann, Tritt & Hardt, 2011). More specifically, the STAI measures individuals' predisposition to perceive diverse stimuli as threatening (i.e. trait anxiety). Thus, an individual with high trait anxiety is more likely to produce an anxiety response. The STAI also assesses the existence of threat-related stimuli in the environment of the respondent around the time of the assessment (state anxiety). This is equivalent to a measure of negative affect (Watson, Clark & Stasik, 2011). Trait anxiety and state anxiety are theoretically independent constructs (Spielberger et al., 1970). Its two subscales include items that refer to behavioral, cognitive, emotional and physiological aspects.

Since its creation, “the STAI has been adapted to sixty different languages and dialects and has been cited in over 14,000 studies” (Spielberger & Reheiser, 2009, p. 280), which makes it one of the questionnaires most widely used. The fact that various systematic reviews and meta-analyses report that this instrument is used in a large number of articles provides further evidence of its popularity. In a meta-analysis on the effect of relaxation on anxiety levels, for example, Manzoni, Pagnini, Castelnuovo and Molinari (2008) mentioned that the STAI was used in 60% of the articles reviewed. This makes it one of the measures most used in this context. The STAI has also been used to study the psychological health of workers (Marine, Ruotsalainen & Verbeek, 2009), showing differences in the level of anxiety experienced by various groups. Reviews have revealed that the STAI is not only used in general population samples but also in clinical groups. For example, Rosa-Alcázar, Sánchez-Meca, Gómez-Conesa and Marín-Martínez (2008) conducted a meta-analysis of obsessive-compulsive disorder.

They found that the STAI was the questionnaire used in the largest number of studies analyzed.

Nevertheless, the STAI has been criticized for various reasons. First, it was pointed out that the trait anxiety subscale had more items reporting the presence of anxiety than its absence. It was also argued that its factor structure was not totally stable (Spielberger, 1983). This led to a new revised version (i.e. *form Y*) in which these shortcomings were corrected. Yet, the main criticism of the STAI is that it does not assess pure anxiety, since it includes items related to depression or negative affect (Endler, Cox, Parker & Bagby, 1992). This statement is based on the fact that the STAI shows a high correlation with several measures of depression (Spielberger & Reheiser, 2009). The study of the possible relationships between the scores of the STAI and other constructs, particularly depression, has led to alternative analyses that include new factors. For example, Caci, Bayle, Dossios, Robert and Boyer (2003) have proposed several factor models for the structure of the trait anxiety subscale. They argue that several items of the STAI in fact assess “restlessness,” “worrying,” “self-confidence,” “unsuccessfulness” and “anhedonia”. For this reason, they use 10 items that assess three of these factors (items 2, 20, 9, 11, 17, 18, 7, 12, 14 and 15), considering that trait anxiety is a second-order factor. Bieling, Antony and Swinson (1998) proposed a factor structure of the trait subscale of the STAI that includes a factor of depression or general negative affect and showed good fit in the confirmatory factor analysis. These authors concluded that items 2, 8, 9, 11, 17, 18 and 20 measure trait anxiety in its pure form.

Several authors have focused on developing a short form of the STAI or of one of its subscales. Although several short forms had already been produced, the first study to do so following certain methodological guidelines was that of van Knippenberg, Duivenvoorden, Bonke and Passchier (1990). These authors shortened the Dutch

version of both subscales to a form with six items for each subscale. More specifically, they obtained three different item selections (depending on the sample used) and verified which of such structures was the most appropriate. They concluded that they were all valid. Two years later, Marteau and Bekker (1992) shortened the English version of state anxiety and concluded that the form including items 1, 3, 6, 15, 16 and 17 was the most appropriate. However, only a sample of pregnant women was used in that study. This structure was later tested using other samples. In fact, van de Bij, de Weerd, Cikot, Steegers and Braspenning (2003) studied whether it was possible to use the short form proposed by Marteau and Bekker (1992) in a Dutch sample. Again, they used a sample of pregnant women but included their partners as well. They concluded that the item selection was valid. Davey, Barrant, Butow and Deeks (2007) also attempted to confirm this short form. They compared the short form proposed by Marteau and Bekker (1992) to one of the eight-item versions proposed by van Knippenberg et al. (1990). According to the results obtained, Davey et al. (1992) concluded that both short forms have good properties and can therefore be used.

When the STAI is applied to certain populations, it is very difficult for respondents to complete 20 items. An example of this is the assessment of the anxiety levels of patients with respiratory problems. With this in mind, Chlan, Savik and Weinert (2003) developed a six-item form (5, 9, 10, 12, 17 and 20) of the state subscale of the STAI. This form showed adequate properties in such sample (Chen, 2006; Chlan et al., 2003).

All the short forms presented so far except for that developed by van Knippenberg et al. (1990) were based on *form Y* of the STAI. By contrast, much fewer short versions are based on *form X*. The first was proposed by van Knippenberg et al. (1990), who developed a short form of both subscales. There are also two short forms of

the Brazilian adaptation of the STAI that include both subscales. Indeed, Fioravanti-Bastos, Cheniaux and Landeira-Fernandez (2011) developed a six-item form of each subscale. They added items of *form Y* to the original questionnaire and chose the best items to broaden this selection. In that case, it was possible to determine the items belonging to *form X*. In addition, Kaipper, Chachamovich, Hidalgo, Torres and Caumo (2010) proposed a form with 13 items for state anxiety and 12 items for trait anxiety.

In Spain, the STAI is one of the instruments most widely used by psychologists. In fact, it is the seventh most broadly used by clinical psychologists (Muñiz & Fernández-Hermida, 2010) and has shown adequate psychometric properties in a normal population (Guillén-Riquelme & Buela-Casal, 2011). Yet, no short forms of this questionnaire have been validated in a normal population. The version proposed by Chlan et al. (2003) has been validated in patients with respiratory intubation. This Spanish adaptation has shown good reliability and validity (Perpiñá-Galvañ, Richart-Martínez & Cabañero-Martínez, 2011) and evidence of content validity (Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez & Martínez-Durá, 2011). Considering all this, the objective of the present research was to develop a short form of both scales of the STAI for the Spanish general population. Another goal was to compare the short form obtained with those developed by other authors to determine which short version is best adapted to the Spanish general population.

Method

Participants

The sample was composed of 1,157 adults. It included geographic representation from the north, center and south of Spain (i.e. regions of Galicia, Basque Country, Cataluña, Madrid, Valencia and Andalusia). The mean age of participants was 29.73 years ($SD = 10.55$) and age ranged from 18 to 69 years. The sex distribution was 448

men and 693 women (17 participants did not provide this information). The mean age of men was 30.33 years ($SD = 11.56$), with the same age range as the total sample. The mean age of women was 29.39 years ($SD = 9.85$), with ages ranging from 18 to 63 years.

Instrument

To achieve the objectives of the study, participants responded to socio-demographic questions on their age, sex and place of residence. In addition, we administered the Spanish adaptation of the STAI (Buela-Casal, Guillén-Riquelme & Seisdedos Cubero, 2011). This version includes 40 items (i.e. 20 for each subscale). In the Spanish adaptation, the response scale ranges from 0 to 3 points, by contrast with the original STAI, in which it ranges from 1 to 4 points (Spielberger et al., 1970). In the state anxiety subscale, item scores range between 0 = *nada* [not at all], 1 = *algo* [somewhat], 2 = *bastante* [moderately] and 3 = *mucho* [very much]. In the trait anxiety subscale, response options range between 0 = *casi nunca* [almost never], 1 = *a veces* [sometimes], 2 = *a menudo* [often] and 3 = *casi siempre* [almost always]. Although each subscale has a theoretical range from 20 to 80, the scores can be compared to those of the original scale by adding 20 to the scores obtained.

Design

The present research was an instrumental study, as we assessed the psychometric characteristics of a psychological questionnaire (i.e. the STAI). The study was drafted following the recommendations made by Hartley (2012).

Procedure

Participants were selected through non-probabilistic sampling. The STAI was applied in public places visited by a large number of people and in various

organizations that cooperated in the study. Prior to administering the STAI, the experimenters verified that participants were over 18 years old and informed participants of the purpose of the study, guaranteeing its confidentiality. Next, the STAI was handed to participants and the experimenter stayed close to them to respond to any questions.

Statistical analyses

Analyses were conducted in two stages, each with a different part of the sample. First, we performed a discriminant analysis separately for the trait anxiety and state anxiety subscales. Our aim was to obtain a shorter selection of items that enabled a correct classification of subjects with high and low anxiety. This was done by forming two groups based on a cut-off score. The cut-off score was set at the mean plus one standard deviation. After that, we calculated the internal consistency (i.e. Cronbach's alpha) using a polychoric correlation matrix. This was done using IBM SPSS Statistics 20 and R software.

Second, we conducted a confirmatory factor analysis to assess the fit of the short form obtained with the discriminant analysis. We also verified the fit of the short forms proposed by other authors. This allowed us to determine which of the short versions demonstrated the best fit with the present sample. To do so, we selected the robust Weighted Least Squares with Mean- and Variance-adjusted chi-square (WLSMV) method, calculated by using a diagonal weight matrix. This method is appropriate for categorical measures (Brown, 2006). We compared the different models using the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI). Although both scales of the STAI were theoretically independent, we developed path diagrams showing relations between both, since many studies have found a correlation between them. This was done by using R software and the lavaan package (Rosseel, 2012).

Results

Discriminant analysis

We randomly extracted a subsample of 200 participants from the total sample and conducted a separate discriminant analysis for state anxiety and trait anxiety. We created two groups (i.e. high anxiety and low anxiety) using the mean score plus one standard deviation as a cut-off score. As regards state anxiety, the cut-off score was 29. The mean anxiety score was 14.67 ($SD = 7.14$) for the low state anxiety group and 39.07 ($SD = 7.02$) for the high state anxiety group. As regards trait anxiety, the cut-off score was 33. The mean anxiety score was 18.73 ($SD = 7.53$) for the low trait anxiety group and 40.20 ($SD = 6.43$) for the high trait anxiety group.

After creating the groups, we adjusted their respective discriminant functions. Five items were retained in the discriminant function in both state and trait anxiety. Table 1 shows adequate fit indices, with significant results in both cases. In addition, the percentage of correct classification exceeded 92% in the two discriminant functions. This indicated that a selection of five items per scale was effective.

Table1.

Summary of the discriminant analyses of the anxiety subscales.

Indicator	State anxiety	Trait anxiety
Eigenvalue	1.68	1.25
Wilks' λ	.374	.445
χ^2 (<i>d.f.</i>)	174.67 (5)	158.27 (5)
<i>p</i>	< .001	< .001
Canonical correlation	0.79	0.75
% correct classification	95.5	92.5
Centroids		
Low fit	-0.537	-0.512
High fit	3.084	2.412

Note. *d.f.* = degrees of freedom; *p* = probability.

Finally, we calculated the reliability of the five items of each subscale. This was done using Cronbach's alpha, calculated from the polychoric correlation matrix. The state anxiety subscale yielded an alpha reliability value of .81, and no items worsened this index. The trait anxiety scale yielded an alpha value of .76; there was one item whose suppression improved this value by two hundredths of a point.

Confirmatory factor analysis

After selecting the items that enabled making the best predictions of groups with high and low anxiety, we verified whether such structure was adequate with a different sample. We also compared the fit of the model with that of other theoretical structures. To do so, we used the part of the sample that had not been used in the discriminant analysis ($n = 957$). Table 2 summarizes the theoretical models tested and the items they include.

Table 2.

Summary of items in each factor and model of the confirmatory factor analysis.

Model	Author	State	Trait
Model 1 (M1)	Own model	3, 8, 16, 18 y 20	3, 8, 10, 12 y 18
Model 2 (M2)	van Knippenberg et al. (1990)	2, 4, 11, 15, 17 y 18	7, 14, 15, 16, 17 y 18
Model 3 (M3)	Kaipper et al. (2010)	1, 2, 5, 6, 7, 8, 11, 13, 14, 16, 17, 18 y 19	1, 2, 5, 7, 8, 9, 10, 12, 16, 17, 18 y 20
Model 4 (M4)	Fioravanti-Bastos et al. (2011)	1, 3, 5, 12, 15 y 17	7, 9, 11, 13, 16 y 20

Note. In M4 we eliminated 2 items that did not belong to *form X* of the STAI and replaced them with the following ones that would have been selected following the methodology of the article of reference.

After establishing the models, we verified which of them demonstrated better fit. Table 3 shows the fit indices of the various theoretical models compared. It reveals that the structure obtained after reducing the number of items with the discriminant analysis showed adequate fit. Models 3 and 4 showed good fit, although Model 2 showed the best fit.

Table 3.
Fit of the models proposed in the confirmatory factor analysis.

Statistic	Model			
	m1	m2	m3	m4
χ^2	616.21	622.62	3,917.89	722.19
Degrees of freedom	34	53	274	53
χ^2 / Degrees of freedom	18.12	11.75	14.30	13.63
p	< .001	< .001	< .001	< .001
Comparative Fit Index (CFI)	.935	.96	.945	.959
Tucker-Lewis Index (TLI)	.914	.95	.94	.949

Note. m1: data from this article; m2: proposed by van Knippenberg et al. (1990); m3: proposed by Kaipper et al. (2010); m4: Fioravanti-Bastos et al. (2011), modified. p = probability.

Finally, we analyzed the correlations between factors obtained from the covariances between trait and state anxiety. The model developed from the discriminant analysis showed a correlation of .84. The remaining models showed similar correlations between factors (i.e. .82 in Model 2, .83 in Model 3 and .72 in Model 4). For this reason, we decided not to verify the fit of each factor separately.

Discussion

The objective of the present study was to develop a short form of both subscales of the STAI. First, we defined a five-item version for each subscale that was able to discriminate between groups with high and low anxiety. The number of items selected is similar to that of other short versions of the STAI, since most short forms include six

items. However, a comparison of the fit of the form obtained through discriminant analysis revealed that other short forms had higher fit indices. Despite this, the items selected through discriminant analysis were adequate.

The short form of five items per subscale was able to correctly classify over 90% of subjects according to their anxiety level. This is an interesting finding considering that most short forms include six items. This number of items is mainly set to replicate the form developed by Marteau and Bekker (1992) or the methodology followed by them. However, these authors selected six items without explaining why they had done so. In addition, as stated by several authors, other forms with a different number of items have been found to be adequate (Fioravanti-Bastos et al., 2011; van Knippenberg, 1990). For this reason, it is interesting to test scales with a different number of items and to try to set a number of items using objective and theoretical criteria. In addition, it is adequate to propose short forms of both subscales because, although state anxiety and trait anxiety are theoretically independent (Spielberger et al, 1970), many studies have shown a relationship between both factors (e.g. Guillén-Riquelme & Buela-Casal, 2011). This is consistent with the high correlations observed in all the models tested in the confirmatory factor analysis.

As regards the comparison between models, we observed that the version of five items per factor had good fit. Yet, other models analyzed had higher fit indices. The items that discriminate the best between high and low anxiety levels may not have the best factor structure throughout all levels of anxiety. A possible explanation to this may be that the cut-off score we used to create the group was the mean plus one standard deviation and not a clinical score. However, no clinical score was available in the Spanish adaptation (Buena-Casal et al., 2011). The values of the original form (Spielberger et al., 1970) were considerably higher than those of the Spanish adaptation

(Guillén-Riquelme & Buela-Casal, 2011), so it would not be appropriate to use this value either.

Among the various structures analyzed, the one that showed better fit to the sample was that of van Knippenberg et al. (1990). This version was composed of six items per scale. By contrast, the version of 12 and 13 items proposed by Kaipper et al. (2010) showed the lowest fit indices. Thus, it is not necessary to have a large number of items to obtain a consistent structure (van Knippenberg et al., 1990; Marteau & Bekker, 1992). In fact, it is possible to develop an adequate short form with five or six items per scale, as we did in the present study. When comparing the form developed by van Knippenberg et al. (1990) to that proposed after the discriminant analysis, we observed very small differences of fit between both. Therefore, both versions seem to be appropriate to assess anxiety in Spanish samples. Other item combinations may also be adequate, as there seem to be several possible combinations with good fit. This explains the fact that we found a valid form of the Spanish adaptation (Perpiñá-Galvañ, Richart-Martínez & Cabañero-Martínez, 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez & Martínez-Durá, 2011) based on the items of the short form proposed by Chlan et al. (2003) even though the latter was developed from *form Y* and the Spanish version was developed from *form X*. Thus, although the items did show an exact match between versions, we found a good fit based on such structure.

It is increasingly common to develop short forms (e.g. Kämpfe et al., 2012; Martínez-López et al., 2012; Kruyen, Emons & Sijtsma, 2013) and cultural adaptations (Muñiz, Elosua & Hambleton, 2013) of psychological assessment questionnaires with adequate reliability and validity guarantees (Delgado-Rico, Carretero-Dios & Ruch, 2012). It is essential to determine the factor structure of personality questionnaires (e.g. de la Fuente et al., 2013; Ginting, Näring, van der Veld, Srisayektu & Becker, 2013). In

this context, proposing a short version of the STAI for a Spanish sample with adequate reliability measures is an important contribution. A short version is very useful for assessments involving a large number of questionnaires, particularly in situations in which respondents cannot or must not undergo long assessments.

References

- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K., & Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology, 11*, 443-457.
- Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behaviour Research and Therapy, 36*, 777-788. doi: 10.1016/S0005-7967(98)00023-0
- Brown, T. A. (2006). Data Issues in CFA: Missing, Non-Normal, and Categorical Data. En D. A. Kenny (Ed.), *Confirmatory Factor Analysis for Applied Research* (pp. 363-411). New York: The Guilford Press.
- Buela-Casal, G., Guillén-Riquelme, A., & Seisdedos Cubero, N. (2011). *Cuestionario de Ansiedad Estado-Rasgo: Adaptación española* (8ª ed.). Madrid: TEA Ediciones.
- Caci, H., Baylé, F. J., Dossios, C., Robert, P. y Boyer, P. (2003). The Spielberger trait anxiety inventory measures more than anxiety. *European Psychiatry, 18*, 394-400. doi:10.1016/j.eurpsy.2003.05.003
- Chen, Y. J. (2006). *Psychophysiological determinants of repeated ventilator weaning failure* (Tesis doctoral). Descargada de <http://www.nursing.arizona.edu>
- Chlan, L., Savik, K., & Weinert, C. (2003). Development of a Shortened State Anxiety Scale From the Spielberger State-Trait Anxiety Inventory (STAI) for Patients Receiving Mechanical Ventilatory Support. *Journal of Nursing Measurement, 11*, 283-293. doi:10.1891/jnum.11.3.283.61269

- Davey, H. M., Barratt, A. L., Butow, P. N., & Deeks, J. J. (2007). A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *Journal of Clinical Epidemiology*, *60*, 356-360. doi: 10.1016/j.jclinepi.2006.07.015
- de la Fuente, E. I., Lozano, L. M., García-Cueto, E., San Luis, C., Vargas, C., Cañadas, G. R.,... Hambleton, R. K. (2013). Development and validation of the Granada Burnout Questionnaire in Spanish police. *International Journal of Clinical and Health Psychology*, *13*, 216-225
- Delgado-Rico, E., Carretero-Dios, H., & Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology*, *12*, 449-460.
- Endler, N. S., Cox, B. J., Parker, J. D. A., & Bagby, R. M. (1992). Self-Reports of Depresión and State-Trait Anxiety: Evidence for Differential Assessment. *Journal of Personality and Social Psychology*, *63*, 832-838. doi:10.1037/0022-3514.63.5.832
- Fioravanti-Bastos, A. C. M., Cheniaux, E., & Landeira-Fernandez, J. (2011). Development and Validation of a Short-Form Version of the Brazilian State-Trait Anxiety Inventory. *Psicologia: Reflexão e Critica*, *24*, 485-494. doi: 10.1590/S0102-79722011000300009
- Ginting, H., Näring, G., van der Veld, W. M., Srisayekti W., & Becker, E. S. (2013). Maladaptive personality traits in adolescence: Psychometric properties of the Personality Diagnostic Questionnaire-4+. *International Journal of Clinical and Health Psychology*, *13*, 235-242.

- Guillén-Riquelme, A., & Buéla-Casal, A. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23, 510-515.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Kaipper, M. B., Chachamovich, E., Hidalgo, M. P. L., Torres, I. L. S., & Caumo, W. (2010). Evaluation of the structures of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research*, 68, 223-233. doi:10.1016/j.jpsychores.2009.09.013
- Kämpfe, C. K., Gloster, A.T., Wittchen, H-U., Helbig-Lang, S., Lang, T., Gerlach, A. L.,... Deckert, J. (2012). Experiential avoidance and anxiety sensitivity in patients with panic disorder and agoraphobia: Do both constructs measure the same? *International Journal of Clinical and Health Psychology*, 12, 5-22.
- Kruyen, P. M., Emons, W. H. M., & Sijtsma, K. (2013). Shortening the S-STAI: Consequences for research and clinical practice. *Journal of Psychosomatic Research*, 75, 167-172. doi:10.1016/j.jpsychores.2013.03.013
- Manzoni, G. M., Pagnini, F., Castelnuovo, G., & Molinari, E. (2008). Relaxation training for anxiety: a ten-years systematic review with meta-analysis. *BioMed Central Psychiatry*, 8, 41. doi:10.1186/1471-244X-8-41.
- Marine, A., Ruotsalainen, C. S., & Verbeek, J. H. (2009). *Preventing occupational stress in healthcare workers (Review)*. Descargado de The Cochrane Collaboration de: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD002892.pub2/pdf>. doi: 10.1002/14651858.CD002892.pub2

- Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology, 31*, 301-306. doi:10.1111/j.2044-8260.1992.tb00997.x
- Martínez-López, P., Durá Ferrandis, E., Andreu Vaillo, Y., Galón Garrido, M. J., Murgui Pérez, S., & Ibáñez Guerra, E. (2012). Structural validity and distress screening potential of the Hospital Anxiety and Depression Scale in cancer. *International Journal of Clinical and Health Psychology, 12*, 435-447.
- Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema, 25*, 151-157.
- Muñiz, J., & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo, 31*, 108-121.
- Perpiñá-Galvañ, J., Richart-Martínez, M., & Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología, 47*, 184-189. doi:10.1016/j.arbres.2010.11.006
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J., & Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem, 19*, A04:1-A04:6.
- Rosa-Alcázar, A. I., Sánchez-Meca, J., Gómez-Conesa, A., & Marín-Martínez, F. (2008). Psychological treatment of obsessive-compulsive disorder: A meta-analysis. *Clinical Psychology Review, 28*, 1310-1325. doi:10.1016/j.cpr.2008.07.001.

- Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48, 1-36.
- Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory: STAI(Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., & Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- van der Bij, A. K., de Weerd, S., Cikot, R. J. L. M., Steegers, E. A. P., & Braspenning, J. C. C. (2003). Validation of the Dutch Short Forma of the State Scale of the Spielberger State-Trait Anxiety Inventory: Considerations for Usage in Screening Outcomes. *Community Genetics*, 6, 84-87. doi: 10.1159/000073003
- van Knippenberg, F. C. E., Duivenvoorden, H. J., Bonke, B., & Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology*, 43, 995-1000. doi:10.1016/0895-4356(90)90083-2
- Watson, D., Clark, L. A., & Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 429-442.

Estudio 4

Versión breve del STAI en adolescentes y universitarios españoles

Alejandro Guillén-Riquelme y Gualberto Buela-Casal

Artículo publicado en *Terapia Psicológica*

Revista indexada en el *Journal Citation Reports*

Factor de impacto 2012 = 0,854.

Referencia:

Guillén-Riquelme, A. y Buela-Casal, A. (2013). Versión breve del STAI en adolescentes y universitarios españoles. *Terapia Psicológica*, 31, 293-299.

Versión breve del STAI en adolescentes y universitarios españoles

Resumen

El STAI es uno de los instrumentos de evaluación psicológica más empleados, generándose diversas versiones breves del mismo. No se han hallado versiones breves en las que se haya comprobado el ajuste para muestras de adolescentes ni universitarios españoles. El objetivo de este artículo es comprobar cuál de las versiones cortas propuestas en la literatura tiene un mejor ajuste en adolescentes y universitarios. Para ello, se aplicó el STAI a 482 adolescentes escolarizados y 510 estudiantes universitarios, de diversas ciudades españolas. Mediante análisis factorial confirmatorio se evaluó el ajuste de cinco versiones breves del STAI. En los universitarios hay dos modelos en los que se observa un buen ajuste. En los adolescentes no ajusta ningún modelo. Puede que el proceso de diferenciación emocional con la edad explique que en los adolescentes las versiones breves halladas (siempre con adultos) no serían válidas, sí pudiendo emplearse con universitarios.

Palabras clave: STAI, versión breve, ansiedad, adolescencia.

Abstract

The STAI is an instrument very used, generating different short versions. We found no short versions in which it is tests the fit for teenagers and college samples Spanish. The aim of this paper is to test which of the short versions proposed in the literature have a better fit in adolescents and college students. To do this, we applied the STAI to 482 adolescent students and 510 university students from different Spanish cities. Through factorial confirmatory analysis we evaluated five short versions of the STAI. At the

university there are two models in which there is a good fit. In adolescents not fit any model. Maybe the emotional differentiation process explains that the found short versions (always with adults) are not valid in teenagers, and can be used with university students.

Keywords: STAI, short version, anxiety, adolescence.

El State-Trait Anxiety Inventory es uno de los instrumentos clásicos para evaluar ansiedad, uno de los trastornos psicológicos con una mayor prevalencia en población general (Beutel, Bleichner, von Heumann, Tritt y Hardt, 2011) así como en universitarios (Micin y Bagladi, 2011). Desarrollado por Spielberger, Gorsuch y Lushene (1970) este cuestionario ha sido citado en más de 14.000 estudios de archivo, siendo uno de los instrumentos más empleados por psicólogos (Spielberger y Reheiser, 2009). Su aplicación permite obtener dos escalas: una de ansiedad rasgo (tendencia estable de percibir estímulos como más amenazantes) y ansiedad estado (presencia de estímulos ansiosos en el entorno del evaluado en los días cercanos a la evaluación). Además, el STAI ha contado con más de 60 adaptaciones o traducciones (Spielberger y Reheiser, 2009), entre ellas, la española (Buela-Casal, Guillén-Riquelme y Seisdedos, 2011), donde es el sexto cuestionario más utilizado por psicólogos clínicos (Muñiz y Fernández-Hermida, 2010). Esta adaptación cuenta con adecuadas propiedades psicométricas en población general (Guillén-Riquelme y Buela-Casal, 2011).

En 1983, se realiza una doble modificación del STAI: a) cambia la dirección de algunos ítems, para equiparar el número de elementos positivos y negativos; b) varía el contenido de otros, para facilitar su comprensión (Spielberger, 1983). Estos cambios dan respuesta a diversas críticas realizadas sobre el STAI. A la primera versión se le denomina como forma X y a la versión con los cambios descritos, como forma Y. Este hecho es importante, porque el contenido de algunos ítems varía según la versión (forma X o Y), pese a que las evaluaciones realizadas con una u otra no difieran (Spielberger, 1983). La adaptación española se realizó a partir de la forma X (Buela-Casal et al., 2011).

Diversos autores realizan una versión reducida de los ítems del cuestionario que permite evaluar ansiedad de forma fiable. La versión breve realizada por van

Knippenberg, Duivenvoorden, Bonke y Passchier (1990), aplica, por primera vez, una metodología científica y unas pruebas estadísticas adecuadas; requisitos que no cumplen las propuestas anteriores. Estos autores realizan nueve versiones breves, a partir de las correlaciones ítem-total corregidas. Las versiones varían en el número de ítems (combinaciones de 4, 6 y 8 ítems). Además, utilizan dos muestras de pacientes médicos y una de estudiantes universitarios. En este trabajo, se concluye que es posible establecer versiones breves del STAI, con adecuadas propiedades psicométricas. También, se establece que las versiones de seis y ocho ítems son las adecuadas para la evaluación de la ansiedad estado y rasgo; mientras que las de cuatro ítems resultan insuficientes. En la muestra de universitarios, la versión breve está formada por los ítems número 2, 4, 8, 11, 12, 15, 17 y 18 para la ansiedad estado, y por los ítems 7, 8, 14, 15, 16, 17, 18 y 20 para la ansiedad rasgo. Hay que precisar que las versiones de 6 ítems se obtienen, eliminando, en la escala de ansiedad estado, los ítems 8 y 12, y, en la de ansiedad rasgo, los ítems 8 y 20.

Recientemente, se han propuesto nuevas versiones breves del STAI. Kaipper, Chachamovich, Hidalgo, Torres y Caumo (2010) realizaron una versión breve de ambas escalas. El banco de ítems estaba formado por todos los ítems de la forma X, más los ítems con diferente contenido de la forma Y. Estos autores establecen que los ítems de la versión breve de la escala de ansiedad estado son: 1, 2, 5, 6, 7, 8, 11, 13, 14, 16, 17, 18 y 19, además de otros ítems de la forma Y. En el caso de la ansiedad rasgo los ítems seleccionados son: 1, 2, 5, 7, 8, 9, 10, 12, 16, 17, 18 y 20, además de otros ítems de la forma Y. Esta versión breve tiene 25 ítems pero, en la actualidad, también se defienden versiones de seis ítems por cada una de las escalas. De esta forma, Fioravanti-Bastos, Cheniaux y Landeira-Fernandez (2011) afirman que la selección de los ítems 1, 3, 5, 12,

15 y 17, para la ansiedad estado y los ítems 7, 9, 11, 13, 16 y 20 para la ansiedad rasgo, resultan adecuados para evaluar estos factores.

En España se realizó una versión breve mediante una muestra de población general, empleando análisis discriminante (Buela-Casal y Guillén-Riquelme, en prensa). Esta versión está compuesta por los ítems 3, 8, 16, 18 y 20 para la ansiedad estado y los ítems 3, 8, 10, 12 y 18 para la ansiedad rasgo. Por otra parte, se comparaba, mediante análisis factorial confirmatorio, si el ajuste de esa versión corta era mejor que el de otras reducciones anteriores, concluyendo que la versión breve con un mejor ajuste era la presentada por van Knippenberg et al. (1990). Además del trabajo de Buela-Casal y Guillén-Riquelme (en prensa), se ha realizado otra versión breve con muestra española (Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez y Martínez-Durá, 2011). Sin embargo, esta versión muestra un par de problemas: a) únicamente emplea muestra de pacientes con intubación respiratoria; b) se realiza la confirmación de la versión propuesta por Marteau y Bekker (1992), derivada de la forma Y del STAI, cuyos ítems sólo coinciden parcialmente con los de la adaptación española, proveniente de la forma X (Buela-Casal et al., 2011).

El STAI se caracteriza por la calidad de sus propiedades psicométricas a la hora de evaluar ansiedad en diversas poblaciones, entre ellas la de adolescentes. Así pues, la adaptación española resulta fiable y válida, presentando baremos específicos para este segmento de la población (Buela-Casal et al., 2011). Además, la adaptación chilena (desarrollada a partir de la versión española) también posee una consistencia interna elevada al ser aplicada en una muestra de adolescentes (Vera-Villaruel, Celis-Atenas, Córdova-Rubio, Buela-Casal y Spielberger, 2007).

Entre las numerosas versiones breves del STAI ninguna sobresale por un mayor apoyo empírico, ni ha sido validada en población adolescente. Por ello, el objetivo del presente trabajo es comprobar si las diferentes versiones breves realizadas, a partir de la forma X del STAI, ajustan adecuadamente en: a) una muestra de adolescentes españoles y b) una muestra de universitarios. Al utilizar ambas muestras, se podrá comparar, de forma más precisa, el ajuste de la versión corta en la adolescencia y al inicio de la edad adulta. La hipótesis inicial es el adecuado ajuste de los diferentes modelos, siendo los hallados por van Knippenberg et al. (1990) y por (Buela-Casal y Guillén-Riquelme, en prensa), los que ajusten de forma más precisa en adolescentes y universitarios, puesto que en ellos se ha observado mejores propiedades psicométricas en una muestra de adultos españoles.

Método

Participantes

Para la consecución del objetivo planteado, se utilizaron dos muestras independientes. La primera de ellas formada por 482 adolescentes de diversos institutos (públicos y privados) de cuatro ciudades españolas. Del total, 186 (42,20%) son varones. La media de edad es de 15,48 ($DT = 1,58$), siendo el mínimo de 12 y el máximo de 19. En el caso de los varones la media es de 15,31 años ($DT = 1,76$). En las mujeres la media es de 15,61 años ($DT = 1,59$).

La segunda muestra está formada por 510 universitarios matriculados en seis universidades españolas (todas ellas públicas), procedentes de diversas carreras. La media de edad es de 23,15 años ($DT = 5,51$) y el rango de edad abarca desde 18 a 54 años, con más de un 90% de los participantes menores de 29 años). En el caso de los varones ($n = 186$; 36,5%) la media de edad es de 25,05 años ($DT = 7,71$). En las mujeres la edad media es de 22,06 años ($DT = 3,96$).

Instrumentos

Tras realizar preguntas relativas al lugar de residencia, sexo y edad, se procedió a aplicar la adaptación española del *State-Trait Anxiety Inventory* (Buela-Casal et al, 2011). Este cuestionario permite evaluar ansiedad rasgo (tendencia a percibir los estímulos como ansiógenos o no) y estado (grado en el que hay elementos ansiosos en el entorno de la persona en el tiempo cercano a la evaluación). La escala de respuesta es de tipo Likert con 4 alternativas (en la adaptación española desde 0 hasta 3). Tanto la ansiedad estado como la ansiedad rasgo cuentan con 20 ítems para su evaluación. Por ello, el rango de respuestas va desde 0 a 60, correspondiéndose mayor nivel de ansiedad con una mayor puntuación. La adaptación española tiene adecuadas propiedades psicométricas (Guillén-Riquelme y Buela-Casal, 2011; Buela-Casal et al., 2011). Además, la escala se puede aplicar en población adolescente así como en universitarios, teniendo evidencias de fiabilidad y validez.

Procedimiento

El procedimiento es diferente para cada una de las muestras obtenidas. En el caso de la muestra de adolescentes, en primer lugar se contactó con diversos centros de cuatro ciudades españolas. La aplicación de los cuadernillos transcurrió en la propia aula, durante las horas de tutorías a grupos completos (entre 25 y 35 alumnos de forma general). El experimentador informaba de la confidencialidad del estudio así como del consentimiento informado. Seguidamente se repartieron los cuadernillos al grupo y el experimentador permaneció en el aula durante la aplicación, para resolver posibles dudas de cumplimentación.

En el caso de la muestra de universitarios, se contactó con los centros para la colaboración. El experimentador se personó en las aulas e informó sobre la confidencialidad y el consentimiento informado. De nuevo se aplicó la prueba a grupos

completos, pese a que en este caso eran, de media, tres veces más numerosos que los de los adolescentes. Durante la prueba, el experimentador permaneció en el aula para contestar posibles dudas.

Análisis de datos

El análisis factorial confirmatorio fue seleccionado como la técnica analítica idónea para el ajuste de los modelos. Esta aproximación permite facilitar la comparaciones de invarianza y equivalencia para futuros estudios, garantizando la utilización del test en diversas muestras (Muñiz, Elosua y Hambleton, 2013). Se pusieron a prueba cinco modelos diferentes: el modelo 1 (m1) se corresponde al presentado por Buela-Casal y Guillén-Riquelme (en prensa); el modelo 2 (m2) es el propuesto por van Knippenberg et al. (1990) de 6 ítems; el modelo 3 (m3) se corresponde con el anterior pero en la versión de 8 ítems; el modelo 4 (m4) es el propuesto por Kaipper et al. (2010) y el modelo 5 (m5) por Fioravanti-Bastos et al. (2011). La naturaleza categórica de los ítems (escala de cuatro alternativas) hizo necesario aplicar una prueba que solventase la falta de normalidad, resultante del uso de este tipo de variables. Así pues, se seleccionó el método robusto a partir del *Weighted Least Squares with Mean- and Variance-adjusted chi-square* (WLSMV; Brown, 2006). La comprobación del ajuste de los modelos a los datos se realizó mediante cuatro estadísticos: el χ^2 dividido por los grados de libertad, el *Comparative Fit Index* (CFI), el *Tucker-Lewis Index* (TLI) y el *Root Mean Square Error of Approximation* (RMSEA). Los dos primeros deben tener valores superiores al 0,95 para que se considere que el ajuste es adecuado. En el RMSEA se exigen valores inferiores a 0,08. A la hora de establecer el modelo se incluyeron, de forma conjunta, ambas escalas (estado y rasgo) y su covariación, ya que estudios previos indicaban una fuerte relación entre ellas (Buela-Casal y Guillén-Riquelme, en prensa). Para los análisis se utilizó el paquete lavaan

(Rosseel, 2012) del programa R. Tras ello, la redacción del artículo se realizó siguiendo las recomendaciones propuestas por Hartley (2012).

Resultados

Análisis factorial confirmatorio en adolescentes

En la muestra de adolescentes se comprobó el ajuste de los cinco modelos teóricos, expuestos anteriormente. En cada uno de ellos y, en función de los índices de modificación, se incluyeron diversas covarianzas de los errores, respetando que: 1) los errores fuesen de ítems de un mismo factor, 2) el índice de modificación tuviese un valor superior a 20 puntos. En el modelo número 4 al tener un mayor número de ítems, para evitar una sobrecarga de covariaciones de errores, se amplió este criterio hasta un valor de 30 puntos. En la Tabla 1, se recoge un resumen de los índices de ajuste de cada uno de los modelos teóricos. Ninguno de ellos alcanza niveles adecuados de ajuste, aunque, en el modelo 5 se observan valores más cercanos a un ajuste adecuado.

Análisis factorial confirmatorio en universitarios

En el caso de los universitarios, hay un correcto ajuste de diferentes modelos. El modelo 1 tiene índices próximos a los criterios establecidos. El modelo 2, conformado por 6 ítems, tiene un mejor ajuste que el resto, en la muestra de universitarios. El modelo 3 tiene un ajuste similar al anterior, siendo ligeramente inferior, y contando con dos ítems más por cada una de las escalas. El modelo 4 es, junto con el 1, el que cuenta con índices de ajuste más bajos. El modelo 5, no llega a ajustar adecuadamente, aunque el CFI resulte superior a 0,95. Un resumen se recoge en la Tabla 2.

Tabla 1.
Ajuste de los modelos propuestos en el análisis factorial confirmatorio en adolescentes.

Estadístico	Modelo				
	m1	m2	m3	m4	m5
χ^2 (g.l.)	231,65(28)	346,20(47)	583,29(95)	1.482,25(255)	230,51(44)
χ^2 /g.l.	8,27	7,37	6,14	5,81	5,24
p	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001
CFI	0,887	0,906	0,889	0,936	0,891
TLI	0,818	0,868	0,86	0,925	0,837
RMSEA	0,122	0,114	0,103	0,099	0,093
I.C. inf. RMSEA	0,108	0,103	0,095	0,094	0,082
I.C. sup. RMSEA	0,137	0,126	0,111	0,104	0,105

Nota. m1: datos de Buela-Casal y Guillén-Riquelme (en prensa); m2: propuesto por van Knippenberg et al. (1990) de 6 ítems; m3: propuesto por van Knippenberg et al. (1990) de 8 ítems, m4: propuesto por Kaipper et al. (2010); m5: Fioravanti-Bastos et al. (2011) modificado. g.l. = grados de libertad; p = probabilidad; I.C. = intervalo de confianza al 95%; CFI = *Comparative Fit Index*; TLI = *Tucker-Lewis Index*; RMSEA = *Root Mean Square Error of Approximation*.

Tras analizar el ajuste se comprobó la covariación entre factores. En el modelo 2 (el que tiene mejores índices) la covariación es de 0,89, lo que indica una fuerte relación entre ambos factores. En el resto de modelos también se encuentra valores similares entre la covariación de la ansiedad estado y rasgo.

Análisis de fiabilidad en el modelo con mejor ajuste en universitarios

A partir de los ítems del modelo que mejor ajuste obtiene, en la muestra de universitarios, se realizó un análisis de la fiabilidad para cada una de las escalas. En el caso de la ansiedad estado, el alfa de Cronbach es de 0,82. En la correlación ítem total corregida los valores oscilan desde 0,52 a 0,73, por lo que todos los ítems están relacionados con el total. En el caso de la ansiedad rasgo el alfa de Cronbach es de 0,69 (la eliminación de un ítem mejoraría el alfa en dos centésimas). En la correlación ítem total corregida se observan resultados dispares puesto que la correlación para los ítems

15, 16, 17 y 18 oscila de 0,42 a 0,55, mientras en el ítem 7, es de 0,27 y en el ítem 14 es de 0,38.

Tabla 2.

Ajuste de los modelos propuestos en el análisis factorial confirmatorio en universitarios.

Estadístico	Modelo				
	m1	m2	m3	m4	m5
χ^2 (g.l.)	160,54(29)	143,97(44)	358,58(94)	1.487,54(259)	202,18(45)
$\chi^2/g.l.$	5,54	3,27	3,81	5,74	4,49
p	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001
CFI	0,942	0,976	0,96	0,964	0,959
TLI	0,909	0,964	0,949	0,959	0,939
RMSEA	0,094	0,067	0,074	0,096	0,083
I.C. inf. RMSEA	0,08	0,055	0,066	0,092	0,071
I.C. sup. RMSEA	0,109	0,079	0,083	0,101	0,095

Nota. m1: datos de Buela-Casal y Guillén-Riquelme (en prensa); m2: propuesto por van Knippenberg et al. (1990) de 6 ítems; m3: propuesto por van Knippenberg et al. (1990) de 8 ítems, m4: propuesto por Kaipper et al. (2010); m5: Fioravanti-Bastos et al. (2011) modificado. g.l. = grados de libertad; p = probabilidad; I.C. = intervalo de confianza al 95%; CFI = *Comparative Fit Index*; TLI = *Tucker-Lewis Index*; RMSEA = *Root Mean Square Error of Approximation*.

Discusión

En el prolegómeno de este estudio se propuso como objetivo “comprobar si las diferentes versiones breves realizadas, a partir de la forma X del STAI, ajustan adecuadamente en: a) una muestra de adolescentes españoles y b) una muestra de universitarios“, también se partió de la hipótesis de que se obtendría un adecuado ajuste de los modelos en ambas muestras. En los universitarios la escala breve que mejor ajuste ha mostrado, coincide con la establecida por Buela-Casal y Guillén-Riquelme (en prensa) en población general española. Se confirma que la versión breve propuesta para población general (Buela-Casal y Guillén-Riquelme, en prensa), a partir de la versión de seis ítems de van Knippenberg et al. (1990), se puede utilizar para evaluar ansiedad en universitarios. El hecho de contar con una versión breve de este cuestionario tan empleado resulta necesario, no solo en los grupos analizados, si no también en pacientes

con problemas físicos y que requieran varias evaluaciones (Martínez-López et al., 2012).

Ninguna de las versiones breves del STAI, incluida la seleccionada para los universitarios, muestra un correcto ajuste en el grupo de adolescentes. La nula incidencia de un buen ajuste en el grupo de adolescentes, de las cinco versiones breves teóricas propuestas, resulta extraña ya que la versión completa del STAI tiene adecuadas propiedades psicométricas en esta población (Buela-Casal et al., 2011). El hecho de que en el grupo de adolescentes no se ajuste ninguna de las versiones breves teóricas tiene varias explicaciones. En primer lugar, cabe indicar que la mayoría de las versiones breves se realizaron a partir de muestras de adultos (con o sin trastornos de ansiedad). Por ello, puede que ninguna de estas versiones se ajuste a una muestra de adolescentes. Sin embargo, este argumento no justifica el hecho de que la versión corta para adultos no se pueda aplicar a adolescentes, mientras que en el STAI completo se observan adecuadas propiedades psicométricas en esta muestra (Buela-Casal et al., 2011; Vera-Villaroel et al., 2007). Una posible explicación a este hecho se puede hallar en la teoría general del afecto, donde se defiende que la evolución de las emociones se produce de un modo similar a la de la inteligencia, según la hipótesis de la diferenciación. Se considera que se parte de afectos muy generales, desde los cuales, a partir de un proceso de diferenciación, se desarrollan las emociones más específicas, permitiendo diferenciarlas entre sí. Así pues, Yik, Russell y Feldman-Barrett (1999) proponen la existencia de dos continuos en los que se enmarcarían las diferentes emociones: el primero de ellos el *arousal*, oscilante desde la activación a la desactivación; el segundo el de la valencia del afecto, indicativo del carácter positivo o negativo de los sentimientos. La ansiedad se vincularía exclusivamente con afecto negativo (Watson, Clark y Stasik, 2011). De acuerdo con este modelo, las emociones

son más generales en la adolescencia que en el inicio de la edad adulta (tramo en el que se encuentra la muestra de universitarios), ya que la diferenciación se incrementa con el desarrollo psicológico del individuo. Por tanto en adolescentes de muy tempranas edades puede ser recomendable emplear versiones breves de cuestionarios más generales sobre sensibilidad ansiosa (v. gr. Kämpfe et al., 2012).

En los universitarios el modelo que mejor ajusta es el presentado por van Knippenberg et al. (1990) lo que confirma la hipótesis inicial, donde se enunciaba que dicha versión tendría el mejor ajuste, puesto que esta propuesta fue la que se seleccionó como versión breve en población general española (Buela-Casal y Guillén-Riquelme, en prensa). Este hecho, confirma que se puede emplear una versión breve en universitarios, pese a que tengan mayores puntuaciones medias en ansiedad que la población general (Buela-Casal et al., 2011). Además, al considerar el grupo de universitarios como una parte de la población general, se aporta una nueva evidencia de que la versión breve seleccionada en la investigación de Buela-Casal y Guillén-Riquelme (en prensa), es la más apropiada. Observando otras versiones breves de cuestionarios de ansiedad que ajustan bien en este tipo de muestras (v. gr. Antúnez y Vinet, 2012). Así pues, la versión de seis ítems de van Knippenberg et al. (1990) permite evaluar ansiedad estado y rasgo en población general (incluyendo al grupo de universitarios).

En resumen, la versión propuesta por van Knippenberg et al. (1990) y confirmada por Buela-Casal y Guillén-Riquelme (en prensa), compuesta por los ítems 2, 4, 11, 15, 17 y 18, para la ansiedad estado y por los ítems 7, 14, 15, 16, 17 y 18 para la ansiedad rasgo, permite evaluar de forma fiable ambos factores en universitarios. En el caso de los adolescentes, pese a que la versión completa sea válida, las versiones breves utilizadas hasta la fecha no proporcionan un correcto ajuste. Es imprescindible hallar una versión breve que permita evaluar ansiedad en adolescentes de forma válida y

fiable. Una de las limitaciones del presente estudio es que las muestras no están seleccionadas de forma completamente aleatoria. Por ello, cabría plantearse en futuros estudios el usar este tipo de selección muestral, así como comprobar la validez de contenido del cuestionario (Delgado-Rico, Carretero-Dios y Ruch, 2012). A pesar de ello, al establecer que la versión breve puede ser empleada para el análisis de la ansiedad rasgo y estado en universitarios, se aporta una nueva evidencia de su correcto funcionamiento en población general. Este hallazgo puede tener grandes aplicaciones favoreciendo la investigación y la evaluación de este colectivo, con un menor número de ítems, pero una fiabilidad similar a la de la versión completa.

Referencias

- Antúnez, Z. y Vinet, E. V. (2012). Escalas de depresión, ansiedad y Estrés (DASS – 21): Validación de la Versión abreviada en Estudiantes Universitarios Chilenos. *Terapia Psicológica, 30*, 49-55.
- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K. y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology, 11*, 443-457.
- Brown, T. A. (2006). Data Issues in CFA: Missing, Non-Normal, and Categorical Data. En Kenny, D.A. (Ed.), *Confirmatory Factor Analysis for Applied Research* (pp. 363-411). New York: The Guilford Press.
- Buela-Casal, G. y Guillén-Riquelme, A. (en prensa). Short version of Spanish adaptation of the State-Trait Anxiety Inventory. *Assessment*.
- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos, N. (2011). *STAI: Cuestionario de Ansiedad Estado-Rasgo. Adaptación española* (8ª ed.). Madrid: TEA Ediciones.
- Delgado-Rico, E., Carretero-Dios, H. y Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology, 12*, 449-460.
- Fioravanti-Bastos, A. C. M., Cheniaux, E. y Landeira-Fernandez, J. (2011). Development and Validation of a Short-Form Version of the Brazilian State-Trait Anxiety Inventory. *Psicologia: Reflexão e Crítica, 24*, 485-494. doi: 10.1590/S0102-79722011000300009

- Guillén-Riquelme, A. y Buela-Casal, A. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23, 510-515.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Kaipper, M. B., Chachamovich, E., Hidalgo, M. P. L., Torres, I. L. S. y Caumo, W. (2010). Evaluation of the structures of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research*, 68, 223-233. doi:10.1016/j.jpsychores.2009.09.013
- Kämpfe, C. K., Gloster, A. T., Wittchen, H-U., Helbig-Lang, S., Lang, T., Gerlach, A. L.,... Deckert, J. (2012). Experiential avoidance and anxiety sensitivity in patients with panic disorder and agoraphobia: Do both constructs measure the same? *International Journal of Clinical and Health Psychology*, 12, 5-22.
- Martínez-López, P., Durá Ferrandis, E., Andreu Vaillo, Y., Galón Garrido, M. J., Murgui Pérez, S. e Ibáñez Guerra, E. (2012). Structural validity and distress screening potential of the Hospital Anxiety and Depression Scale in cancer. *International Journal of Clinical and Health Psychology*, 12, 435-447.
- Marteau, T. M. y Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31, 301-306. doi:10.1111/j.2044-8260.1992.tb00997.x
- Micin, S. y Bagladi, V. (2011). Salud Mental en Estudiantes Universitarios: Incidencia de Psicopatología y Antecedentes de Conducta Suicida en Población que Acude a un Servicio de Salud Estudiantil. *Terapia Psicológica*, 29, 53-64.

- Muñiz, J., Elosua, P. y Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25, 151-157.
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.
- Perpiñá-Galvañ, J., Richart-Martínez, M. y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189. doi:10.1016/j.arbres.2010.11.006
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J. y Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem*, 19, A04:1-A04:6.
- Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48, 1-36.
- Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory: STAI(Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Gorsuch, R. L. y Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- van Knippenberg, F. C. E., Duivendoorn, H. J., Bonke, B. y Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology*, 43, 995-1000. doi:10.1016/0895-4356(90)90083-2

- Vera-Villarroel, P., Celis-Atenas, K., Córdova-Rubio, N., Buela-Casal, G. y Spielberger, C. D. (2007). Preliminary Analysis and Normative Data of the State-Trait Anxiety Inventory (STAI) in Adolescent and Adults of Santiago, Chile. *Terapia Psicológica*, 25, 155-162. doi:10.4067/S0718-48082007000200006
- Watson, D., Clark, L. A. y Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 429-442.
- Yik, M., Russell, J. y Feldman-Barrett, L. (1999). Structure of self-reported current affect: integration and beyond. *Journal of Personality and Social Psychology*, 77, 600-619. doi:10.1037/0022-3514.77.3.600

OBJETIVO 3
FIABILIDAD Y VALIDEZ DEL
STATE-TRAIT ANXIETY INVENTORY
EN MUESTRAS CLÍNICAS

Estudio 5

**Dimensionalidad del *State Trait Anxiety Inventory* en pacientes diagnosticados con
depresión**

Alejandro Guillén-Riquelme y Gualberto Buela-Casal

Manuscrito en proceso de revisión en *Salud Mental*

Dimensionalidad del *State Trait Anxiety Inventory* en pacientes diagnosticados con depresión

Resumen

Introducción: Pese a ser uno de los cuestionarios para evaluar ansiedad más empleados por profesionales de la salud mental, el STAI ha sido objeto de críticas, entre las que destaca la posible existencia de un conjunto de ítems que por evaluar depresión conformarían un factor independiente. El objetivo de este trabajo es evaluar la factorización del STAI en una muestra de pacientes con diagnóstico de depresión.

Material y métodos: Para ello se aplicó la adaptación española del STAI a 266 pacientes españoles, diagnosticados con diferentes trastornos depresivos.

Resultados: Mediante un análisis factorial exploratorio se determinaron tres factores subyacentes: ansiedad estado, ansiedad rasgo positiva y ansiedad rasgo negativa.

Conclusiones: Así pues, la factorización realizada no permite confirmar la presencia de conjuntos de ítems específicos para la depresión, señalada anteriormente como la principal crítica a este cuestionario. Además, los elevados valores del alfa categórico, tanto en la estructura factorial obtenida, como en las subescalas teóricas, son indicios de una elevada fiabilidad para el empleo del STAI en pacientes con diagnóstico de depresión.

Palabras clave: Ansiedad, depresión, depresivos, STAI

Abstract

Introduction: The STAI, despite to be one of the questionnaires to assess anxiety more employees for mental health professionals, has been criticized, the most notably critic is the possible existence of a set of items to assess depression would form an independent factor. The aim of this work is to evaluate the STAI factoring in a sample of patients diagnosed with depression.

Material and methods: We applied the Spanish adaptation of the STAI to 266 Spanish patients diagnosed with various depressive disorders.

Results: In the exploratory factor analysis were identified three underlying factors: state anxiety, trait anxiety positive and negative trait anxiety.

Conclusions: Thus, factorization does not confirm the presence of specific item sets for depression, noted above as the main criticism of this questionnaire. Furthermore, the high values of alpha categorical, in the factor structure obtained, as in the theoretical subscales are highly reliable indications for the use of the STAI in patients diagnosed with depression.

Key words: Anxiety, depression, depressive, STAI.

En 1970 se desarrolla la primera versión comercial del *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch y Lushene, 1970). Este cuestionario ha sido citado en más de 14.000 documentos de archivo y adaptado a más de 60 idiomas (Spielberger y Reheiser, 2009). Entre las diferentes adaptaciones realizadas se encuentra la versión española (Buela-Casal, Guillén-Riquelme y Seisdedos Cubero, 2011). A lo largo de las diferentes investigaciones se ha demostrado que el STAI tiene fiabilidad y validez adecuadas; características que se mantienen en la adaptación española, que cuenta con una fiabilidad elevada, una correcta validez discriminante (Buela-Casal et al., 2011) y ausencia de funcionamiento diferencial en sus ítems (Guillén-Riquelme y Buela-Casal, 2011). Ello es causa de que este cuestionario sea uno de los más empleados por psicólogos clínicos en este país (Muñiz y Fernández-Hermida, 2010) y que recientemente se haya desarrollado una versión breve (Buela-Casal y Guillén-Riquelme, en prensa) validada en varias muestras (Guillén-Riquelme y Buela-Casal, 2013). A pesar de dichas cualidades la factorización presenta un doble problema: 1) la separación en factores diferentes de los ítems que han sido invertidos para la corrección, de los ítems directos; 2) saturaciones de algunos ítems en un factor suelto con confusión entre estado y rasgo.

Diversas investigaciones con el STAI en muestras no españolas, obtienen factorizaciones diferentes, tanto en el número de factores a extraer, como en los ítems que conforman cada factor. De hecho, algunos autores, defienden reducciones factoriales muy robustas con cuatro factores: ansiedad rasgo y ansiedad estado, afirmativo y negativo respectivamente (Hishinuma, Miyamoto, Nishimura y Nahulu, 2000). Otros intentos más complejos señalan que algunos de los ítems inversos generan confusión a la hora de evaluar ansiedad positiva o negativa. El ejemplo más representativo de esta línea de investigación es el trabajo de Vautier (2004), donde se

concluye que diversos ítems del STAI, pese a ser negativos, evalúan tanto presencia como ausencia de ansiedad (tanto estado como rasgo) y provocan los problemas en los análisis factoriales, observados en otros trabajos.

Otros autores destacan el hecho de que el STAI correlaciona de forma notable con medidas de depresión. Uno de los métodos para comprobar esta correlación consistió en unificar los ítems de la ansiedad rasgo con los del cuestionario de depresión de Beck, para aplicar un análisis factorial exploratorio sobre el banco de ítems final. Así pues, Endler, Cox, Parker y Bagby (1992), empleando el STAI completo, obtienen un factor en el que saturan los ítems del cuestionario de Beck y otro con los ítems del STAI y otros ítems de ansiedad, lo que indica que este conjunto de ítems estaría evaluando la misma dimensión teórica. Posteriormente, Andrade, Gorenstein, Vieira Filho, Tung y Artes (2001) replican esta metodología, pero empleando únicamente la subescala de ansiedad rasgo. En este caso se observó que los ítems 1, 10, 15 y 16 de la subescala de ansiedad rasgo saturaban predominantemente en el factor formado por los ítems de depresión. Finalmente se ha estudiado la correlación directa entre el STAI y cuestionarios de depresión, obteniendo correlaciones de 0,45 y superiores (Grös, Antony, Simms y McCabe, 2007) y se defiende que el contenido de algunos de los ítems no evaluaría propiamente ansiedad, lo que constituye el primer paso para establecer la validez de contenido (Delgado-Rico, Carretero-Dios y Ruch, 2012)

Para estudiar los problemas de validez de contenido del STAI, también se empleó el análisis factorial confirmatorio, para establecer si el contenido de algún grupo de ítems estaba más relacionado con diversos factores diferentes a la ansiedad. Bieling, Antony y Swinson (1998), determinaron que, en la subescala de ansiedad rasgo el modelo que mejor ajuste obtuvo es el bifactorial y, a partir del análisis del contenido de los ítems, uno de los factores mediría ansiedad y el otro depresión. Otra de las

propuestas alternativas es que un grupo de ítems permitiría evaluar afecto negativo. Siguiendo un procedimiento similar otros autores, defienden la existencia de una subescala de depresión (Bados, Gómez-Benito y Balaguer, 2010; Caci, Baylé, Dossios, Robert y Boyer, 2003) o de afecto negativo general (Hill, Musso, Jones, Pella y Gouvier, 2012) entendiendo éste como un componente característico, tanto de los trastornos depresivos como de los ansiosos (Watson, Clark y Stasik, 2011). Otros autores defienden, pese a obtener una estructura unifactorial, que el STAI rasgo estaría evaluando afecto general negativo (Balsamo et al., 2013).

Pese a las críticas recibidas, cabe destacar que la mayoría de los trabajos emplean únicamente la subescala de ansiedad rasgo y utilizan muestras no clínicas. Ante la diversidad de modelos, del número de factores y de las combinaciones de ítems defendidos, resulta difícil establecer una conclusión clara, por lo cual, el objetivo de la presente investigación es analizar la factorización de STAI con una muestra de pacientes depresivos españoles.

Método

Participantes

En el presente estudio participaron 266 pacientes diagnosticados con algún trastorno de depresión, 182 mujeres (68,4%) y 84 hombres (31,6%). La muestra procede de diez ciudades españolas. Las edades oscilaron de los 15 a los 63 años, siendo la edad media de 33,51 ($DT = 11,2$). En el grupo de mujeres la edad media fue de 31,9 años ($DT = 10,63$) y en el de hombres de 36,99 ($DT = 11,66$). En la frecuencia por el tipo de trastorno hay 191 pacientes (71,8% del total) con diagnóstico principal de depresión; 64 (24%) con distimia; 9 (3,4%) con trastorno mixto de depresión y ansiedad y 2 (0,8%) con trastorno maniaco-depresivo.

Instrumento

Para lograr los objetivos del estudio se aplicó la adaptación española del *State-Trait Anxiety Inventory* (Spielberger et al., 1970) desarrollada por Buela-Casal et al., (2011). Este cuestionario evalúa la ansiedad estado y la ansiedad rasgo, mediante 20 ítems, para cada una, con escala de respuesta tipo Likert de 4 alternativas. En el caso de la ansiedad estado la escala oscila desde 0 (*Nada*) a 3 (*Mucho*), mientras que en la ansiedad rasgo se corresponde con 0 (*Casi nunca*) a 3 (*Casi siempre*). Tanto en la ansiedad rasgo como estado un porcentaje de los ítems están invertidos y evalúan bienestar o ausencia de ansiedad, mientras que el resto de los ítems se refieren a la presencia de ansiedad. La puntuación total se obtiene mediante la suma de los ítems, tras la inversión de los que están redactados en positivo.

Además del STAI, los clínicos respondieron a tres cuestiones: el diagnóstico principal, el criterio para su evaluación y el tiempo de tratamiento. Estas cuestiones se presentaban de forma separada del cuadernillo del paciente.

Procedimiento

Para realizar esta investigación ex-post facto con un solo grupo (Montero y León, 2007) en primer lugar, se redactó la hoja de respuesta sobre el diagnóstico tal y como se ha descrito arriba; en segundo lugar, se contactó con diversos clínicos de diez ciudades españolas para presentar el estudio y se les proporcionó las hojas de respuesta del STAI. Los clínicos seleccionaron a aquellos pacientes que habían sido diagnosticados previamente con algún trastorno de depresión, para aplicarle el cuadernillo. Los criterios de exclusión era no estar diagnosticados como diagnóstico principal de uno de los trastornos de depresión y ser menor de edad. Los cuestionarios iban acompañados por unas instrucciones por escrito, para garantizar la uniformidad de la aplicación. En primer lugar el clínico explicaba el consentimiento informado a los

pacientes, con el fin de que éstos entendiesen las garantías de su participación. El paciente podía cumplimentar el cuadernillo, junto con los datos sociodemográficos, en la propia consulta o fuera de ella, y en este caso lo entregaba en la siguiente sesión. La hoja de información diagnóstica y el cuadernillo con el STAI y los sociodemográficos se cumplimentaban en documentos separados, para evitar que el paciente pudiese acceder a la información diagnóstica, en el caso de que el terapeuta así lo deseara. Para evitar la confusión de ambos documentos entre pacientes se estableció un código duplicado en: a) cuadernillo y b) hoja de diagnóstico. Dicho código, a cargo del personal sanitario, garantizaba el anonimato del paciente respecto al examinador e impedía el cambio involuntario de las hojas entre pacientes. Finalmente, los clínicos enviaron los documentos al experimentador para comenzar la corrección y el pase a la base de datos. Para la redacción del manuscrito se siguieron las recomendaciones propuestas por Hartley (2012).

Análisis de datos

El primer paso fue realizar un análisis factorial exploratorio. Para la retención del número de factores se realizó mediante un análisis paralelo. Este análisis consiste en determinar unos eigenvalores, calculados a partir de una matriz generada de forma aleatoria. Estos valores son comparados con los resultantes del análisis factorial. Se mantendrá en el análisis factorial aquellos factores cuyo eigenvalor (superior a uno según el criterio de Kayser) sea también superior al eigenvalor obtenido con la matriz generada de forma aleatoria (Patil, Singh, Mishra y Donovan, 2008; Timmerman y Lorenzo-Seva, 2011). El procedimiento de extracción fue de ejes principales y el método de rotación fue varimax, ya que ambas subescalas son teóricamente independientes.

Para el análisis de la fiabilidad se empleó el alfa categórico, calculado a partir de la matriz de correlaciones policóricas, tal y como se especifica en Elosua Oliden y Zumbo (2008).

Resultados

Primeramente se calculó la puntuación media para ambas subescalas y se comparó si era estadísticamente superior al valor medio establecido en el manual de la adaptación española para población general. En el caso de la ansiedad estado la media fue de 32,26 ($DT = 14,4$). Al calcular las medias por sexo, en el grupo de hombres fue de 32,85 ($DT = 14,3$; comparación con la media de la adaptación española: $t(83) = 10,64$; $p < 0,001$), y en el de mujeres de 31,98 ($DT = 14,48$; comparación con la media de la adaptación española: $t(181) = 12,73$; $p < 0,001$). En el caso de la ansiedad rasgo la media fue de 36,41 ($DT = 11,96$). Al realizar los análisis por la media del grupo de hombres fue de 36,39 ($DT = 13,89$; comparación con la media de la adaptación española: $t(83) = 10,64$; $p < 0,001$), en el de mujeres fue de 36,42 ($DT = 12,2$; comparación con la media de la adaptación española: $t(181) = 14,43$; $p < 0,001$).

En segundo lugar se realizó la media para cada uno de los trastornos depresivos (excluyendo el trastorno maniaco-depresivo, ya que únicamente había dos sujetos con dicho trastorno). En el caso de personas diagnosticadas con depresión la media de ansiedad estado fue de 31,92 ($DT = 14,40$) y para la ansiedad rasgo de 35,84 ($DT = 11,98$). En personas con diagnóstico de distimia la media de ansiedad estado fue de 32,7 ($DT = 15,13$) y en ansiedad rasgo fue de 38,77 ($DT = 12,11$). En el grupo diagnosticado con trastorno mixto de ansiedad y depresión la media para la ansiedad estado fue de 35,78 ($DT = 15,43$) y para la ansiedad rasgo de 36,67 ($DT = 9,53$). Se realizó un ANOVA, tras comprobar los supuestos, para comparar si las puntuaciones medias diferían según el tipo de trastorno, no encontrando diferencias significativas ni en

ansiedad estado ($F(2, 254) = 0,34$; $p = 0,713$) ni en ansiedad rasgo ($F(2, 254) = 1,42$; $p = 0,244$).

Seguidamente se realizó un análisis factorial. En primer lugar, se comprobó la adecuación de la matriz para el análisis, obteniendo adecuados índices ($\chi^2(780; 266) = 6.236,99$; $KMO = 0,936$). Mediante el análisis paralelo se determinó que se debía extraer tres factores. Al realizar el análisis factorial se observó que los tres factores explicaron un 52,26% de la varianza total explicada. Al observar las comunalidades, salvo en el caso del ítem 7 de la subescala de ansiedad rasgo, todos los ítems tuvieron valores superiores a 0,25. Tras ello se analizó las saturaciones de los ítems en cada uno de los factores. Un resumen de las saturaciones se recoge en la tabla 1.

Por último, se realizó un análisis de fiabilidad. En el caso de los 40 ítems de forma conjunta el alfa fue de 0,936. En este caso únicamente el ítem 7 de la subescala de ansiedad rasgo mejoraba este valor si era eliminado. En el caso de la ansiedad estado el alfa era igual a 0,964. En el caso de la ansiedad rasgo la fiabilidad fue de 0,927, se observó, de nuevo, que la eliminación del ítem 7 mejoraba este valor (en tres milésimas). Tras estos análisis se calculó la fiabilidad para la factorización obtenida mediante análisis factorial. Así pues, en el primer factor el alfa fue de 0,96 (el ítem 7 de la subescala de estado mejoraba el alfa en una milésima). En el segundo factor el alfa fue de 0,889 (la eliminación del ítem 19 mejoraba una milésima ese valor) y en el tercero de los factores el alfa fue de 0,902, no habiendo ítems cuya eliminación mejorase la consistencia interna.

Tabla 1.
 Matriz de saturaciones de los ítems del STAI en
 cada uno de los factores del análisis factorial

Ítem	Factor		
	1	2	3
Estado3	0,805		
Estado12	0,791		
Estado6	0,776		
Estado4	0,766		
Estado18	0,719		
Estado14	0,661	0,317	
Estado1	0,654	0,365	
Estado15	0,637	0,488	
Estado9	0,629		
Estado13	0,564		
Estado17	0,559		0,370
Estado5	0,539	0,505	
Estado8	0,534	0,470	
Estado10	0,511	0,510	
Rasgo16		0,740	
Rasgo10		0,734	
Rasgo1		0,709	
Estado19	0,378	0,678	
Estado16	0,403	0,671	
Estado11		0,651	
Rasgo13		0,645	
Estado2	0,455	0,615	
Estado20	0,539	0,589	
Rasgo6		0,512	
Rasgo15		0,488	0,427
Rasgo19		0,442	
Rasgo4		0,434	0,423
Rasgo7			
Rasgo9			0,667
Rasgo20			0,627
Rasgo17			0,567
Rasgo8		0,398	0,558
Rasgo11			0,540
Rasgo18			0,475
Estado7	0,414		0,442
Rasgo12		0,421	0,434
Rasgo14			0,432
Rasgo3		0,372	0,387
Rasgo2		0,317	0,356
Rasgo5			0,335

Nota. Se eliminaron las saturaciones inferiores a 0,3. La negrita indica el factor en el que se sitúa el ítem.

Discusión

Mediante estos datos se ha establecido que los depresivos tienen puntuaciones elevadas en lo que respecta a los niveles de ansiedad evaluados mediante el STAI, siendo unas puntuaciones medias significativamente diferentes a los valores obtenidos en la adaptación española del cuestionario (Buela-Casal et al., 2011). Las diferencias están entre los 15 y los 20 puntos, lo cual coincide con otros estudios donde se compara la puntuaciones medias de pacientes con trastornos depresivos (depresión mayor y trastorno mixto de ansiedad y depresión) observando puntuaciones medias dentro de ese rango (v.gr. Kennedy, Schwab, Morris y Beldia, 2001). Este hecho se puede deber a dos causas: que los trastornos de ansiedad y depresión tienen una alta comorbilidad entre sí (Beutel, Bleichner, von Heymann, Tritt y Hardt, 2011; Lamers et al., 2011), siendo común observar correlaciones entre las puntuaciones de los instrumentos de evaluación de la ansiedad con los de depresión (Kämpfe et al., 2012). La segunda causa de esos 15-20 puntos de diferencia sería que algunos de los ítems del STAI evalúan depresión o malestar general (Bados et al., 2010; Bieling et al., 1998; Hill et al., 2012).

A través del procedimiento empleado en el presente artículo no se puede especificar categóricamente si las puntuaciones medias de los pacientes con depresión se deben a que también sufren ansiedad o a que el STAI evalúa depresión o malestar general. Por ello, se decidió analizar la estructura factorial subyacente a los datos para verificar la similitud entre los factores extraídos de dicho análisis y los defendidos por quienes consideran que el STAI tiene subescalas de depresión, malestar general... A partir de los resultados se determinó la existencia de tres factores subyacentes. En ellos, pese a que hay 4 de los 40 ítems que no encajan en su factor teórico, la estructura obtenida es un modelo trifactorial, donde el primero de los factores se corresponde con la subescala de la ansiedad rasgo. La subescala de ansiedad estado se divide en dos

factores: uno que aglutina a los ítems que han sido invertidos (factor 2 en la tabla 1) y el otro a los ítems positivos (factor 3 en la tabla 1). Así pues, empleando una muestra de pacientes depresivos, no se observa una agrupación estadística de los ítems que miden predominantemente depresión u otros factores diferentes a la ansiedad rasgo.

En resumen, al observar la factorización obtenida en la muestra de depresivos no se detecta agrupación alguna de los ítems que teóricamente miden depresión (Bados et al., 2010; Bieling et al., 1998) o afecto negativo (Hill et al., 2012, Reiss, 1997). Sin embargo, tanto en las saturaciones como en el caso de los análisis de fiabilidad hay cinco ítems en los que se observan problemas en los análisis, siendo el ítem siete de la subescala de ansiedad rasgo el más conflictivo. En este caso este ítem no satura por encima de 0,3 en ninguno de los factores teóricos y su eliminación mejora el alfa de Cronbach en todos los factores de los que forma parte. El contenido del ítem siete (“Soy una persona tranquila, serena y sosegada”) hace referencia a que la persona normalmente esté calmada, lo que se vincula principalmente con ausencia de ansiedad, pero podría estar relacionado con el bien estar psicológico. El otro problema de este ítem es que al tener tres adjetivos, éstos pueden confundir al sujeto evaluado, provocando que sólo tenga en cuenta uno de ellos o incluso una respuesta aleatoria sin conexión alguna con el contenido del ítem.

En lo que respecta a la fiabilidad se puede observar que, en todos los casos, los valores del alfa categórico son excelentes. Este es un dato muy relevante ya que implica que en pacientes depresivos el STAI es un instrumento fiable para detectar altos niveles de ansiedad comórbida con el trastorno. Este dato tiene sentido ya que el STAI muestra adecuadas medidas de consistencia interna tanto en la versión original (Spielberger et al., 1970; Barnes, Harp y Sik, 2002) como en la adaptación española empleando población general (Buela-Casal et al., 2011; Guillén-Riquelme y Buela-Casal, 2011).

Además, en otros trabajos con muestras españolas donde se evalúa la fiabilidad de la versión breve del instrumento en pacientes españolas con intubación respiratoria la fiabilidad es excelente (Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011)

En este artículo se aporta una nueva evidencia del correcto funcionamiento del STAI en pacientes depresivos, aun cuando, en esta población, las puntuaciones suelen ser más elevadas que en personas sin trastornos, lo que exige una prudencia extrema en su empleo sobre estos pacientes. Respecto a si esto se debe a que el STAI mide depresión o malestar general, la factorización realizada encaja con las estructuras teóricas de la versión original (Spielberger et al., 1970), en la adaptación española (Buela-Casal et al., 2011) y en estudios similares (v. gr. Hishinuma et al., 2000). Así pues, el presente documento aporta la evidencia de que su uso en pacientes con diagnóstico de depresión es adecuado y ofrece una nueva evidencia de su validez de constructo. Pese a ello el trabajo conlleva diversas limitaciones, la más relevante la descompensación entre los grupos de pacientes en función del tipo de trastorno, lo que impide comparar el funcionamiento del STAI para cada uno de ellos. Futuras investigaciones deberán abordar el alcance de estas diferencias, para aportar las evidencias necesarias y determinar si la adaptación española del STAI tiene ítems que evalúan constructos diferentes a la ansiedad o no. Pese a las limitaciones descritas, la estructura subyacente de la adaptación española del STAI obtenida en la presente investigación, no implica la existencia de agrupaciones de ítems cuyo contenido haga suponer que se evalúe depresión, afecto negativo general... Este hecho ayudará a los clínicos a la hora de aplicar el cuestionario y llevar a cabo su práctica profesional.

Referencias

- Andrade, L., Gorenstein, C., Vieira Filho, A. H., Tung, T. C. y Artes, R. (2001). Psychometric properties of the Portuguese version of the State-Trait Anxiety Inventory applied to college students: factor analysis and relation to the Beck Depression Inventory. *Brazilian Journal of Medical and Biological Research*, 34, 367-374. doi:10.1590/S0100-879X2001000300011
- Bados, A., Gómez-Benito, J. y Balaguer, J. (2010). The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety? *Journal of Personality Assessment*, 92, 560-567. doi: 10.1080/00223891.2010.513295
- Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L. y Saggino, A. (2013). The State-Trait Anxiety Inventory: Shadows and Lights on its Construct Validity. *Journal of Psychopathology and Behavioral Assessment*, descargado el 13 de agosto de 2013 de <http://link.springer.com/article/10.1007/s10862-013-9354-5>. doi:10.1007/s10862-013-9354-5
- Barnes, L., Harp, D. y Sik, W. (2002). Reliability Generalization of Scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618. doi: 10.1177/0013164402062004005
- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K. y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology*, 11, 443-457.
- Bieling, P. J., Antony, M. M. y Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behavior Research and Therapy*, 36, 777-788. doi: 10.1016/S0005-7967(98)00023-0

- Buela-Casal, G. y Guillén-Riquelme, A. (en prensa). Short form of the Spanish adaptation of the State-Trait Anxiety Inventory. *Assessment*.
- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos Cubero, N. (2011). *Cuestionario de ansiedad estado-rasgo* (8ª edición). Madrid: TEA ediciones.
- Caci, H., Baylé, F. J., Dossios, C., Robert, P. y Boyer, P. (2003). The Spielberger trait anxiety inventory measures more than anxiety. *European Psychiatry*, 18, 394-400. doi:10.1016/j.eurpsy.2003.05.003
- Delgado-Rico, E., Carretero-Dios, H. y Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology*, 12, 449-460.
- Elosua Oliden, P. y Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, 20, 896-901.
- Endler, N. S., Cox, B. J., Parker, J. D. A. y Bagby, R. M. (1992). Self-Reports of Depression and State-Trait Anxiety: Evidence for Differential Assessment. *Journal of Personality and Social Psychology*, 63, 832-838. doi:10.1037/0022-3514.63.5.832
- Grös, D. F., Antony, M. M., Simms, L. J. y McCabe, R. E. (2007). Psychometric Properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment*, 19, 369-381. doi: 10.1037/1040-3590.19.4.369
- Guillén-Riquelme, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23, 510-515.

- Guillén-Riquelme, A. y Buela-Casal, A. (2013). Versión breve del STAI en adolescentes y universitarios españoles. *Terapia Psicológica*, 31, 293-299.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Hill, B. D., Musso, M., Jones, G. N., Pella, R. D. y Gouvier Wm. D. (2012). A Psychometric Evaluation of the STAI-Y, BDI-II, and PAI Using Single and Multifactorial Models in Young Adults Seeking Psychoeducational Evaluation. *Journal of Psychoeducational Assessment*, XX, 1-13. doi: 10.1177/0734282912462670
- Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T. y Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory Scores for Ethnically Diverse Adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83. doi:10.1037/1099-9809.6.1.73
- Kämpfe, C. K., Gloster, A. T., Wittchen, H.-U., Helbig.Lang, S., Lang, T., Gerlach, A. L.,... Deckert, J. (2012). Experimental avoidance and anxiety sensitivity in patients with panic disorder and agoraphobia: Do both constructs measure the same? *International Journal of Clinical and Health Psychology*, 12, 5-22.
- Kennedy, B. L., Schwab, J. J., Morris, R. L. y Beldia, G. (2001). Assessment of the state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatric Quarterly*, 72, 263-276. doi:10.1023/A:1010305200087
- Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J., ... Penninx, B. W. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and

- Anxiety (NESDA). *Journal of Clinical Psychiatry*, 72, 341-348. doi:10.4088/JCP.10m06176blu
- Montero, I. y León, O. G. (2007) A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862.
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.
- Patil, V. H., Singh, S. N., Mishra, S. y Donovan, D. T. (2008). Efficient Theory Development and Factor Retention Criteria: A Case for Abandoning the 'Eigenvalue Greater Than One' Criterion. *Journal of Business Research*, 61, 162-170. doi:10.1016/j.jbusres.2007.05.008
- Perpiñá-Galvañ, J., Richart-Martínez, M. y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189. doi:10.1016/j.arbres.2010.11.006
- Reiss, S. (1997). Trait Anxiety: it's not what you think it is. *Journal of Anxiety Disorders*, 11, 201-214. doi:10.1016/S0887-6185(97)00006-6
- Spielberger, C. D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- Timmerman, M. E. y Lorenzo-Seva, U. (2011). Dimensionality Assessment of Ordered Polytomous Items with Parallel Analysis. *Psychological Methods*, 16, 209-220. doi:10.1037/a0023353

Vautier, S. (2004). A Longitudinal SEM Approach to STAI Data: Two Comprehensive Multitrait-Multistate Models. *Journal of Personality Assessment*, 83, 167-179.
doi:10.1207/s15327752jpa8302_11

Watson, D., Clark, L. A. y Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 429-442.

Estudio 6

**Fiabilidad y validez del *State-Trait Anxiety Inventory* en pacientes diagnosticados
con ansiedad**

Gualberto Buela-Casal, Alejandro Guillén-Riquelme y Carmen T. Pitti

Manuscrito en proceso de revisión en *Depression and Anxiety*

Fiabilidad y validez del *State-Trait Anxiety Inventory* en pacientes diagnosticados con ansiedad

Resumen

Introducción: La adaptación española del STAI ha demostrado ser fiable y válida en población general española, pero no ha sido validada para pacientes con diagnóstico de ansiedad. Esta carencia lleva al objetivo de analizar los ítems, la fiabilidad y la validez de constructo en pacientes con diagnóstico de ansiedad. Método: Para lograrlo se aplicaron, junto al STAI, otros tres instrumentos (BAI, IDER y CBD) a una muestra de 251 pacientes con diagnóstico principal de ansiedad. Resultados: Las medias de la ansiedad rasgo y estado son elevadas, siendo significativamente diferentes del valor de la adaptación española para población normal. La fiabilidad es superior a 0,9 en ambas subescalas. Hay relaciones moderadas entre el STAI y el BAI, IDER y CBD. Discusión: En la muestra de pacientes con ansiedad, la correlación del STAI con el BAI es indicio de que la validez es adecuada, sin embargo las correlaciones con los instrumentos de depresión, implican que se puede estar evaluando otros constructos diferentes. Además, el STAI es fiable en pacientes con ansiedad.

Palabras clave: Ansiedad; ansiosos; STAI; psicometría; fiabilidad; validez.

Abstract

Introduction: The Spanish adaptation of the STAI has proven reliable and valid in Spanish general population, however has not been validated for patients diagnosed with anxiety. The objective is to analyze the items the reliability and validity of constructs in patients diagnosed with anxiety. Method: To achieve the objective the STAI and several

questionnaires of anxiety and depression were applied in a sample of 251 patients with a diagnosis of anxiety. Results: Means to state and trait anxiety are high, being significantly different from the value of the Spanish adaptation to normal population. Reliability is above 0.9 on both scales. There are moderate relationships between the STAI and BAI, IDER and CBD. Discussion: In anxiety patients sample, the correlation between STAI and BAI is an indication that the validity is adequate, however the correlations with depression instruments imply that may be evaluating different constructs. In addition, the STAI is reliable in patients with anxiety.

Keywords: Anxiety; anxious; STAI; psychometric; reliability; validity.

El *State-Trait Anxiety Inventory* (STAI; Spielberger, Gorsuch y Lushene, 1970) fue adaptado en España, en 1982, siendo revisadas sus propiedades psicométricas por Buela-Casal, Guillén-Riquelme y Seisdedos Cubero (2011). La adaptación española del STAI es uno de los cuestionarios más usados en este país, siendo el sexto más empleado por psicólogos clínicos y el séptimo por psicólogos sin tener en cuenta la especialización (Muñiz y Fernández-Hermida, 2010). Esa preferencia cuantitativamente notable de los psicólogos españoles se explica porque dicha adaptación es sencilla de aplicar, corregir e interpretar para sus dos constructos: ansiedad estado y ansiedad rasgo. Además, la consistencia interna del STAI en población general es elevada tanto para ansiedad rasgo como para ansiedad estado (Barnes, Harp y Jung, 2002; Rossi y Pourtois, 2012), y lo mismo sucede en población general española (Buela-Casal et al., 2011; Guillén-Riquelme y Buela-Casal, 2011). A estas características que favorecen su difusión se añade que las correlaciones con medidas de ansiedad son elevadas (Spielberger y Reheier, 2009).

Una vez confirmada la fiabilidad y validez de la adaptación española del STAI en población general, se estudió su comportamiento en otras muestras, con características similares a las de población general, por ejemplo, universitarios (Guillén-Riquelme y Buela-Casal, 2013). Paralelamente, se ha estudiado el establecimiento de una versión breve del STAI para la evaluación de la ansiedad estado y rasgo (Buela-Casal y Guillén-Riquelme, en prensa; Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez y Martínez-Durá, 2011), en población general, pacientes intubados, adolescentes y universitarios (Guillén-Riquelme y Buela-Casal, 2013).

La versión original del STAI ha sido validada en pacientes con ansiedad (Spielberger y Reheier, 2009). Así, en un meta-análisis, centrado en su uso en esta

población, se concluyó que sus puntuaciones eran significativamente superiores a los valores en muestras control sanas (Guillén-Riquelme y Buela-Casal, 2014). Además, en sujetos con ansiedad el alfa de Cronbach fue de 0,92 en ansiedad estado y 0,91 en ansiedad rasgo. Contrariamente, en la adaptación española no se ha demostrado su validez para ese grupo de pacientes. De hecho, ni siquiera se informa de su fiabilidad en estudios donde se aplica este cuestionario a pacientes con diagnóstico de ansiedad (Guillén-Riquelme y Buela-Casal, 2014).

Pese a que en el STAI se observa una adecuada fiabilidad y que su relación con las medidas de ansiedad es elevada, diversos autores han criticado, que este cuestionario correlaciona con instrumentos de depresión (Grös, Antony, Simms y McCabe, 2007) o presenta saturaciones factoriales que permiten establecer un factor de depresión (Bados, Gómez-Benito y Balaguer, 2010; Bieling, Antony y Swinson, 1998). En el caso de los pacientes con diagnóstico de ansiedad, debido a la comorbilidad entre la ansiedad y la depresión, estos problemas pueden verse acentuados. No obstante, no se ha encontrado ninguna investigación que estudie la relación entre las puntuaciones del STAI y de otros cuestionarios de depresión, en muestras de pacientes con diagnóstico de ansiedad.

Así, frente a la validez contrastada del STAI y su adaptación española en población general, se ha señalado la existencia de determinadas críticas respecto a su idoneidad en pacientes diagnosticados con ansiedad; por lo que el objetivo de la presente investigación es analizar la fiabilidad y validez de constructo de la adaptación española del STAI en una muestra de pacientes con diagnóstico principal de ansiedad. Los objetivos específicos son analizar: 1) consistencia interna del STAI; 2) la estructura factorial y 3) la validez de constructo. Las hipótesis de partida son: 1) la fiabilidad de ambas subescalas del STAI (ansiedad estado y ansiedad rasgo) serán iguales o superiores a 0,9; 2) la estructura factorial tanto de la ansiedad estado como de la

ansiedad rasgo será unifactorial y 3) el STAI correlacionará de forma elevada con una media de ansiedad y de forma moderada o baja con medidas de depresión.

Método

Participantes

En el presente estudio participaron 247 pacientes con algún trastorno de ansiedad y habitantes en nueve comunidades autónomas de España (Andalucía, Asturias, Baleares, Canarias, Cataluña, Extremadura, Madrid, Murcia y Valencia). Del total de la muestra, 157 participantes (62,55% del total) son mujeres. La edad media fue de 35,93 ($DT = 11,83$). En la muestra aparecen los siguientes 12 trastornos: ansiedad generalizada (35,46% del total), crisis de angustia (12,35%), estrés agudo (11,95%), trastorno obsesivo-compulsivo (9,96%), angustia sin agorafobia (6,77%), angustia con agorafobia (5,18%), fobia específica (5,18%), agorafobia (3,98%), fobia social (3,19%), de estrés post-traumático (3,19%), de ansiedad debido a enfermedad médica (2,39%) y de ansiedad inducido por sustancias (0,40%).

Instrumentos

Para lograr tanto el objetivo general como los específicos se aplicaron cuatro instrumentos de ansiedad y depresión:

- *Cuestionario de Ansiedad Estado y Rasgo (State-Trait Anxiety Inventory; STAI; Spielberger et al., 1970)*: Este cuestionario fue adaptado en España en 1982, revisando este trabajo en 2011, momento en el que se modificaron los baremos de dicha adaptación (Buela-Casal et al., 2011). Está formado por 20 ítems para la evaluación de la ansiedad estado y 20 para la ansiedad rasgo. La escala de respuesta oscila de 0 a 3 en todos los ítems y la categoría de respuesta va, en el caso de la ansiedad estado, de “nada” a “mucho” y en la ansiedad rasgo de “casi nunca” a “casi siempre”. El total de

cada una de las subescalas se obtiene mediante la suma de los 20 ítems y oscila de 0 a 60.

- *Inventario de Ansiedad de Beck* (BAI; Beck, Epstein, Brown y Steer, 1988): La adaptación española fue llevada a cabo por Sanz, Vallar, De la Guía y Hernández (2011). Este cuestionario permite evaluar ansiedad general por medio de ítems con una escala de 4 puntos que oscila desde 0 (“nada en absoluto”) hasta 3 (“gravemente”). La puntuación total del cuestionario queda comprendida entre 0 y 63, correspondiéndose a una mayor puntuación mayores niveles de ansiedad.

- *Inventario de Depresión Estado-Rasgo* (IDER; Spielberger y Ritterband, 1996): La adaptación española fue realizada por Buela-Casal y Agudelo Vélez (2008). Este cuestionario cuenta con 20 ítems (10 para evaluar depresión estado y 10 para depresión rasgo). El cuestionario tiene tres opciones de respuesta desde 0 hasta 3, abarcando para la depresión estado desde “nada” hasta “mucho” y para la depresión rasgo desde “nunca” hasta “siempre”. El total de ambas subescalas oscila de 0 a 30.

- *Cuestionario Básico de Depresión* (CBD; Peñate, 2001): Está formado por 21 síntomas específicos de la depresión, a los que la persona tiene que responder cuándo los ha sentido siendo las opciones de respuesta “nunca” (0 puntos), “semanas” (1), “meses” (2) y “años” (3). La puntuación total abarca desde 0 a 63, correspondiéndose a una mayor puntuación mayores niveles de depresión.

Procedimiento

El establecimiento del objetivo fue el primer paso, tras lo cual se determinaron los cuestionarios necesarios para cumplir el mismo. Seguidamente, se contactó con diversos psicólogos y psiquiatras para presentar el estudio. Tras informar del procedimiento de aplicación a los clínicos se reclutó a los pacientes. Los criterios de inclusión eran que los pacientes fuesen mayores de edad y que estuviesen

diagnosticados con uno de los trastornos de ansiedad, recogidos en el DSM-IV-TR (American Psychological Association, 2002), como diagnóstico principal. Así pues, cuando los clínicos diagnosticaban a un paciente con uno de los trastornos de ansiedad, le presentaban el estudio y le informaban de la confidencialidad del mismo. En el caso de que el paciente decidiese participar, el clínico le proporcionaba el cuadernillo con los cuestionarios seleccionados y cinco preguntas sociodemográficas (sexo, edad, lugar de residencia, nivel de estudios y estado civil). Además de ello, los clínicos añadían un código a los cuadernillos de los pacientes, para vincularlos con una hoja de diagnóstico donde se informaba del diagnóstico principal, secundario, el tipo de tratamiento y si el paciente tomaba ansiolíticos. Este procedimiento mediante códigos tiene dos ventajas: 1) garantizar que solo el clínico pudiese vincular las respuestas de los cuadernillos con los datos del participante; 2) que los pacientes no viesen la hoja de diagnóstico en aquellos casos en los que se deseaba que así fuera. Posteriormente, en el centro de investigación se introducían en la base de datos para los análisis. La redacción del artículo se realizó siguiendo las recomendaciones propuestas por Hartley (2012).

Análisis estadísticos

En primer lugar, para el análisis de ítems se empleó la frecuencia de elección de las categorías de respuesta para cada uno de los ítems. Respecto a la fiabilidad se calculó el alfa ordinal para variables categóricas, debido a que los ítems del STAI son de tipo Likert con cuatro opciones de respuesta. Este procedimiento es el más indicado con este tipo de datos (Elosua Oriden y Zumbo, 2008; Gadermann, Guhn y Zumbo, 2012). El análisis de validez se realizó mediante dos pruebas, la primera de ellas un análisis factorial exploratorio, empleando *máxima verosimilitud* y rotación *oblimin*; el criterio para determinar el número de factores fue mediante análisis paralelo (Cho, Li y Bandalos, 2009). Dicho análisis se llevó a cabo de forma independiente para ansiedad

estado y rasgo. En segundo lugar, se realizaron correlaciones de Pearson entre medidas similares y relacionadas para evaluar la validez de contenido del STAI.

Resultados

En primer lugar, se calculó el índice de atracción para cada uno de los ítems del STAI que se recoge en la tabla 1, donde se puede observar que hay una mayor frecuencia en las categorías intermedias que en las extremas, donde las frecuencias en algunos ítems son muy bajas (llegando a un porcentaje del 0,8 en el ítem 11 de la subescala de ansiedad rasgo). En algunos casos (ítems 2 y 18 de estado y 1, 15 y 16 de rasgo) hay categorías elegidas por más del 45% de la muestra.

En segundo lugar, se realizó un análisis factorial exploratorio. Dentro de esta técnica se empleó el análisis paralelo para determinar el número de factores a extraer, y se estableció que cada una de las subescalas del STAI estaba compuesta por dos factores. En la ansiedad estado el porcentaje de varianza total explicada es del 65%; mientras que en la ansiedad rasgo es del 50%. En la tabla 2 se presentan las saturaciones rotadas en cada uno de los factores, donde se observa que, en la mayoría de los ítems se satura por encima de 0,3 en ambos y, también, que hay cierta influencia de los ítems inversos en el segundo de los factores.

En tercer lugar, se realizó la correlación de Pearson entre la ansiedad estado y la ansiedad rasgo, obteniendo un valor de 0,57 ($p < 0,001$), lo que indica una correlación moderada entre ambos factores (ver tabla 3). La puntuación media en el caso de la ansiedad estado es de 27,53 ($DT = 13,56$) y para la ansiedad rasgo de 32,58 ($DT = 12,09$). Las medias para los varones fueron de 27,53 ($DT = 13,56$) en ansiedad estado, siendo significativamente mayores que el valor de la adaptación española para el mismo grupo ($t(90) = 7,93$; $p < 0,001$) y de 32,58 ($DT = 12,09$; $t(90) = 11,26$; $p < 0,001$). En el

Tabla1.
Frecuencia (%) de respuesta a cada categoría de respuesta de los ítems de STAI.

Ítem	Categoría de respuesta			
	0	1	2	3
Estado 1*	20 (8,1)	78 (31,7)	113 (45,9)	35 (14,2)
Estado 2*	21 (8,5)	79 (32)	101 (40,9)	46 (18,6)
Estado 3	65 (26,6)	89 (36,5)	60 (24,6)	30 (12,3)
Estado 4	87 (36,4)	71 (29,7)	45 (18,8)	36 (15,1)
Estado 5*	23 (9,3)	76 (30,8)	94 (38,1)	54 (21,9)
Estado 6	95 (39,1)	68 (28)	50 (20,6)	30 (12,3)
Estado 7	54 (21,9)	70 (28,3)	71 (28,7)	52 (21,1)
Estado 8*	22 (8,9)	67 (27,1)	79 (32)	79 (32)
Estado 9	79 (32)	78 (31,6)	55 (22,3)	35 (14,2)
Estado 10*	19 (7,7)	64 (26)	105 (42,7)	58 (23,6)
Estado 11*	18 (7,3)	86 (34,8)	101 (40,9)	42 (17)
Estado 12	65 (26,3)	80 (32,4)	70 (28,3)	32 (13)
Estado 13	56 (23,6)	98 (41,4)	58 (24,5)	25 (10,5)
Estado 14	102 (41,3)	63 (25,5)	41 (16,6)	41 (16,6)
Estado 15*	14 (5,7)	70 (28,5)	88 (35,8)	74 (30,1)
Estado 16*	13 (5,3)	76 (31)	95 (38,8)	61 (24,9)
Estado 17	39 (15,8)	78 (31,6)	73 (29,6)	57 (23,1)
Estado 18	120 (48,6)	59 (23,9)	46 (18,6)	22 (8,9)
Estado 19*	8 (3,3)	64 (26,1)	105 (42,9)	68 (27,8)
Estado 20*	18 (7,3)	80 (32,5)	100 (40,7)	48 (19,5)
Rasgo 1*	21 (8,5)	84 (34,1)	112 (45,5)	29 (11,8)
Rasgo 2	55 (22,4)	81 (32,9)	74 (30,1)	36 (14,6)
Rasgo 3	54 (22)	109 (44,5)	58 (23,7)	24 (9,8)
Rasgo 4	32 (13,1)	62 (25,4)	66 (27)	84 (34,4)
Rasgo 5	65 (26,3)	99 (40,1)	48 (19,4)	35 (14,2)
Rasgo 6*	16 (6,5)	58 (23,7)	105 (42,9)	66 (26,9)
Rasgo 7*	25 (10,2)	49 (20,1)	91 (37,3)	79 (32,4)
Rasgo 8	39 (15,8)	82 (33,2)	72 (29,1)	54 (21,9)
Rasgo 9	19 (7,7)	58 (23,5)	79 (32)	91 (36,8)
Rasgo 10*	34 (13,8)	72 (29,1)	103 (41,7)	38 (15,4)
Rasgo 11	2 (0,8)	48 (19,4)	81 (32,8)	116 (47)
Rasgo 12	28 (11,3)	81 (32,8)	59 (23,9)	79 (32)
Rasgo 13*	23 (9,3)	59 (24)	105 (42,7)	59 (24)
Rasgo 14	63 (25,6)	91 (37)	50 (20,3)	42 (17,1)
Rasgo 15	29 (11,7)	117 (47,4)	63 (25,5)	38 (15,4)
Rasgo 16*	16 (6,5)	69 (28,2)	112 (45,7)	48 (19,6)
Rasgo 17	23 (9,4)	74 (30,2)	81 (33,1)	67 (27,3)
Rasgo 18	46 (18,7)	60 (24,4)	55 (22,4)	85 (34,6)
Rasgo 19*	42 (17,1)	69 (28)	91 (37)	44 (17,9)
Rasgo 20	9 (3,6)	65 (26,3)	84 (34)	89 (36)

Nota. * El ítem ha sido invertido.

Tabla 2.

Saturaciones del análisis factorial exploratorio de los ítems del STAI.

Ítems	Factor 1	Factor 2
Estado 1*	0,589	0,553
Estado 2*	0,39	0,668
Estado 3	0,796	0,303
Estado 4	0,677	0,381
Estado 5*	0,484	0,583
Estado 6	0,856	
Estado 7	0,541	
Estado 8*		0,563
Estado 9	0,819	0,372
Estado 10*	0,469	0,734
Estado 11*		0,698
Estado 12	0,831	0,316
Estado 13	0,491	
Estado 14	0,734	
Estado 15*	0,639	0,64
Estado 16*	0,321	0,817
Estado 17	0,68	0,432
Estado 18	0,821	0,326
Estado 19*		0,869
Estado 20*	0,52	0,745
Rasgo 1*	0,824	
Rasgo 2	0,507	0,365
Rasgo 3	0,529	0,375
Rasgo 4	0,608	0,404
Rasgo 5	0,374	0,322
Rasgo 6*	0,596	
Rasgo 7*	0,395	0,323
Rasgo 8	0,598	0,504
Rasgo 9		0,855
Rasgo 10*	0,893	
Rasgo 11		0,506
Rasgo 12	0,569	0,333
Rasgo 13*	0,721	
Rasgo 14	0,301	
Rasgo 15	0,675	0,306
Rasgo 16*	0,889	
Rasgo 17		0,747
Rasgo 18	0,504	0,414
Rasgo 19*	0,602	0,308
Rasgo 20	0,516	0,558

Nota. Se eliminaron las saturaciones por debajo de 0,3; los análisis factoriales exploratorio se realizaron por separado para estado y rasgo. Se empleó máxima versosimilitud y rotación oblimin. * El ítem ha sido invertido

grupo de mujeres la media en ansiedad estado fue de 31,34 ($DT = 13,32$; $t(154) = 11,55$; $p < 0,001$) y en ansiedad rasgo de 34,12 ($DT = 11,37$; $t(154) = 11,77$; $p < 0,001$). Calculadas las medias por sexo se analizó si las mismas mostraban diferencias significativas entre hombres y mujeres. En el caso de la ansiedad estado la diferencia de medias resultó significativa siguiendo un criterio liberal ($t(185,98) = 2,14$; $p = 0,034$). En la ansiedad rasgo las diferencias por sexo no fueron significativas ($t(179,52) = 0,98$; $p = 0,335$).

En cuarto lugar se calculó la fiabilidad de las subescalas del STAI. El alfa policórico fue de 0,951 en la ansiedad estado (no habiendo ningún ítem cuya eliminación mejorase el alfa) y de 0,916 en la ansiedad rasgo (la eliminación del ítem 11 mejoraría el alfa en 0,003 puntos). Para complementar este valor se calculó el coeficiente de correlación ítem-total corregido. En el caso de la ansiedad estado todos los valores son superiores a 0,48. En la ansiedad rasgo todos los ítems están por encima de 0,45, salvo el ítem 7 ($r = 0,28$; $p < 0,001$) y el ítem 11 ($r = 0,35$; $p < 0,001$).

Y por último, en quinto lugar se calculó la correlación de Pearson entre la ansiedad estado y la ansiedad rasgo con tres medidas de ansiedad y depresión (ver tabla 3). Las correlaciones entre la ansiedad estado, rasgo y el BAI son moderadas. De la misma forma, las correlaciones entre las subescalas del STAI y las medidas de depresión (IDER y CBD) son moderadas en todos los casos, salvo en la relación entre la ansiedad estado y la puntuación en depresión estado del IDER, que es alta ($r = 0,81$; $p < 0,001$).

Tabla 3.

Correlaciones de Pearson entre el STAI, el BAI el CBD y el IDER.

Variables	STAI Estado	STAI Rasgo	IDER Estado	IDER Rasgo	CBD
STAI Rasgo	0,57**				
IDER Estado	0,81**	0,60**			
IDER Rasgo	0,50**	0,78**	0,62**		
CBD	0,53**	0,72**	0,58**	0,69**	
BAI	0,54**	0,54**	0,56**	0,41**	0,51**

Nota. STAI: State-Trait Anxiety Inventory; IDER: Inventario de Depresión Estado-Rasgo; CBD: Cuestionario Básico de Depresión; BAI: Inventario de Ansiedad de Beck. ** = $p < 0,001$.

Discusión

En el presente artículo se aborda las propiedades psicométricas del STAI en una muestra de pacientes con diagnóstico de ansiedad. En los resultados se puede observar, que en esta población las medias son significativas, tomando como referencia el valor medio de las puntuaciones de la adaptación española para una muestra de adultos (Buela-Casal et al., 2011). Al mismo tiempo se observa que las diferencias de las medias en ansiedad estado y rasgo entre hombres y mujeres son de menos de dos puntos, mientras que en otros estudios con población general las mujeres puntúan significativamente más que los hombres (Guillén-Riquelme y Buela-Casal, 2011). En conclusión, con altos niveles de ansiedad estas diferencias tienden a desaparecer, por lo tanto, si bien las mujeres tienen, por lo general, mayores niveles de ansiedad estado y rasgo (Spielberger y Reheier, 2009), cuando la ansiedad es elevada tiende a igualarse por sexo.

En lo que respecta a la fiabilidad, los valores alfa son elevados para ambas subescalas, aunque ligeramente inferiores a los observados en población general, tanto española (Guillén-Riquelme y Buela-Casal, 2011) como en muestras de diferentes países (Barnes et al., 2002). Además, al comparar los valores del alfa con los obtenidos por muestras de pacientes con ansiedad las puntuaciones son similares (Guillén-

Riquelme y Buela-Casal, 2014). No obstante, es necesario indicar que en el presente artículo se ha empleado el alfa categórico, mientras que en Guillén-Riquelme y Buela-Casal (2014) se estudian los valores del alfa de Cronbach, debido a que son los más empleados en la literatura científica. En cualquier caso, los alfa son superiores a 0,9 tanto en ansiedad estado como rasgo, por lo que la consistencia interna del STAI en pacientes con diagnóstico de ansiedad es elevada.

Otro de los temas controvertidos en la investigación del STAI es la factorización. Desde su creación los autores originales defienden la estructura teórica de ansiedad estado y ansiedad rasgo (Spielberger et al., 1970; Spielberger y Reheier, 2009). Otros autores afirman que dichas subescalas están formadas a su vez por dos factores (presencia y ausencia de ansiedad) separados (Hishinuma, Miyamoto, Nishimura y Nahulu, 2000). También, hay estudios en los que se concluye que existen agrupaciones de ítems que pueden ser indicio de contenidos diferentes a la ansiedad (Bieling et al., 1998; Caci, Baylé, Dossios, Robert y Boyer, 2003; Hill, Musso, Jones, Pella y Gouvier, 2012). En los resultados obtenidos en el presente estudio se puede observar una ligera influencia de la inversión de los ítems en las agrupaciones de las saturaciones factoriales. A pesar de lo cual, la mayoría de los ítems tienen elevadas cargas en ambos factores, por lo que cabe defender que cada una de las subescalas del STAI es unifactorial. Además, no se observan agrupaciones de ítems que permitan inferir estructuras factoriales en las que se evalúe otro contenido diferente a la ansiedad.

De forma independiente a la estructura factorial subyacente, se calculó el total de la ansiedad estado y la ansiedad rasgo, las cuales se correlacionaron moderada y significativamente entre sí, lo que niega la tesis clásica de la independencia de la ansiedad estado y la ansiedad rasgo entre sí. Así, en la adaptación española se aconseja no invertir el orden de aplicación de las escalas de respuesta, respecto a su presentación

en la plantilla, para evitar que las puntuaciones en ansiedad rasgo influyan en la ansiedad estado (Buela-Casal et al., 2011).

Por último, se analizó las correlaciones entre la ansiedad estado y rasgo con el BAI, el IDER y el CBD. En este caso se observa que todas las correlaciones son moderadas. Es decir, que el STAI correlaciona con medidas de ansiedad, pero también tiene una relación moderada con medidas de depresión. Estos datos son similares a los encontrados en otros estudios. Por ejemplo, Grös et al. (2002), obtienen correlaciones similares a las observadas en el presente artículo, pero en población no clínica, entre el STAI e instrumentos de depresión. Pese a esa relación, no está claro si se debe a un problema del STAI o a la elevada comorbilidad entre la ansiedad y la depresión (Beutel, Bleichner, von Heymann, Tritt y Hardt, 2011; Lamers et al., 2011), algo que debería ser de especial relevancia en las revisiones de las clasificaciones de trastornos mentales y del comportamiento (Obiols, 2012; Reed, Anaya y Evans, 2012).

En resumen, en este trabajo se han aportado evidencias de la consistencia interna y la validez de constructo de la adaptación española del STAI en pacientes con diagnóstico de ansiedad. Estos hallazgos resultan muy útiles para los clínicos a la hora de aplicar el STAI con garantías psicométricas en pacientes ansiosos, pero existen dos limitaciones: el hecho de que la prevalencia de trastornos en la muestra sea dispar, lo que dificulta su generalización, y el que las correlaciones entre las puntuaciones del STAI y la depresión sean elevadas, lo que cuestiona el constructo evaluado. Pese a estas dos limitaciones, en este estudio se realizan dos aportaciones concretas: se determina la puntuación media en pacientes con ansiedad, y aporta evidencias de la fiabilidad y validez del STAI en pacientes con diagnóstico de ansiedad, lo que supone una garantía para su uso por los profesionales de la salud mental.

Referencias

- American Psychological Association (2002). *Manual diagnóstico y estadístico de los trastornos mentales. DSM-IV-TR*. Barcelona: Masson.
- Bados, A., Gómez-Benito, J. y Balaguer, J. (2010). The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety? *Journal of Personality Assessment*, *92*, 560-567. doi: 10.1080/00223891.2010.513295
- Barnes, L. L. B., Harp, D. y Jung, W. S. (2002). Reliability Generalization of Scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, *62*, 603-618. doi: 10.1177/0013164402062004005
- Beck, A. T., Epstein, N., Brown, G. y Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*, 893-897. doi:10.1037/0022-006X.56.6.893
- Beutel, M.E., Bleichner, F., von Heymann, F., Tritt, K. y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology*, *11*, 443-457.
- Bieling, P. J., Antony, M. M. y Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behavior Research and Therapy*, *36*, 777-788. doi: 10.1016/S0005-7967(98)00023-0
- Buela-Casal, G. y Agudelo-Vélez, D. (2008). *IDER: Inventario de Depresión Estado-Rasgo*. Madrid: TEA Ediciones, S.A.
- Buela-Casal, G. y Guillén-Riquelme, A. (en prensa). Short form of the Spanish adaptation of the State-Trait Anxiety Inventory. *Assessment*.

- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos Cubero, N. (2011). *Cuestionario de Ansiedad Estado-Rasgo: Adaptación española* (8ª ed.). Madrid: TEA Ediciones.
- Caci, H., Baylé, F. J., Dossios, C., Robert, P. y Boyer, P. (2003). The Spielberger trait anxiety inventory measures more than anxiety. *European Psychiatry*, *18*, 394-400. doi:10.1016/j.eurpsy.2003.05.003
- Cho, S., Li, F. y Bandalos, D. (2009). Accuracy of the parallel analysis procedure with polychoric correlations. *Educational and Psychological Measurement*, *69*, 748-759. doi:10.1177/0013164409332229
- Elosua Oliden, P. y Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema*, *20*, 896-901.
- Gadermann, A. M., Guhn, M. y Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, *17*. Descargado de <http://www.pareonline.net/getvn.asp?v=17&n=3>.
- Grös, D. F., Antony, M. M., Simms, L. J. y McCabe, R. E. (2007). Psychometric Properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment*, *19*, 369-381. doi: 10.1037/1040-3590.19.4.369
- Guillén-Riquelme, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory. *Psicothema*, *23*, 510-515.
- Guillén-Riquelme, A. y Buela-Casal, A. (2013). Versión breve del STAI en adolescentes y universitarios españoles. *Terapia Psicológica*, *31*, 293-299.

- Guillen-Riquelme, A. y Buela-Casal, G. (2014). Metaanálisis de comparación de grupos y metaanálisis de generalización de la fiabilidad del cuestionario State-Trait Anxiety Inventory (STAI). *Revista Española de Salud Pública*. Avance on-line.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Hill, B. D., Musso, M., Jones, G. N., Pella, R. D. y Gouvier W. D. (2012). A Psychometric Evaluation of the STAI-Y, BDI-II, and PAI Using Single and Multifactorial Models in Young Adults Seeking Psychoeducational Evaluation. *Journal of Psychoeducational Assessment*, XX, 1-13. doi: 10.1177/0734282912462670
- Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T. y Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory Scores for Ethnically Diverse Adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83. doi:10.1037/1099-9809.6.1.73
- Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J., ... Penninx, B. W. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *Journal of Clinical Psychiatry*, 72, 341-348. doi:10.4088/JCP.10m06176blu
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.

- Obiols, J. E. (2012). DSM 5: Precedents, present and prospects. *International Journal of Clinical and Health Psychology*, 12, 281-290.
- Peñate, W. (2001). Presentación de un cuestionario básico para evaluar los síntomas genuinos de la depresión. *Análisis y Modificación de Conducta*, 27, 679-731.
- Perpiñá-Galvan, J., Richart-Martínez, M. y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189. doi:10.1016/j.arbres.2010.11.006
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J. y Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem*, 19, A04:1-A04:6.
- Reed, G. M., Anaya, C. y Spencer, C. E. (2012). ¿Qué es la Cie y por qué es importante en la psicología? *International Journal of Clinical and Health Psychology*, 12, 461-473.
- Rossi, V. y Pourtois, G. (2012). Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety, Stress, & Coping*, 25, 603-645. doi: 10.1080/10615806.2011.582948
- Sanz, J., Vallar, F., de la Guía, E. y Hernández, A. (2011). *Inventario de Ansiedad de Beck*. Madrid: Pearson Educación, S.A.
- Spielberger, C. D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.

Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being, 1*, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.

Spielberger, C. D. y Ritterband, L. (1996). *Preliminary Test Manual for the State-Trait Depression Scale*. Tampa: University of South Florida.

DISCUSIÓN

En esta tesis se pretende alcanzar tres objetivos mediante seis estudios. El primero de los tres objetivos de esta tesis es analizar la fiabilidad del STAI. Para ello, se han realizado dos estudios, en uno de ellos, mediante un meta-análisis se observó que el STAI es fiable en muestras de pacientes diagnosticados con alguno de los trastornos de ansiedad recogidos en el DSM-IV-TR. Sin embargo, en la revisión sistemática de este primer estudio, no se encontró ningún documento con muestra española de pacientes con ansiedad, hecho que resulta destacable puesto que, si bien la versión española del STAI no cuenta con un punto de corte clínico, es uno de los instrumentos más empleados por psicólogos. Así pues, el STAI tiene una adecuada fiabilidad en muestra general española y en muestras internacionales de pacientes con diagnóstico principal de ansiedad (Estudio 2).

Tras comprobar que la versión completa del STAI era fiable, se pasó a abordar el segundo objetivo: si algún conjunto de ítems permitía la evaluación de la ansiedad rasgo y la ansiedad estado (Estudio 3). Se determinó una selección de diez ítems que permitía la evaluación de la ansiedad, no obstante al comparar, mediante análisis factorial confirmatorio, esta versión con otras selecciones de ítems teóricas, se observó que la versión de diez ítems no era la que tenía el mejor ajuste, aunque todas ellas, tenían unos índices adecuados. Pese a ello, no se obtuvo un buen ajuste de ésta ni de otras formas breves en la muestra de adolescentes. En el caso de los universitarios, los resultados coinciden plenamente con los obtenidos en muestra general (Estudio 4).

El tercer objetivo fue analizar las propiedades psicométricas de la versión completa del STAI para muestras de pacientes depresivos (Estudio 5) y ansiosos (Estudio 6). En ambos casos la fiabilidad es elevada y la dimensionalidad es bifactorial (ansiedad estado y ansiedad rasgo) correspondiéndose con los ítems positivos y aquellos que han sido invertidos. Este dato es relevante ya que diversos autores defienden que

varios ítems del STAI evalúan otros factores como depresión o malestar general, y en el caso de pacientes con trastornos depresivos en la factorización no se observan agrupaciones de ítems cuyo contenido pueda indicar que se evalúa algo diferente a ansiedad estado y rasgo positiva y negativa. Además, en el caso de los pacientes de ansiedad la relación del STAI con el cuestionario de ansiedad de Beck es moderada-alta.

Más allá de este breve resumen, a continuación, se presenta una discusión compartida para la Tesis siguiendo los tres objetivos principales señalados al inicio:

El establecimiento de la Psicología como ciencia hace necesario el uso de instrumentos de medida fiables y válidos. La consistencia interna de un instrumento varía en función de la muestra y la población, por lo cual, es común en estudios aplicados que se informe de la fiabilidad en la muestra del mismo. En este estudio se han seleccionado diferentes estudios en los que se diese el valor de alfa de Cronbach para muestras con pacientes ansiosos. Los valores alfa medios son muy elevados para ambas escalas lo que indica que el STAI es adecuado para evaluar los niveles de ansiedad en pacientes con este tipo de trastornos. Estos datos encajan con los de la adaptación original (Spielberger, Gorsuch y Lushene, 1970), en su segunda versión modificada, donde el alfa mejoraba en unas décimas (Spielberger, Gorsuch, Lushene, Vagg y Jacobs, 1983; Spielberger y Reheier, 2009). Estos resultados son similares a los obtenidos en la adaptación española (Buela-Casal, Guillén-Riquelme y Seisdedos Cubero, 2011), que se sitúa como el séptimo instrumento más utilizado por psicólogos españoles (Muñiz y Fernández-Hermida, 2009), y a los de diferentes adaptaciones internacionales (Barnes, Harp y Jung, 2002; Rossi y Pourtois, 2012).

En los análisis realizados se observa que la forma X del STAI es más fiable que la forma Y, lo que resulta contrario a la mayoría de resultados en población normal (Spielberger et al., 1970; Spielberger et al., 1983; Spielberger y Reheier, 2009),

probablemente porque las muestras de depresivos suelen ser menos numerosas que las de población general y el tamaño muestral influye en el alfa de Cronbach. En el caso del porcentaje de mujeres no se encuentra influencia de esta variable por lo que la fiabilidad no varía entre hombres y mujeres, lo cual de nuevo encaja con los resultados originales del STAI (Spielberger y Reheier, 2009). En el caso de los diferentes países se observa que el alfa de Cronbach siempre es adecuado en este tipo de muestras, aunque cabe destacar que no se encuentra ningún estudio en el que se utilice el STAI en pacientes con ansiedad españoles.

Las diferencias de medias en el STAI entre pacientes con ansiedad y grupos control son significativas, siendo moderadas por país, forma del STAI aplicada (X o Y) y por tipo de trastorno de ansiedad diagnosticado. En la versión original, así como en la forma Y existe un punto de corte diagnóstico para facilitar la evaluación de niveles de ansiedad patológicos (Spielberger et al., 1970). En el caso de España, pese a que el STAI no disponga de punto de corte clínico (Buela-Casal et al., 2011), en su uso se observan diferencias en las puntuaciones de grupos de sujetos sanos y pacientes de ansiedad.

Tras el establecimiento de los valores generales de fiabilidad en el STAI (Barnes et al., 2002; Rossi y Pourtois, 2012), se pasó a analizar los ítems y la consistencia interna de la adaptación original en 1982 ya que los niveles de ansiedad en la población podían haber variado (Buela-Casal et al., 2011). En los resultados del Estudio 2 se obtuvo que los valores medios de la ansiedad estado han disminuido cinco puntos (diferencia estadísticamente significativa), respecto al valor de la primera adaptación española (Seisdedos, 1982). Sin embargo en el caso de la ansiedad rasgo dicha disminución no sobrepasa los dos puntos; indicio de que el STAI ha detectado las diferencias en la población producidas por el paso del tiempo.

En este segundo estudio se analizó la estructura factorial del STAI. Este tema resulta más controvertido, encontrando las siguientes propuestas: 1) estructura de dos factores (ansiedad estado y rasgo) independientes entre sí (Spielberger et al., 1970); 2) Estructura bifactorial de la ansiedad estado y monofactorial de la ansiedad rasgo (presencia y ausencia de ansiedad estado; Hishinuma, Miyamoto, Nishimura y Nahulu, 2000) 3) tetrafactorial (presencia y ausencia de ansiedad estado y rasgo) como el caso de la adaptación española (Seisdedos Cubero, 1982), u otras adaptaciones en castellano (Vera-Villarroel, Buena-Casal y Spielberger, 2007); 4) tetrafactorial con cuatro variables diferentes (ansiedad estado, rasgo, presencia y ausencia de ansiedad) de tal forma que por parejas definiesen cada uno de los ítems (Vigneau y Cormier, 2008). En dicho Estudio 2 de la presente tesis se ha obtenido un modelo tetrafactorial, donde destacan los dos factores teóricos del STAI (estado y rasgo) y cierta influencia de la inversión de los ítems, confirmando el modelo de Seisdedos Cubero (1982) y Vera-Villarroel et al. (2007)

En los resultados del Estudio 2 se continuó con el análisis del funcionamiento diferencial de los ítems, con el fin de conocer si las diferencias por sexo observadas en el STAI se podían deber a un posible sesgo. Se observó que solo el ítem 3 de la ansiedad rasgo (“Siento ganas de llorar”) presenta funcionamiento diferencial, siendo los hombres más propensos a la selección de la opción “casi nunca”. Por último, se procedió al análisis de fiabilidad, obteniendo valores superiores a 0,9, lo que es indicio de que la adaptación española del STAI tiene una adecuada consistencia interna (Buena-Casal et al., 2011), similar a las obtenidas en otras muestras internacionales (Barnes et al., 2002; Rossi y Pourtois, 2012).

Una vez confirmada la fiabilidad del STAI en población general española, el siguiente objetivo fue determinar una versión breve del mismo. Mediante un análisis

discriminante, se establecieron diez ítems (cinco para ansiedad rasgo y otros cinco para ansiedad estado) que clasificaban bien a la mayoría de los sujetos. Este dato es indicio de que se puede establecer una versión breve del STAI en población general española, al igual que en la original (Marteau y Bekker, 1992). A ello cabe añadir que en la mayoría de las versiones breves del STAI se establece seis ítems por escala, pero el criterio para seleccionar este número de ítems es replicar el mismo procedimiento que Marteau y Bekker (1992), por ello, resulta importante determinar el número de ítems en función de criterios estadísticos, ya que, en los estudios publicados se encuentran alternativas con diferente número de ítems que resultan fiables y válidas (Kaipper, Chachamovich, Hidalgo, Torres y Caumo, 2010).

Se comparó la versión breve establecida, con otras tres, provenientes siempre a partir de la forma X del STAI. En los resultados se observa que la versión que mejor ajusta es la de seis ítems propuesta por van Knippenberg, Duivenvoorden, Bonke y Passchier (1990), que muestra un ajuste superior a la determinada mediante análisis discriminante, así como a la selección de ítems teórica con el mismo número de ítems por escala (Fioravanti-Bastos, Cheniaux y Landeira-Fernandez, 2011) e incluso con el doble de ítems (Kaipper et al., 2010); por lo que es posible establecer un subconjunto de ítems que permita evaluar ansiedad estado y rasgo. De hecho existe una versión breve previamente validada en población española por Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez (2011), no incluida por dos motivos: 1) proviene de una versión breve realizada a partir de la forma Y del STAI y 2) fue realizada, tanto en su versión original (Chlan, Savik y Weinert, 2003; Chen 2006) como en la adaptación española, para pacientes hospitalarios con intubación respiratoria. En el caso de la adaptación española muestra elevados índices de fiabilidad y validez (Perpiñá-Galvañ et al., 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez y Martínez-Durá, 2011).

Una vez establecida una selección de ítems del STAI en población general, se procedió a su análisis con una muestra de adolescentes y otra de universitarios. En los resultados se observa que en la muestra de universitarios varios de los modelos tienen índices adecuados, siendo el de van Knippenberg et al. (1990) el que mejor ajusta. El modelo establecido por medio de análisis discriminante en el Estudio 3 también obtiene índices adecuados.

En la muestra de adolescentes no se observan ajustes adecuados en ninguno de los modelos. Este hecho resulta extraño puesto que en la versión completa del STAI se observan elevados índices de fiabilidad y validez en muestras de adolescentes (Spielberger y Reheier, 2009), así como en la adaptación española (Buela-Casal et al., 2011) y en otras en español (Vera-Villaruel et al., 2007). Así pues, el STAI completo permite evaluar ansiedad en adolescentes pero no así las versiones breves del mismo. Una explicación posible se halla en la teoría general del afecto, desde la que se defiende que las emociones comienzan siendo muy generales para ir diferenciándose paralelamente al desarrollo personal (Yik, Russell y Feldman-Barrett, 1999). Así, para poder evaluar a adolescentes con las mismas garantías que en muestras de adultos sería necesario el uso de todos los ítems del STAI, pues las versiones breves del mismo presentan carencias en esta muestra.

Como se señaló al principio de la discusión el tercer objetivo de la presente tesis era estudiar la consistencia interna y validez del STAI en dos muestras clínicas: depresivos y ansiosos. En el Estudio 5 se analizó el alfa de Cronbach y la factorización de una muestra de pacientes con alguno de los diagnósticos de depresión recogidos en el DSM-IV-TR (American Psychological Association, 2002). En la consistencia interna los valores fueron elevados. Así pues el STAI es fiable en esta muestra de depresivos, donde las medias son significativamente mayores que el valor de referencia de la

adaptación original. Posiblemente se debe a que la ansiedad y la depresión tienen una elevada comorbilidad entre sí, compartiendo aspectos del afecto negativo (Beutel, Bleichner, von Heymann, Tritt y Hardt, 2011; Lamers et al., 2011). Otra posible explicación sería que algunos de los ítems del STAI estén evaluando depresión y malestar general (v.gr. Andrade, Gorenstein, Vieira Filho, Tung y Artes, 2001).

En el caso de la factorización se encuentran de nuevo, cuatro factores: dos de ansiedad estado y dos de ansiedad rasgo, en los que hay cierta influencia de la inversión de algunos de los ítems, tal y como se presenta en otros estudios (Hishinuma et al., 2000; Vautier, 2004). De hecho, las saturaciones factoriales son moderadas o altas en la mayoría de los ítems en ambos factores, por ello cabría defender la bifactorialidad del STAI, frente a otros estudios donde se señala que este inventario tiene ítems en los que se evalúa depresión (Andrade et al., 2001; Endler, Cox, Parker y Bagby, 1992; Grös, Antony, Simms y McCabe, 2007) o malestar general (Balsamo et al., 2013; Hill, Musso, Jones, Pella y Gouvier, 2012); entendiéndolo como un componente característico, tanto de los trastornos depresivos como de los ansiosos (Watson, Clark y Stasik, 2011).

Entre los defensores de que el STAI evalúa afecto negativo o depresión, hay diversos autores que emplean el análisis factorial exploratorio para confirmar que hay ítems del cuestionario en los que se evalúa depresión o malestar general. Por ejemplo, Bieling, Antony y Swinson (1998), obtuvieron una matriz de saturaciones factoriales a partir de la cual afirmaban que el STAI rasgo estaría compuesto por dos subescalas: depresión y ansiedad, lo que ha sido corroborado en otros trabajos, que muestran selecciones distintas de los ítems incluidos en el factor de depresión (Bados, Gómez-Benito y Balaguer, 2010; Caci, Baylé, Dossios, Robert y Boyer, 2003). En el Estudio 5, se ha realizado el análisis factorial con una muestra de depresivos, sin observar agrupaciones de los ítems criticados por evaluar depresión. En una muestra de

depresivos es de suponer que los ítems de depresión saturasen con mayor fuerza en un factor independiente. Sin embargo, como se acaba de afirmar, en los resultados no se observan este tipo de agrupaciones. Por ello, se defiende la estructura bifactorial clásica del STAI (estado y rasgo) incluso en una muestra de pacientes depresivos, lo que supone una nueva aportación para su uso en muestras españolas.

En estudio sexto sirve para completar lo sustancial de esta Tesis, abordando el último de sus tres objetivos: analizar la fiabilidad y validez del STAI en una muestra de pacientes españoles con diagnóstico de ansiedad. En primer lugar, se observa que las medias son significativamente superiores al valor de referencia de la media para población normal, obtenido de su adaptación española (Buela-Casal et al., 2011), lo que indica que en este cuestionario se obtienen mayores puntuaciones en personas con diagnóstico de ansiedad.

Al analizar la fiabilidad del STAI se observa que en ambas subescalas la consistencia interna es ligeramente inferior que en la versión original (Barnes et al., 2002; Rossi y Pourtois, 2012; Spielberger et al., 1970), así como a su adaptación española en población general (Buela-Casal et al., 2011) y a los datos obtenidos en el Estudio 2 con una muestra de población general. Más importante es el hecho de que estos valores sean similares o superiores a los obtenidos en el meta-análisis del Estudio 2, en el que se estimaba la fiabilidad media para diversas muestras de pacientes con ansiedad. Para aquellos casos en los que las muestras eran similares, los alfa de Cronbach obtenidos son semejantes, por lo que la adaptación española del STAI tiene la misma consistencia interna en este tipo de pacientes.

El cálculo de la fiabilidad fue seguido por un análisis factorial para cada una de las subescalas del mismo, obteniendo matrices de saturaciones similares a las de los Estudios 2 y 5; ello indica que la factorización de la adaptación española del STAI es

bifactorial, no observando saturaciones de ítems de las que se pueda inferir que hay factores subyacentes de depresión o malestar general, como afirman otros autores (Bados et al., 2010; Bieling et al., 1998; Caci et al., 2003).

Para concluir el sexto y último estudio se procedió a correlacionar el STAI con diversos instrumentos: el Inventario de Ansiedad de Beck (BAI; Beck, Epstein, Brown y Steer, 1988), el Inventario de Depresión Estado-Rasgo (IDER; Buena-Casal y Agudelo-Vélez, 2008; Spielberger y Ritterband, 1996) y el Cuestionario Básico de Depresión (CBD; Peñate, 2001). Las correlaciones entre el STAI y el BAI son moderadas lo que es indicio de que el primero se relaciona con medidas de ansiedad. A la vez, las correlaciones con los cuestionarios de depresión también son moderadas, lo que encaja con la literatura previa (Grös et al., 2007; Spielberger y Reheier, 2009). Así pues, el STAI presenta cierta confusión en la evaluación de la ansiedad y la depresión, aunque en el análisis factorial no se observó ningún conjunto de ítems en los que se evalúe, de forma específica, depresión. En el Estudio 6, se concluye que la adaptación española del STAI: 1) es fiable; 2) sus factorizaciones en diversas muestras son estables; 3) las puntuaciones medias son superiores a las de población general y 4) las correlaciones con medidas de ansiedad y depresión son moderadas.

Líneas futuras de investigación

Pese a los estudios realizados en la presente tesis doctoral y a las conclusiones obtenidos en los mismos, el trabajo cuenta con algunos aspectos mejorables en los que cabe trabajar. Así pues, en primer lugar, cabría analizar otros moderadores que influyesen en las fluctuaciones de la fiabilidad del STAI, e incluir medidas de correlación test-retest.

En el caso de las muestras clínicas, sería muy interesante ampliar el número de pacientes con el fin de cubrir un número representativo, para cada uno de los trastornos

incluidos y poder analizar, de forma independiente, para cada uno de ellos, la fiabilidad y validez. Así pues, con mayores muestras de ansiedad y depresión se podría determinar si existen diferencias entre ambas muestras, en los niveles de ansiedad obtenidos.

Por otra parte, en lo que respecta a las versiones breves, cabría estudiar si existe alguna de ellas, que permita evaluar la ansiedad estado y rasgo en muestras de adolescentes. Para ello, se podría probar con alguna de las versiones establecidas a partir de la forma Y del STAI o partiendo del cuestionario completo y realizar análisis discriminante o de correlación ítem-total. Independientemente, en el caso de las versiones breves sería necesario analizar su funcionamiento en pacientes hospitalarios, respecto a la versión alternativa, propuesta para este tipo de pacientes.

Conclusiones

Al realizar una lectura sistemática de los seis estudios que componen la presente Tesis Doctoral, se extraen una serie de conclusiones sustanciales:

- El STAI tiene adecuada consistencia interna en muestras de pacientes con diagnóstico de ansiedad de diferentes países.
- Las puntuaciones en el STAI de los pacientes de ansiedad son superiores a las de población general.
- En la adaptación española del STAI las puntuaciones medias en ansiedad estado han variado significativamente, tomando como referencia las de la primera adaptación española.
- En la adaptación española del STAI las puntuaciones medias en ansiedad rasgo no han variado significativamente, tomando como referencia las de la primera adaptación española.
- El ítem 3 de la subescala de ansiedad rasgo presenta funcionamiento diferencial entre mujeres y hombres.
- La consistencia interna de la adaptación española del STAI es superior a 0,9 en ambas subescalas.
- Hay diferentes versiones breves del STAI que ajustan adecuadamente en población general.
- El modelo de van Knipenberg et al. (1990) es el que mejor ajuste presenta en población general.
- En muestras de universitarios el modelo de van Knipenberg et al. (1990) es el que mejor ajuste presenta.

- En adolescentes ninguna de las versiones breves resulta adecuada.
- En la factorización del STAI en muestras de depresivos no se observa ningún patrón que permita establecer un factor de “depresión” o “malestar general”
- En depresivos la consistencia interna es superior a 0,9.
- En pacientes con ansiedad la consistencia interna es superior a 0,9.
- En pacientes con ansiedad las medias en estado y rasgo son superiores al valor medio de la adaptación española para población general.

Conclusions

By conducting a systematic reading of the six studies that make up this Thesis, a number of substantial conclusions were extracted:

- The STAI has adequate internal consistency in samples of patients diagnosed with anxiety in different countries.
- Scores on the STAI anxiety patients are higher than the general population.
- In the Spanish adaptation of the STAI state anxiety the mean scores changed significantly, with reference to those of the first Spanish adaptation.
- In the Spanish adaptation of the STAI mean scores on trait anxiety have not changed significantly, with reference to those of the first Spanish adaptation.
- The item 3 of the trait anxiety subscale presents differential functioning between women and men.
- The internal consistency of the Spanish adaptation of the STAI is greater than 0.9 in both subscales.
- There are different versions of the STAI short to fit properly in the general population.
- The model of van Knipenberg et al. (1990) has the best fit in general population.
- In samples of university model van Knipenberg et al. (1990) has the best fit.
- In teenagers none short version has a good fit.
- In the factorization of the STAI in depressive samples, any pattern factor "depression" or " malaise " is observed
- In patients with depression the internal consistency is higher than 0.9.

- In patients with anxiety the internal consistency is higher than 0.9.
- In patients with anxiety the state and trait averages are higher than the average value of the Spanish adaptation in general population.

Referencias

- American Psychological Association (2002). *Manual diagnóstico y estadístico de los trastornos mentales. DSM-IV-TR*. Barcelona: Masson.
- Andrade, L., Gorenstein, C., Vieira Filho, A. H., Tung, T. C. y Artes, R. (2001). Psychometric properties of the Portuguese version of the State-Trait Anxiety Inventory applied to college students: factor analysis and relation to the Beck Depression Inventory. *Brazilian Journal of Medical and Biological Research*, 34, 367-374. doi:10.1590/S0100-879X2001000300011
- Bados, A., Gómez-Benito, J. y Balaguer, J. (2010). The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety? *Journal of Personality Assessment*, 92, 560-567. doi: 10.1080/00223891.2010.513295
- Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L. y Saggino, A. (2013). The State-Trait Anxiety Inventory: Shadows and Lights on its Construct Validity. *Journal of Psychopathology and Behavioral Assessment*, descargado el 13 de agosto de 2013 de <http://link.springer.com/article/10.1007/s10862-013-9354-5>. doi:10.1007/s10862-013-9354-5
- Barnes, L. L. B., Harp, D. y Jung, W. S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational Psychological Measurement*, 62, 603-818. doi: 10.1177/0013164402062004005
- Beck, A.T., Epstein, N., Brown, G. y Steer, R.A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56, 893-897. doi:10.1037/0022-006X.56.6.893

- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K. y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology, 11*, 443-457.
- Bieling, P. J., Antony, M. M. y Swinson, R. P. (1998). The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behavior Research and Therapy, 36*, 777-788. doi: 10.1016/S0005-7967(98)00023-0
- Buela-Casal, G. y Agudelo-Vélez, D. (2008). *IDER: Inventario de Depresión Estado-Rasgo*. Madrid: TEA Ediciones, S.A.
- Buela-Casal, G., Guillén-Riquelme, A. y Seisdedos Cubero, N. (2011). *Cuestionario de Ansiedad Estado-Rasgo: Adaptación española* (8ª ed.). Madrid: TEA Ediciones.
- Caci, H., Baylé, F. J., Dossios, C., Robert, P. y Boyer, P. (2003). The Spielberger trait anxiety inventory measures more than anxiety. *European Psychiatry, 18*, 394-400. doi:10.1016/j.eurpsy.2003.05.003
- Chen, Y. J. (2006). *Psychophysiological determinants of repeated ventilator weaning failure* (Tesis doctoral). Descargada de <http://www.nursing.arizona.edu>
- Chlan, L., Savik, K. y Weinert, C. (2003). Development of a Shortened State Anxiety Scale From the Spielberger State-Trait Anxiety Inventory (STAI) for Patients Receiving Mechanical Ventilatory Support. *Journal of Nursing Measurement, 11*, 283-293. doi:10.1891/jnum.11.3.283.61269
- Endler, N. S., Cox, B. J., Parker, J. D. A. y Bagby, R. M. (1992). Self-Reports of Depression and State-Trait Anxiety: Evidence for Differential Assessment. *Journal of Personality and Social Psychology, 63*, 832-838. doi:10.1037/0022-3514.63.5.832

- Fioravanti-Bastos, A. C. M., Cheniaux, E. y Landeira-Fernandez, J. (2011). Development and Validation of a Short-Form Version of the Brazilian State-Trait Anxiety Inventory. *Psicologia: Reflexão e Crítica*, 24, 485-494. doi: 10.1590/S0102-79722011000300009
- Grös, D. F., Antony, M. M., Simms, L. J. y McCabe, R. E. (2007). Psychometric Properties of the State–Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State–Trait Anxiety Inventory (STAI). *Psychological Assessment*, 19, 369–381. doi: 10.1037/1040-3590.19.4.369
- Hill, B. D., Musso, M., Jones, G. N., Pella, R. D. y Gouvier Wm. D. (2012). A Psychometric Evaluation of the STAI-Y, BDI-II, and PAI Using Single and Multifactorial Models in Young Adults Seeking Psychoeducational Evaluation. *Journal of Psychoeducational Assessment*, XX, 1-13. doi: 10.1177/0734282912462670
- Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T. y Nahulu, L. B. (2000). Differences in State-Trait Anxiety Inventory Scores for Ethnically Diverse Adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83. doi:10.1037/1099-9809.6.1.73
- Kaipper, M. B., Chachamovich, E., Hidalgo, M. P. L., Torres, I. L. S. y Caumo, W. (2010). Evaluation of the structures of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research*, 68, 223-233. doi:10.1016/j.jpsychores.2009.09.013
- Lamers, F., van Oppen, P., Comijs, H. C., Smit, J. H., Spinhoven, P., van Balkom, A. J., ... Penninx, B. W. (2011). Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and

- Anxiety (NESDA). *Journal of Clinical Psychiatry*, 72, 341-348. doi:10.4088/JCP.10m06176blu
- Marteau, T. M. y Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31, 301-306. doi:10.1111/j.2044-8260.1992.tb00997.x
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.
- Peñate, W. (2001). Presentación de un cuestionario básico para la evaluación de los síntomas genuinos de la depresión. *Análisis y Modificación de Conducta*, 27, 679-731.
- Perpiñá-Galvañ, J., Richart-Martínez, M. y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189. doi:10.1016/j.arbres.2010.11.006
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J. y Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem*, 19, A04:1-A04:6.
- Rossi, V. y Pourtois, G. (2012). Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety, Stress, & Coping*, 25, 603-645. doi:10.1080/10615806.2011.582948
- Seisdedos Cubero, N. (1982). *Manual del Cuestionario de Ansiedad Estado/Rasgo (STAI)*. Madrid: TEA Ediciones S.A.

- Spielberger, C. D., Gorsuch, R. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R. y Jacobs, G. A. (1983). *State-Trait Anxiety Inventory (Form Y)*. Palo Alto: Consulting Psychologists Press.
- Spielberger, C. D. y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being, 1*, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- Spielberger, C. D. y Ritterband, L. (1996). *Preliminary Test Manual for the State-Trait Depression Scale*. Tampa: University of South Florida.
- van Knippenberg, F. C. E., Duivenvoorden, H. J., Bonke, B. y Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology, 43*, 995-1000. doi:10.1016/0895-4356(90)90083-2
- Vautier, S. (2004). A Longitudinal SEM Approach to STAI Data: Two Comprehensive Multitrait-Multistate Models. *Journal of Personality Assessment, 83*, 167-179. doi:10.1207/s15327752jpa8302_11
- Vera-Villaruel, P., Buela-Casal, G. y Spielberger, C. D. (2007). Preliminary Analysis and Normative Data of the State-Trait Anxiety Inventory (STAI) in Adolescent and Adults of Santiago, Chile. *Terapia Psicológica, 25*, 155-162. doi:10.4067/S0718-48082007000200006
- Vigneau, F. y Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment, 90*, 280-285. doi:10.1080/00223890701885027

Watson, D., Clark, L. A. y Stasik, S. M. (2011). Emotions and the emotional disorders:

A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, *11*, 429-442.

Yik, M., Russell, J. y Feldman-Barrett, L. (1999). Structure of self-reported current

affect: integration and beyond. *Journal of Personality and Social Psychology*, *77*, 600-619. doi:10.1037/0022-3514.77.3.600

ANEXOS

REVISIÓN SISTEMÁTICA Y METAANÁLISIS

METAANÁLISIS DE COMPARACIÓN DE GRUPOS
Y METAANÁLISIS DE GENERALIZACIÓN DE LA FIABILIDAD
DEL CUESTIONARIO STATE-TRAIT ANXIETY INVENTORY (STAI)

Alejandro Guillén-Riquelme y Gualberto Buela-Casal.

Departamento Personalidad, evaluación y tratamiento psicológico. Universidad de Granada.

RESUMEN

Fundamentos: Desde su creación, el STAI se ha citado en más de 14.000 documentos, contando con más de 60 adaptaciones en diversos países. En algunas de ellas este cuestionario no cuenta con puntuaciones clínicas. El objetivo de este trabajo es determinar si el cuestionario *State-Trait Anxiety Inventory* (STAI) tiene puntuaciones superiores en personas diagnosticadas de ansiedad respecto a la población general. Además, se pretende analizar si la consistencia interna es adecuada en personas con ansiedad.

Método: Se realizó una búsqueda bibliográfica en Tripdatabase, Cochrane, Web of Knowledge, Scopus, PsycINFO y Scholar Google de documentos publicados entre 2008 y 2012. Se seleccionaron 131 artículos para la comparación entre pacientes diagnosticados de ansiedad respecto a la población general y 25 para la generalización de la fiabilidad. En los análisis se utilizó la *d* de Cohen para la comparación de medias (método de efectos aleatorios) y para la generalización de la fiabilidad el alfa de Cronbach (método de efectos fijos).

Resultados: En la comparación entre grupos, en la ansiedad estado ($d = 1,39$; IC95%: 1,22-1,56) y en la ansiedad rasgo ($d = 1,74$; IC95%: 1,56-1,91) las diferencias fueron estadísticamente significativas. Además de ello, la fiabilidad para pacientes con algún trastorno de ansiedad abarcó de 0,87 a 0,93.

Conclusiones: El STAI es un cuestionario sensible para medir el nivel de ansiedad y fiable en personas diagnosticadas de crisis de angustia, fobia específica, fobia social, fobia social generalizada, trastorno de ansiedad generalizada, trastorno de estrés post-traumático, trastorno obsesivo compulsivo o trastorno por estrés agudo.

Palabras clave: Ansiedad. Metaanálisis. Fiabilidad.

Correspondencia

Alejandro Guillén-Riquelme
Centro de Investigación Mente, Cerebro y Comportamiento-CIMCYC
Universidad de Granada
Campus de Cartuja s/n
18011 Granada.
agr@ugr.es

ABSTRACT

Meta-analysis of Group Comparison and
Meta-analysis of Reliability Generalization
of the *State-Trait Anxiety Inventory*
Questionnaire (STAI)

Background: Since its creation the STAI has been cited in more than 14,000 documents, with more than 60 adaptations in different countries. In some adaptations this instrument has no clinical scores. The aim of this work is to determine if the State-Trait Anxiety Inventory (STAI) has higher scores in patients diagnosed with anxiety than in general population. In addition, we want to examine if the internal consistency is adequate in anxious patient samples.

Methods: We performed a literature search in Tripdatabase, Cochrane, Web of Knowledge, Scopus, PsycINFO and Scholar Google, for documents published between 2008 and 2012. We selected 131 scientific articles to compare between patients diagnosed with anxiety and general population, and 25 for the generalization of reliability. For the analysis we used Cohen's *d* for means comparisons (random-effects method) and Cronbach's alpha for the reliability generalization (fixed-effects method).

Results: In the groups comparison the differences in state anxiety ($d = 1.39$; CI95%: 1.22-1.56) and in the trait anxiety ($d = 1.74$; CI95%: 1.56-1.91) were significant. The reliability for patients of some anxiety disorder was between 0.87 and 0.93.

Conclusions: So it seems that the STAI is sensitive to the level of anxiety of the individual and reliable for patients with diagnosis of panic attack, specific phobia, social phobia, generalized social phobia, generalized anxiety disorder, post-traumatic stress disorder, obsessive compulsive disorder or acute Stress disorder.

Keywords: Anxiety. Meta-analysis. Reliability.

INTRODUCCIÓN

Los trastornos de ansiedad son uno de los problemas psicológicos con mayor prevalencia. En una muestra de más de 9.000 estadounidenses se observaron tasas de prevalencia que alcanzaban el 15,6% en el caso de la fobia específica¹. Estos resultados coinciden con medidas epidemiológicas previas². Además, pese a que existen diferencias entre los continentes, la tasa de diagnósticos de ansiedad es muy elevada³⁻⁴.

Entre los distintos instrumentos disponibles para evaluar la ansiedad general se encuentra el cuestionario *State-Trait Anxiety Inventory* (STAI)⁵ que evalúa dos facetas de la ansiedad. En primer lugar, la ansiedad rasgo, entendida como los factores personales que influyen en la percepción del nivel de ansiedad de los estímulos que se presentan. Es decir, una persona con alta ansiedad rasgo percibirá un estímulo como más amenazante y ansiógeno que una persona con baja ansiedad rasgo. Por el contrario, la ansiedad estado se corresponde con los estímulos desencadenantes de ansiedad que hay en el entorno cercano del sujeto. La ansiedad rasgo y estado son teóricamente independientes⁵. Tras su creación se modificaron algunas preguntas y la escala de respuesta del STAI original, llamando a la nueva versión forma Y y nombrando desde entonces a la versión original como forma X⁶. En lo que respecta a su consistencia interna, el alfa de Cronbach medio para 52 estudios incluidos en un metaanálisis fue de 0,91 en ansiedad estado y para un metaanálisis con 51 artículos fue de 0,89 para la ansiedad rasgo⁷.

El STAI es uno de los cuestionarios más empleados, contando con más de 60 adaptaciones culturales y lingüísticas y con más de 14.000 citas en artículos de sus versiones para adultos⁷. Además, el STAI ha sido utilizado para medir el nivel de ansiedad en pacientes con diversos trastornos físicos y psicológicos⁸. Por ejemplo, el STAI se utilizó en un metaanálisis para evaluar el mejor tra-

tamiento para el trastorno de ansiedad generalizada, empleando la subescala de rasgo como principal medida de eficacia del tratamiento⁹. En dicho metaanálisis se incluyeron seis estudios en los que se empleó el STAI con esta muestra. Otro de los ejemplos, en los que se hace uso del STAI para evaluar los niveles de ansiedad en personas que sufren alguno de los trastornos de ansiedad es el artículo de Kvaal y cols¹⁰. En este estudio se aplicó el STAI a personas ancianas con diversos trastornos de ansiedad, hallando que las que tenían trastornos de ansiedad puntuaron significativamente más que los que no tenían ningún trastorno psicológico.

Por último, en un metaanálisis en el que se incluyeron casi 200 estudios se concluyó que el STAI estado es una medida sensible del nivel de ansiedad en grupos clínicos, así como en personas con diversos trastornos físicos¹¹. Además de ello, este instrumento es lo suficientemente preciso como para detectar diferencias en las puntuaciones de sujetos al ser sometidos a estímulos estresantes o procedimientos para aumentar sus niveles de ansiedad¹¹.

Los objetivos de este trabajo son en primer lugar analizar si el STAI detecta las diferencias entre personas con diferentes trastornos de ansiedad y sujetos sanos. En segundo lugar, determinar si la consistencia interna del STAI en muestras de personas con ansiedad es adecuada y si está influida por la forma del STAI (X o Y), por el país o por el tipo de trastorno de ansiedad diagnosticado.

MATERIAL Y MÉTODOS

Tipo de estudio. Con el fin de llevar a cabo los objetivos propuestos en este estudio, clasificado como un estudio teórico de revisión con metaanálisis¹³, se siguieron las recomendaciones PRISMA, ya que este documento presenta una guía para este tipo de artículos¹⁴ y para su redacción se adoptaron las propuestas presentadas por Hartely¹⁵.

Búsqueda de los estudios. En primer lugar se localizaron revisiones sistemáticas o metaanálisis mediante las opciones de búsqueda avanzada de las bases de datos en línea Tripdatabase y Cochrane. Tras no identificar ninguna revisión previa que emplease únicamente el STAI, se pasó a la segunda fase de la búsqueda. Para ello se buscó en diversas fuentes generales y específicas, con el fin de hallar todos los documentos relevantes¹⁶⁻¹⁷. Así pues, se emplearon la Web of Knowledge, Scopus y PyscINFO (a través de la plataforma OVID). Se acotó la búsqueda a los artículos publicados entre los años 2008 y 2012, ya que al ser un cuestionario muy utilizado, la inclusión de cinco años permitía encontrar un número representativo de artículos. Los términos de búsqueda se incluyeron únicamente en inglés.

La ecuación de búsqueda se formó, por una parte, con el nombre completo del STAI así como con el acrónimo y, por otra, con cada uno de los trastornos de ansiedad, indicados en el DSM-IV-TR¹². Se utilizó el DSM-IV-TR en lugar de la nueva versión DSM-5 debido a que en el momento de las búsquedas todavía no estaba publicada y además los trabajos incluidos empleaban la cuarta edición o anteriores. El término STAI se buscó en todo el artículo, mientras que el trastorno de ansiedad únicamente en título, resumen y palabras clave. Se realizó una búsqueda independiente para cada uno de los trastornos. Se incluyeron truncadores siempre que era posible (por ejemplo *phobi** o *obsessi**) e igualmente se utilizó el nombre completo del trastorno y el acrónimo del mismo en el caso de que éste fuese de uso común (por ejemplo PTSD o GAD). La última fase consistió en la búsqueda en Scholar Google con el fin de localizar la literatura gris y evitar un sesgo de publicación. Para realizar esta búsqueda se identificó el artículo original del STAI y, a través de las citas del mismo, se filtraron los años de interés y se verificaron los artículos resultantes. Este proceso se realizó desde el 21 de julio al 30 de septiembre de 2012. Tras descartar los

artículos que se podían excluir por título, en total se encontraron 674 referencias.

Tras la búsqueda, se trató de localizar el texto completo de los artículos. Para ello, se empleó su versión electrónica (incluyendo manuscritos publicados) mediante préstamo interbibliotecario. Para los artículos a los que no se pudo tener acceso se escribió un correo electrónico al autor, donde se informaba del objetivo del estudio y se solicitaba una copia del artículo. Los trabajos no obtenidos fueron descartados (3,12% respecto al total de artículos).

Criterios de selección de los estudios.

Una vez seleccionados los artículos, para determinar su inclusión definitiva se emplearon diversos criterios explícitos¹⁸: a) debían estar escritos en inglés, español o portugués, incluyendo muestras de cualquier país y publicados en cualquier región geográfica; b) únicamente se incluyeron artículos cuasi-experimentales y experimentales, excluyendo revisiones teóricas y estudios de caso único. Las revisiones sistemáticas y los metaanálisis también se incluyeron inicialmente, no encontrando ningún estudio de este tipo. En las citas de estos trabajos se buscaron artículos que reuniesen los criterios para ser incluidos en este estudio; c) en lo que respecta a la muestra sólo se seleccionaron aquellos estudios realizados con personas mayores de edad cuyo diagnóstico principal fuese uno de los trastornos de ansiedad establecidos en el DSM-IV-TR¹² (crisis de pánico, agorafobia, trastorno de angustia sin agorafobia, trastorno de angustia con agorafobia, agorafobia sin historia de trastorno de angustia, fobia específica, fobia social, trastorno obsesivo-compulsivo, trastorno por estrés postraumático, trastorno por estrés agudo y trastorno de ansiedad generalizada; excluyendo los trastornos de ansiedad debido a sustancia y a causas médica e incluyendo la fobia social generalizada). Además, dicho diagnóstico no podía estar justificado mediante las puntuaciones de STAI como único criterio y d)

se descartaron los artículos donde los problemas ansiosos de las personas de la muestra fuesen secundarios a causas médicas.

Respecto a la versión aplicada del STAI, en los artículos seleccionados se podía emplear la versión completa o una de las dos subescalas (estado o rasgo) indistintamente. De la misma forma, la versión del STAI aplicada podía ser tanto la original como una adaptación, siempre que hubiese sido publicada y respetase el número de ítems y la escala de respuesta se mantuviese en cuatro alternativas. En diversas adaptaciones del STAI se recodifica la escala original (de 1 a 4) a una nueva escala de 0 a 3. Tal y como se indica en la literatura⁶, para realizar comparaciones entre diferentes adaptaciones del STAI o con la versión original se debe sumar 20 a la puntuación obtenida, en cada una de las subescalas. Así pues, en los casos en los que se empleaban adaptaciones con este tipo de escala de respuesta se realizó esta transformación.

Por otra parte, también se incluyeron algunos criterios específicos para cada metaanálisis. Para la generalización de la fiabilidad se seleccionaron artículos en los que se incluyese un grupo clínico sin importar o no la existencia de un grupo control. Mientras en el caso de la comparación entre grupos únicamente se seleccionaron artículos con la existencia de un grupo control. Si los grupos se establecían mediante puntuaciones de un cuestionario de algún trastorno de ansiedad estos artículos eran descartados si el grupo control estaba formado únicamente por personas con puntuaciones bajas en dicha medida. En los artículos que informaban de un rango del valor de alfa (para varios trastornos de ansiedad sin especificar un valor para cada uno de ellos) se tomó el menor de los valores. En el caso de contar con medias pre y post-tratamiento siempre se seleccionó la media pre-tratamiento de ambos grupos. Además de todo ello, se excluyó un artículo en el que se preguntaba a familiares y amigos de suicidas la puntuación que el fallecido habría puesto en el STAI, y otra investiga-

ción cuyas puntuaciones medias eran superiores a 80 (puntuación teórica máxima de la escala). En la figura 1 se muestra un resumen del proceso de selección de artículos.

En los estudios en los que se emplearon varias muestras clínicas la inclusión fue diferente en función de la finalidad del análisis (generalización de la fiabilidad o comparación entre grupos). En la generalización de la fiabilidad se incluyeron los trastornos de forma independiente, ya que al tratarse de muestras independientes no implicaba un incumplimiento de los supuestos del metaanálisis. En el caso de la comparación de medias no se podía optar por esta solución, ya que al mantenerse el mismo grupo control habría que duplicarlo y el cálculo del efecto de este grupo no cumpliría la independencia de las medidas. En este caso se optó por eliminar el grupo de pacientes de ansiedad con menor número de sujetos.

Índices del tamaño del efecto. Para el análisis de la comparación de medias se utilizó la *d* de Cohen y para la generalización de la fiabilidad el alfa de Cronbach.

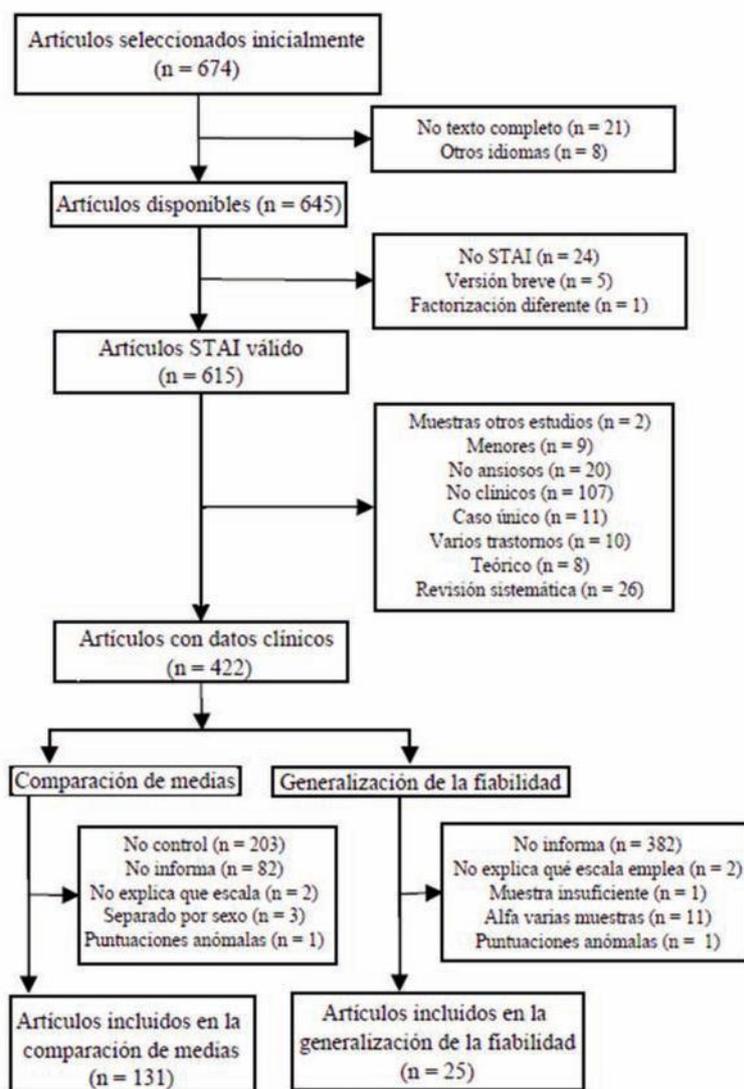
Extracción de datos. De cada uno de los artículos seleccionados se extrajeron las siguientes medidas para efectuar la media en el metaanálisis:

- Medidas *d* de Cohen para el tamaño del efecto en las comparaciones entre grupos. En el caso de no informar del mismo, se extrajeron de los valores *t*-student o *F* y los grados de libertad, para poder estimarlo. En la mayoría de los casos se emplearon las medias, las desviaciones típicas o varianzas y los tamaños de cada grupo, con el fin de estimar el tamaño del efecto para cada uno de los estudios.

- En el caso de la generalización de la fiabilidad se extrajeron los valores del alfa de Cronbach en todos los casos, ya que no se hallaron estudios que informasen de otras medidas similares para el cálculo de la consistencia interna.

METAANÁLISIS DE COMPARACIÓN DE GRUPOS Y METAANÁLISIS DE GENERALIZACIÓN DE LA FIABILIDAD DEL CUESTIONARIO (STAI)

Figura 1
Algoritmo de selección de los artículos



Además de las variables para evaluar el efecto, también se incluyeron diversos moderadores para ver su influencia sobre una posible homogeneidad de la generalización de la fiabilidad o del tamaño del efecto. En primer lugar, se empleó el país donde se seleccionó a los participantes. En segundo

lugar se codificó la forma del STAI utilizada, pudiendo ser la forma X o Y. No se incluyó el uso de adaptaciones a los diferentes países (usando sólo forma X o Y) ya que estadiferencia se analizó en la variable del país de la muestra. Por último, se empleó el tipo de trastorno de ansiedad estudiado,

empleando la codificación del DSM-IV-TR¹². Como moderadores el tamaño muestral de los estudios se extrajo la edad media de los participantes y su distribución por sexo (codificada como porcentaje de mujeres en la muestra). Las diferentes medidas fueron extraídas por dos codificadores independientes con el fin de valorar la fiabilidad del proceso de codificación. Para ello se empleó el índice Kappa y el coeficiente de correlación intraclase, siendo el valor medio de todos los índices 0,844.

Análisis estadístico. Respecto a los análisis, en el caso de la comparación de grupos, para seleccionar un método de estimación de efectos fijos o aleatorios no se empleó ningún estadístico¹⁹. Así pues, para el presente estudio, se puede asumir variabilidad entre cada uno de los tamaños del efecto debido a que proceden de poblaciones diferentes¹⁹. Esta asunción se basa en que los participantes proceden de diferentes países, por tanto han cumplimentado diferentes adaptaciones del STAI e, igualmente, cada estudio se conforma por un grupo clínico, incluyendo diferentes trastornos de ansiedad en la base final. Por ello se optó por utilizar un método de efectos aleatorios. La medida del tamaño del efecto, seleccionada para realizar el metaanálisis, fue la *d* de Cohen²⁰. El método seleccionado es el DerSimonian-Laird²¹ que se considera adecuado para modelos de efectos aleatorios con una medida del tamaño del efecto continua como es la *d* de Cohen²². Se empleó el estadístico *Q* para comprobar la homogeneidad de los datos y se complementó con el índice *I*²¹⁸. En el caso de hallar heterogeneidad en los resultados y determinar la influencia de diversos moderadores sobre la misma se empleó ANOVAs para las variables cualitativas y se aplicó un modelo de regresión para las variables moderadoras continuas.

Para la generalización del alfa de Cronbach, se transformó este estadístico a valores *T* mediante la fórmula: $T=(1-\alpha)^{1/3}$, donde α es el coeficiente obtenido en cada estudio.

Estos valores permiten transformar el alfa de Cronbach en una medida con distribución normal, requisito necesario en los metaanálisis²³. Además, estas puntuaciones se ponderaron por el inverso de la varianzas de los estudios para calcular el tamaño medio de las puntuaciones *T*. Tras obtenerlo, los valores finales se volvieron a convertir en puntuaciones α con el fin de lograr un índice más fácilmente interpretable. El método fue de efectos fijos, ya que al contar con un número reducido de artículos no cabe asumir el cumplimiento de los supuestos necesarios para utilizar un modelo de efectos aleatorios. Para la realización de ambos metaanálisis se empleó el programa R, concretamente el paquete Metafor²⁴.

RESULTADOS

El número de participantes en total fue de 2.476 (mínimo 6 y máximo 122) en el grupo clínico y 3.780 (mínimo 7 y máximo 1.329) en el grupo control en el caso de la ansiedad estado. De 3.104 (mínimo 6 y máximo 122) en el grupo clínico y 4.513 (mínimo 7 y máximo 1.324) en el grupo control, para la ansiedad rasgo. De 1.182 (mínimo 20 y máximo 406) en el grupo clínico de generalización de la fiabilidad estado y 1.631 (mínimo 20 y máximo 406) en la fiabilidad de la ansiedad rasgo. La media de edad osciló entre 30,78 y 40,94 en el conjunto de los cuatro metaanálisis. La media de ansiedad estado fue de 44,94 y en ansiedad rasgo de 50,26. En el caso de la fiabilidad, el α medio sin transformar fue 0,9 para la ansiedad estado y 0,88 para ansiedad rasgo.

Con el test de Egger hubo indicios de un posible sesgo en el caso de la ansiedad rasgo ($Z=7,03$, $p<0,001$), pero no en la ansiedad estado ($Z=1,76$, $p=0,08$). Así pues, en el caso de la ansiedad estado se observó que el tamaño medio del efecto fue de $d=1,39$ (IC95%: 1,22-1,56) resultando estadísticamente significativo ($p<0,001$). Tras ello, se analizó la homogeneidad de los tamaños del efecto con el fin de determinar si había variabilidad entre los estudios. El test de

heterogeneidad resultó estadísticamente significativo ($Q(93)=634,51$ $p<0,001$). El indicador I^2 fue igual a 85,34%, lo que implicó una heterogeneidad alta.

En el caso de la ansiedad rasgo el tamaño del efecto medio fue de $d=1,74$ (IC95%: 1,56-1,91) siendo esta medida estadísticamente significativa ($p<0,001$). En las pruebas de heterogeneidad se observó que había variabilidad en las aportaciones de los diferentes estudios a la media ($Q(109)=910,52$; $p<0,001$). Este resultado se confirmó mediante el resto de estadísticos empleados ($I^2=88,03\%$).

En cuanto a los resultados de los AOVAs, en lo que respecta a ansiedad estado, el país ($Qm(17)=49,02$; $p<0,001$; $Qe(76)=419,08$; $p<0,001$) y la forma del STAI ($Qm(1)=7,61$; $p=0,006$; $Qe(91)=596,45$; $p<0,001$) resultaron estadísticamente significativos, mientras que el tipo de trastorno no ($Qm(7)=6,48$; $p=0,49$; $Qe(86)=602,57$; $p<0,001$). En el caso de la ansiedad rasgo el país ($Qm(18)=22,69$; $p=0,2$; $Qe(91)=742,25$; $p<0,001$) y la forma del STAI ($Qm(1)=0,85$; $p=0,356$; $Qe(107)=902,61$; $p<0,001$) no fueron estadísticamente significativos, al contrario que el tipo de trastorno ($Qm(7)=6,72$; $p<0,001$; $Qe(102)=176,39$; $p<0,001$). Seguidamente, se analizaron los niveles en los que se producía la variabilidad. En la tabla 1 se recoge un resumen de los coeficientes obtenidos para cada una de las categorías de dichas variables. Además se emplearon moderadores continuos: el tamaño muestra, la media en la puntuación del STAI y su desviación típica, la edad media y el porcentaje de mujeres en la muestra. De las variables continuas empleadas únicamente se observó capacidad predictiva de la variabilidad en la media para la ansiedad estado ($b=0,08$ [CI95%: 0,06-0,11]; $Q=44,01$, $p<0,001$; $Qe=507,65$, $p<0,001$; $R^2=0,22$), y de la media ($b=0,09$ [CI95%: 0,08-0,11]; $Qm=106,54$, $p<0,001$; $Qe=510,46$,

$p<0,001$; $R^2=0,5$) y la desviación típica en el caso de la ansiedad rasgo ($b=-0,10$ [CI95%: -0,16-0,14]; $Qm=9,89$, $p=0,002$; $Qe=827,42$; $p<0,001$; $R^2=0,09$).

El segundo tipo de metaanálisis incluido fue el de generalización de la fiabilidad. En este caso se obtuvieron 13 artículos para la escala de ansiedad estado y 19 artículos para la ansiedad rasgo, en los que se informa del α de Cronbach. De nuevo se comenzó por el análisis de la influencia de un posible sesgo de selección, observando problemas tanto en estado ($Z=4,93$, $p<0,001$) como en rasgo ($Z=6,09$, $p<0,001$). En los artículos sobre ansiedad estado el valor medio de α calculado sin ningún tipo de ponderación fue de 0,87 (DT=0,08); de 0,89 (DT=0,05) para la ansiedad rasgo.

Seguidamente, se calculó la media de las puntuaciones T ponderadas, observando un valor de 0,43 en ansiedad estado (EE=0,006) y de 0,45 (EE=0,005) para la ansiedad rasgo. Al transformar esta puntuación de nuevo a valores α de Cronbach, se obtuvo un $\alpha=0,92$ en ansiedad estado (IC95%: 0,91-0,93) y de 0,91 en ansiedad rasgo (IC95%: 0,90-0,92). En ambos casos este valor resultó estadísticamente significativo (ansiedad estado: $Z=71,24$, $p<0,001$; ansiedad rasgo: $Z=83,76$, $p<0,001$). Tras ello, se analizó la homogeneidad de los valores α en los diferentes artículos. En ambos casos se observa heterogeneidad: ansiedad estado: $Q(12)=106,19$ ($p<0,001$) y ansiedad rasgo: $Q(18)=70,86$ ($p<0,001$).

El siguiente paso fue analizar las fuentes de variabilidad. En la ansiedad estado todas las variables analizadas resultaron significativas: país ($Qm(4)=65,09$; $p<0,001$; $Qe(8)=41,1$; $p<0,001$), forma del STAI ($Qm(1)=6,88$; $p=0,009$; $Qe(11)=99,31$; $p<0,001$) y tipo de trastorno ($Qm(3)=19,26$; $p=0,001$; $Qe(9)=86,93$; $p<0,001$). Mientras, en la ansiedad rasgo, el país ($Qm(6)=40,11$;

Tabla 1
Coefficientes de los moderadores en el meta-ANOVA de comparación de grupos
para la ansiedad estado y ansiedad rasgo

Variables (k estado k rasgo) significativo estado /significativo rasgo	Ansiedad estado				Ansiedad rasgo			
	<i>d</i> de Cohen	Error estándar	Z	Intervalo de confianza 95%	<i>d</i> de Cohen	Error estándar	Z	Intervalo de confianza 95%
País del estudio								
Alemania (17;27)	1,07	0,19	5,76 [†]	0,71-1,44	1,51	0,18	8,55 [†]	1,17-1,86
Australia (6;6)	1,30	0,33	3,98 [†]	0,66-1,95	1,89	0,39	4,85 [†]	1,13-2,66
Canadá (1;1) [¶]	3,19	0,85	3,76 [†]	1,53-4,85	0,65	0,94	0,7	-1,18-2,49
China (2;2)	1,38	0,53	2,6 [†]	0,34-2,42	1,39	0,64	2,18 [§]	0,14-2,63
EEUU (21;29)	1,49	0,17	8,76 [†]	1,15-1,81	2,09	0,17	12,19 [†]	1,75-2,42
España (7;3)	1,02	0,28	3,62 [†]	0,47-1,57	1,14	0,52	2,2 [§]	0,12-2,15
Francia (4;4) [¶]	2,53	0,42	6,01 [†]	1,71-3,37	2,71	0,5	5,44 [†]	1,73-3,69
Holanda (2;1)	2,16	0,56	3,88 [†]	1,07-3,25	2,34	0,92	2,55 [§]	0,54-4,13
Israel (1;2)	1,14	0,72	1,6	-0,26-2,55	1,63	0,62	2,62 [†]	0,41-2,85
Italia (4;3)	0,95	0,41	2,32 [§]	0,15-1,74	0,87	0,54	1,61	-0,19-1,92
Japón (6;7)	1,76	0,32	5,47 [†]	1,13-2,40	1,93	0,35	5,45 [†]	1,24-2,63
Korea (3;3)	2,34	0,47	5,01 [†]	1,43-3,26	1,51	0,53	2,85 [†]	0,47-2,56
Nueva Zelanda (0;1)	---	---	---	---	3,22	0,94	3,43 [†]	1,38-5,06
Países Bajos (2;2)	1,62	0,55	2,96 [†]	0,55-2,69	1,16	0,64	1,8	-0,1-2,42
Polonia (1;1)	1,57	0,77	2,05 [§]	0,07-3,07	1,52	0,91	1,66	-0,27-3,3
Suecia (1;2)	1,16	0,76	1,54	-0,32-2,64	1,91	0,66	2,89 [†]	0,62-3,2
Suiza (1;1) ^{¶¶}	5,01	0,82	6,13 [†]	3,41-6,61	4,57	0,94	4,84 [†]	2,72-6,41
Turquía (3;4)	0,54	0,44	1,23	-0,32-1,41	2,08	0,47	4,41 [†]	1,15-3,0
UK (11;13)	1,6	0,23	6,87 [†]	1,14-2,05	2,01	0,26	7,72 [†]	1,5-2,52
Forma del STAI								
Forma Y (39;50) [¶]	1,78	0,14	12,56 [†]	1,51-2,06	2,01	0,14	14,46 [†]	1,74-2,28
Forma X (53;61)	1,24	0,12	10,28 [†]	1,01-1,48	1,71	0,13	13,66 [†]	1,47-1,96
Trastorno psicológico								
Crisis de angustia (16;16)	1,54	0,23	6,67 [†]	1,09-1,99	1,62	0,21	7,66 [†]	1,2-2,03
Fobia específica (11;17) [¶]	1,18	0,28	4,25 [†]	0,64-1,73	1,05	0,21	5,06 [†]	0,64-1,46
Fobia social (16;19) [¶]	1,89	0,23	8,07 [†]	1,43-2,35	2,33	0,2	11,56 [†]	1,93-2,72
Fobia social generalizada (4;4)	1,32	0,46	2,9 [†]	0,43-2,21	1,97	0,44	4,49 [†]	1,11-2,83
Ansiedad generalizada (2;12) [¶]	0,79	0,63	1,25	-0,44-2,02	2,76	0,26	10,71 [†]	2,26-3,27
Estrés post-traumático (16;19)	1,59	0,23	6,99 [†]	1,15-2,04	1,79	0,2	9,08 [†]	1,4-2,17
Obsesivo-compulsivo (26;23)	1,33	0,18	7,48 [†]	0,98-1,68	1,84	0,18	10,22 [†]	1,49-2,19
Trastorno por estrés agudo (2;2)	1,25	0,63	2 [§]	0,03-2,48	0,61	0,58	1,05	-0,53-1,75

Z=puntuación Z del coeficiente; [†] p <0,001; [‡] p <0,01; [§] p <0,05. [¶] la *d* de Cohen media para dicha categoría es estadísticamente diferente de la *d* de Cohen media total.

METAANÁLISIS DE COMPARACIÓN DE GRUPOS Y METAANÁLISIS DE GENERALIZACIÓN DE LA FIABILIDAD DEL CUESTIONARIO (STAI)

Tabla 2
Coefficientes de los moderadores en el meta-ANOVA de generalización de la fiabilidad para la ansiedad estado y rasgo

Variables (k estado k rasgo) significativo estado/significativo rasgo	Ansiedad estado				Ansiedad rasgo			
	α^*	Error estándar	Z	Intervalo de confianza 95%	α^*	Error estándar	Z	Intervalo de confianza 95%
País del estudio								
Alemania (-;3)	---	---	---	---	0,91	0,02	21,65 [†]	0,89-0,94
Canadá (2;1) [¶]	0,89	0,01	44,72 [†]	0,87-0,90	0,94	0,01	41,59 [†]	0,93-0,95
EEUU (7;10) [¶]	0,94	0,01	47,5 [†]	0,93-0,95	0,91	0,01	55,87 [†]	0,9-0,92
Francia (-;2)	---	---	---	---	0,87	0,02	24,08 [†]	0,83-0,90
Holanda (1;1) [¶] / [¶]	0,93	0,02	22,89 [†]	0,91-0,95	0,81	0,03	22,89 [†]	0,76-0,85
Italia (1;1)	0,83	0,04	13,17 [†]	0,74-0,90	0,92	0,03	13,17 [†]	0,88-0,95
Turquía (-;1)	---	---	---	---	0,92	0,03	14,7 [†]	0,88-0,95
Forma del STAI								
Forma Y (7;13) [¶]	0,9	0,03	15,33 [†]	0,87-0,94	0,9	0,02	24,89 [†]	0,88-0,93
Forma X (4;6)	0,93	0,04	11,24 [†]	0,88-0,96	0,89	0,03	17,32 [†]	0,85-0,92
Trastorno psicológico								
Fobia específica (0;3)	---	---	---	---	0,93	0,04	11,35 [†]	0,88-0,96
Fobia social (0;2) [¶]	---	---	---	---	0,9	0,05	10,49 [†]	0,82-0,95
Fobia social generalizada (0;1) [¶]	---	---	---	---	0,84	0,06	8,65 [†]	0,70-0,93
Obsesivo-compulsivo (2;5)	0,95	0,05	7,12 [†]	0,90-0,98	0,9	0,03	16,1 [†]	0,86-0,93
Estrés post-traumático (6;5) [¶]	0,9	0,03	15,05 [†]	0,84-0,93	0,91	0,03	15,71 [†]	0,88-0,95
Ansiedad generalizada (2;1) [¶]	0,91	0,05	8,3 [†]	0,83-0,96	0,84	0,06	8,62 [†]	0,70-0,93
Varios trastornos (1;2) [¶] / ^{**}	0,93	0,07	5,72 [†]	0,83-0,98	0,88	0,04	11,91 [†]	0,8-0,93

[†] Los coeficientes han sido retransformados mediante $\alpha=1-T3$. Z: puntuación Z del coeficiente; [†] $p<0,001$; ^{**} Incluye artículos en los que se emplean varios trastornos juntos siempre que éstos sean exclusivamente trastornos de ansiedad. [¶] el α medio para dicha categoría es estadísticamente diferente del α medio total.

$p<0,001$; $Q_e(12)=30,76$; $p=0,002$) y el tipo de trastorno ($Q_m(5)=33,68$; $p<0,001$; $Q_e(13)=37,18$; $p<0,001$) resultan significativas pero no la forma del STAI ($Q_m(1)=0,1$; $p=0,751$; $Q_e(17)=70,76$; $p<0,001$). Un resumen de los estadísticos por cada nivel de las variables se recoge en la tabla 2. Para los moderadores continuos, debido al bajo número de estudios, los resultados fueron estadísticamente significativos pese a lo cual las betas eran inferiores a 0,01, salvo en el caso de la desviación típica como moderador de la variabilidad en la fiabilidad de la ansiedad rasgo ($b=-0,03$ [CI95%: -0,4--0,02]; $Q_m=31,62$, $p<0,001$; $Q_e=38,55$, $p<0,001$).

DISCUSIÓN

Tras el análisis se observa que, en ambos casos, ansiedad estado y rasgo, los tamaños del efecto medio son elevados. Así pues, el STAI es sensible a los niveles de ansiedad clínicos, detectando puntuaciones significativamente más elevadas que en un grupo control con sujetos sanos. En cuanto a la fiabilidad media de la escala para pacientes diagnosticados de ansiedad los valores también son elevados, observando una fiabilidad media superior a 0,9 en ambas subescalas. Por ello, se puede concluir que el STAI es un instrumento fiable a la hora de evaluar los niveles de ansiedad en pacientes con

ansiedad y que es sensible a la hora de detectar los niveles superiores de ansiedad (estado y rasgo) de los mismos, a través de sus diferentes formas y en diversos países. Además de ello, el STAI es sensible a la ansiedad provocada por diversos trastornos ansiosos.

Al realizar el metaanálisis sobre el efecto medio se ha observado variabilidad de las medidas a lo largo de los estudios. A la hora de explicar la variación de los tamaños del efecto hay diferencias entre países. No obstante, los tamaños del efecto son adecuados (abarcando desde 0,44 a 5,01 en ansiedad estado y de 0,54 a 4,57 en ansiedad rasgo). Además, salvo en los artículos de Alemania, Italia y Turquía, todos los tamaños de efecto medios son superiores a 1. Así pues el STAI detecta diferencias entre los sujetos con ansiedad y los sujetos sanos tal y como cabría esperar^{3,5}. En lo que respecta a la forma del STAI, en ambos casos los tamaños del efecto son similares, pese a que la forma X tiene un tamaño del efecto medio menor que la forma Y, tanto en ansiedad estado como en ansiedad rasgo. Lo que indica que ambas formas son adecuadas para detectar los niveles de ansiedad, pudiéndose utilizar el STAI para la evaluación de la ansiedad⁸⁻⁹. El último moderador de la variabilidad empleado fue el tipo de trastorno, mostrando que el STAI permite detectar ansiedad general de forma adecuada con puntuaciones mayores entre diversos trastornos, respecto a población general, siendo las diferencias estadísticamente significativas.

La generalización de la fiabilidad permite inferir si un test es fiable a lo largo de muestras procedentes de diversas poblaciones²³. En este caso se observa que la fiabilidad media es elevada. Además, al analizar por países, la fiabilidad media es adecuada. Al estimar la fiabilidad media por cada uno de los trastornos los valores superan el 0,86 en todos los casos. Estos resultados son coherentes ya que se encuentra un notable número de artículos en los que se afirma que el

STAI tiene una elevada fiabilidad¹¹ y un metaanálisis en el que se establecen valores medios de alfa de Cronbach de 0,91 en ansiedad estado y de 0,89 en ansiedad rasgo⁷. Así pues en población general la fiabilidad es similar a la obtenida, pese a que en el trastorno de ansiedad generalizada en el caso de la subescala de ansiedad rasgo el alfa sea ligeramente inferior. Por último, al analizar la forma del STAI se observa que la forma Y tiene 0,91 para ambas escalas, mientras que la forma X tiene una fiabilidad de 0,93 para ansiedad estado y de 0,91 para ansiedad rasgo. Generalmente, se afirma que la versión Y del STAI tiene una mayor fiabilidad⁵⁻⁶. El hecho de que en este metaanálisis se observe lo contrario puede deberse a que el tamaño de las muestras de sujetos con diagnóstico de ansiedad suele ser menor que el de las muestras de población general, lo que puede afectar a la fiabilidad.

En lo que respecta a España se observa que hay diversos artículos en los que se emplea una muestra clínica, obteniendo un tamaño del efecto medio elevado. Así pues, la adaptación española del STAI detecta puntuaciones superiores en sujetos con ansiedad. Pese a ello, dicha adaptación no cuenta con una muestra de personas con diagnóstico de ansiedad ni de baremos específicos para ellas²⁵. De la misma forma no hay fiabilidad para sujetos con ansiedad en la adaptación española. De hecho, en la revisión realizada no se ha encontrado ningún artículo con muestra española que informe de la fiabilidad. Pese a ello, en población general la fiabilidad del STAI es adecuada²⁶.

En resumen, el STAI parece ser un instrumento fiable en muestras formadas por sujetos ansiosos detectando mayores niveles de ansiedad en estas respecto a muestras control. Por ello, sería conveniente que las diversas adaptaciones del STAI tuviesen baremos con personas con ansiedad así como versiones breves del mismo donde la fiabilidad varía²⁷.

Como limitación del estudio cabe destacar la posible influencia de otros moderadores que expliquen la variabilidad hallada en los tamaños del efecto y los coeficientes de fiabilidad medios. Otra limitación es el hecho de que pese a haber incluido búsquedas que permitiesen localizar la literatura gris, en los resultados se observa un sesgo de selección, por lo que sería interesante replicar este metaanálisis incluyendo nuevas bases y búsquedas. Pese a ello, este estudio permite valorar la fiabilidad media del STAI, uno de los más utilizados en la práctica psicológica.

BIBLIOGRAFÍA

1. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen, HU. Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *Int J Methods Psychiatr Res.* 2012; 21: 169-184. <http://dx.doi.org/10.1002/mpr.1359>
2. Kessler RC, Ruscio AM, Shear K, Wittchen HU. Epidemiology of Anxiety Disorders. En: Behavioral Neurobiology of Anxiety and Its Treatment. Libro electrónico, 2009, http://dx.doi.org/10.1007/7854_2009_9
3. Somers JM, Goldner EM, Waraich P, Hsu L. Prevalence and Incidence Studies of Anxiety Disorders: A Systematic Review of the Literature. *Can J Psychiatry.* 2006;51:100-113.
4. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet.* 2007;370:841-850.
5. Spielberger CD, Gorsuch R, Lushene R. Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologist Press; 1970.
6. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. State-Trait Anxiety Inventory (Form Y). Palo Alto: Consulting Psychologists Press; 1983.
7. Barnes LLB, Harp D, Jung WS. Reliability generalization of scores on the Spielberger State-Trait anxiety Inventory. *Educ Psychol Meas.* 2002;62:603-618.
8. Spielberger CD, Reheier EC. Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Appl Psychol Health Well Being.* 2009;1:271-302. <http://dx.doi.org/10.1111/j.1758-0854.2009.01017.x>
9. Fisher PL, Durham RC. Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychol Med.* 1999;29:1425-1434.
10. Kvaal K, Ulstein I, Nordhus IH, Engedal K. The Spielberger State-Trait Anxiety Inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *Int J Geriatr Psychiatry.* 2005;20:629-634.
11. Rossi V, Pourtois G. Transient state-dependent fluctuations in anxiety measured using STAI, POMS, PANAS or VAS: a comparative review. *Anxiety Stress Coping.* 2012; 25: 603-645. <http://dx.doi.org/10.1080/10615806.2011.582948>
12. American Psychological Association. Manual diagnóstico y estadístico de los trastornos mentales. DSM-IV-TR. Barcelona: Masson; 2002.
13. Montero I, León OG. A guide for naming research studies in Psychology. *Int J Clin Health Psychol.* 2007;7:847-862.
14. Urrútia G, Bonfill X. La Declaración PRISMA: un paso adelante en la mejora de las publicaciones de la Revista Española de Salud Pública. *Rev Esp Salud Pública.* 2013;87: 99-102.
15. Hartley J. New ways of making academic articles easier to read. *Int J Clin Health Psychol.* 2012;12:143-160.
16. Perestelo-Pérez L. Standards on how to develop and report systematic reviews in Psychology and Health. *Int J Clin Health Psychol.* 2013;13:49-57.
17. Fernández-Ríos L, Buela-Casal G. Standards for the preparation and writing of Psychology review articles. *Int J Clin Health Psychol.* 2009;9:329-344.
18. Sánchez-Meca J. Cómo realizar una revisión sistemática y un meta-análisis. *Aula Abierta.* 2010;38:53-64.
19. Sánchez-Meca J, Marín Martínez F, Huedo Medina TB. Modelos de efectos fijos y modelo de efectos aleatorios. En: Revisiones sistemáticas en las ciencias de la vida. Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-La Mancha; 2006. p. 189-204.
20. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1969.
21. DerSimonian R. y Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7:177-188.

Alejandro Guillén-Riquelme et al.

22. Espallargues i Carreras M, Tebé Cordoní C. Qué resultados combinar y cómo combinarlos. En: Revisiones sistemáticas en las ciencias de la vida. Castilla la Mancha: Fundación para la Investigación Sanitaria en Castilla-La Mancha; 2006. p. 173-187

23. Sánchez-Meca J, López-Pina JA. El enfoque meta-analítico de generalización de la fiabilidad. *Acción Psicológica*. 2008;31:7-17.

24. Viechtbauer, W. metafor: Meta-Analysis Package for R. " R package version 1.4-0. 2010. <http://CRAN.R-project.org/package=metafor>.

25. Buéla-Casal G, Guillén-Riquelme A, Seisdedos Cubero N. Cuestionario de Ansiedad Estado-Rasgo: Adaptación española. 8ª ed. Madrid: TEA Ediciones; 2011.

26. Guillén-Riquelme A, Buéla-Casal G. Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory. *Psicothema*. 2011;23:510-515.

27. Guillén-Riquelme A, Buéla-Casal G. Versión Breve del STAI en Adolescentes y Universitarios Españoles. *Ter Psicol*. 2013;31:293-299.

Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI)

Alejandro Guillén-Riquelme y Gualberto Buela-Casal
Universidad de Granada

La ansiedad es uno de los problemas psicológicos con mayor prevalencia. Entre los instrumentos para medirla se encuentra el State Trait Anxiety Inventory. Este cuestionario mide ansiedad rasgo (factor de personalidad que predispone a sufrir o no ansiedad) y ansiedad estado (los factores ambientales que protegen o generan ansiedad). La adaptación española del cuestionario se realizó en 1982, por ello el objetivo del estudio es realizar una revisión del STAI. Para ello, se reunió una muestra de 1.036 adultos. Se realizó un análisis de fiabilidad mediante alfa de Cronbach (0,90 para ansiedad rasgo y 0,94 para ansiedad estado). También se realizó una reducción factorial con unos resultados similares a los de la adaptación original. Además, se comprobó si existía funcionamiento diferencial de ítem por sexo y únicamente una de las 40 preguntas mostró problemas. Por último se realizó comparaciones t-Student con los valores de la adaptación; mientras el rasgo varía en 1 punto, el estado tiene diferencias de hasta 6 puntos. En estos resultados de forma general se observa que el STAI mantiene unas adecuadas propiedades métricas y que, además, ha sido sensible al aumento de estímulos ambientales que producen estrés.

Psychometric revision and differential item functioning in the State Trait Anxiety Inventory (STAI). One of the psychological problems with highest prevalence is anxiety. The State Trait Anxiety Inventory is one of the instruments to measure it. This questionnaire assesses Trait Anxiety (understood as a personality factor that predisposes one to suffer from anxiety) and State Anxiety (refers to environment factors that protect from or generate anxiety). The questionnaire was adapted in Spain in 1982. Therefore, the goal of the study is to review the current psychometric properties of the STAI. A total of 1036 adults took part in the study. Cronbach's alpha reliability was .90 for Trait and .94 for State Anxiety. Factor analysis showed similar results compared with the original data. Moreover, differential item functioning (DIF) was carried out to explore sex bias. Only one of the 40 items showed DIF problems. Lastly, a t-Test was run, comparing the original and current values; whereas Trait Anxiety varied in 1 point, State Anxiety had differences of up to 6 points. In general, this result shows that the STAI has maintained adequate psychometric properties and has also been sensitive to increased environmental stimuli that produce stress.

En la actualidad, la ansiedad es uno de los problemas psicológicos con mayor prevalencia. Alonso et al. (2004b) indican que, en una muestra representativa de seis países europeos, un 13,6% había tenido algún problema de ansiedad a lo largo de su vida y un 6,4% en el último año; resultado similar al obtenido por Andlín-Sobocki y Wittchen (2005), nuevamente en europeos, donde un 12% presentaba trastornos de ansiedad. De la misma forma Wang et al. (2007) hallaron que un 6,8% de los encuestados españoles había tenido problemas por ansiedad y trastornos emocionales en el último año. Además, un 26,1% de las consultas psiquiátricas formales se deben a trastornos de ansiedad (Alonso et al., 2004c) y un 28,3% del total de casos de

ansiedad muestran comorbilidad con otros trastornos (Alonso et al., 2004a).

Existen numerosos cuestionarios para medir la ansiedad (Inventario de Ansiedad de Beck, Inventario de Situación y Respuestas Ansiosas, Escala de Ansiedad Manifiesta de Taylor...). Entre los cuestionarios para evaluar este trastorno se encuentra el State-Trait Anxiety Inventory (STAI), desarrollado por Spielberger, Gorsuch y Lushene (1970). Este instrumento fue creado para evaluar la ansiedad según el modelo del propio Spielberger, que postulaba que dicho trastorno está constituido por dos componentes: un factor de personalidad que comprendería las diferencias individuales, relativamente estables, para responder ante situaciones percibidas como amenazantes con una subida en la ansiedad. También se define como una tendencia a percibir las situaciones como más amenazantes (ansiedad rasgo). El segundo factor (ansiedad estado) hace referencia a un período transitorio caracterizado por un sentimiento de tensión, aprensión y un aumento de la actividad del sistema nervioso autónomo, pudiendo variar tanto en el tiempo como en la intensidad (Spielberger et al., 1970). La elección de

Fecha recepción: 20-10-10 • Fecha aceptación: 16-2-11
Correspondencia: Alejandro Guillén-Riquelme
Facultad de Psicología
Universidad de Granada
18011 Granada (Spain)
e-mail: agr@ugr.es

este cuestionario sobre el resto de pruebas para evaluar ansiedad se debe a que el STAI es uno de los más utilizados; de hecho Muñoz y Fernández-Hermida (2010) encontraron que es el 7º cuestionario más utilizado en España.

Cada una de las escalas del STAI (ansiedad rasgo y ansiedad estado) está compuesta por 20 ítems; una parte de los mismos redactada de forma positiva y otra de forma negativa. Diversos estudios mostraron reducciones factoriales muy robustas con cuatro factores. Estos factores estaban basados en la inversión de parte de las preguntas y se definían como ansiedad rasgo afirmativa y negativa y ansiedad estado afirmativa y negativa (Hishinuma, Miyamoto, Nishimura y Nahulu, 2000; Suzuki, Tsukamoto y Abe, 2000; Vigneau y Cormier, 2008). En la adaptación española (Spielberger, Gorsuch y Lushene, 1982) se encuentran resultados muy similares. El número de factores que subyacen en el STAI han sido objeto de numerosos análisis, existiendo defensores del modelo bifactorial, tetrafactorial y, en menor medida, un modelo penta-factorial, compuesto por los factores presentados anteriormente más un factor, compuesto por dos ítems, llamado *happiness* (Suzuki et al., 2000).

Este instrumento ha sido traducido a más de cuarenta lenguas (Spielberger y Díaz-Guerrero, 1976, 1983, 1986; Spielberger, Díaz-Guerrero y Strelau, 1990) y se han realizado más de 3.000 investigaciones que lo utilizan (Spielberger, 1989). En el 2002, Barnes, Harp y Jung realizaron una revisión del STAI y encontraron 909 artículos (publicados desde el año 1990 hasta el 2000) que lo citaban. Así pues, hasta el año 2000 se realizaron, al menos, 3.909 investigaciones con el cuestionario.

El STAI se adaptó al español a partir de los trabajos de Bermúdez (1978a; 1978b) y su versión comercial se realizó por la sección de estudios de TEA Ediciones (Spielberger et al., 1982). En estos artículos se comprueba que las medias y la fiabilidad (alfa de Cronbach y correlación test-retest) resultan similares a la versión original. Respecto a la dimensionalidad de la escala, se mantiene el modelo de cuatro factores presentado anteriormente.

El objetivo del presente estudio es analizar las propiedades psicométricas del STAI, puesto que su adaptación española se realizó hace casi treinta años, y comprobar su fiabilidad. Entre los análisis se incluirá una prueba del funcionamiento diferencial del ítem para analizar posibles sesgos, producidos por el efecto del sexo. Este análisis es pertinente puesto que no se ha realizado anteriormente en la adaptación española. Se espera que los índices psicométricos del cuestionario sean adecuados; pese a ello también es posible que los valores normativos hayan variado respecto al original por el paso del tiempo.

Método

Participantes

La muestra está compuesta por 1.036 participantes (352 hombres y 667 mujeres). Las edades oscilan desde 18 hasta los 69 años. La edad media es de 29,64 (DT= 10,51). Las mujeres tienen una edad media de 29,83 años (DT= 10,31) y los hombres de 29,30 años (DT= 10,91).

Instrumentos

Para el estudio se utilizó el State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1982). La adaptación española fue realizada por la sección de estudios de TEA.

Este inventario se compone de veinte ítems para cada una de las subescalas (ansiedad estado y ansiedad rasgo). La escala de respuesta es de tipo Likert, puntuando desde 0 (nada) hasta 3 (mucho). Los totales se obtienen sumando los valores de los ítems (tras las inversiones de las puntuaciones en los ítems negativos). Por ello, los totales de ansiedad rasgo y de ansiedad estado abarcan desde 0 hasta 60, correspondiéndose una mayor puntuación con mayor ansiedad detectada. No cuenta con un criterio clínico en su versión española, aunque sí lo tiene en la versión original, así como en muchas adaptaciones.

Procedimiento

La muestra se recogió en Granada, Córdoba, Jaén, Madrid, Barcelona, Valencia, Ourense, Santiago de Compostela y Bilbao. El tipo de muestreo fue no probabilístico por cuotas (ser mayor de edad y no estar cursando una carrera universitaria). Para el muestreo se contó con la colaboración de organizaciones públicas y privadas. También se administró el cuestionario en lugares con una gran afluencia de personas, como aeropuertos, estaciones, centros comerciales, etc. El cuestionario se administró por un único investigador. Se daba el test a los sujetos y era completado en presencia del administrador para posibles dudas o problemas. En todos los casos se informó a los participantes de la confidencialidad del estudio.

Diseño

Según la clasificación propuesta por Montero y León (2007), el presente estudio sería instrumental, puesto que evalúa las propiedades psicométricas de un cuestionario. Se siguieron las recomendaciones metodológicas de Carretero-Dios y Pérez (2007).

Análisis de datos

En el análisis de la fiabilidad se empleó el alfa de Cronbach. La comprobación de la estructura de la escala se realizó mediante análisis factorial. El método de extracción ha sido máxima verosimilitud y la rotación varimax. Para establecer el número de factores se realizó un análisis paralelo al 95% de confianza.

Para comparar los datos del estudio con los de la adaptación española se utilizó la prueba t-Student para una sola muestra, complementados con el tamaño del efecto mediante la *d* de Cohen.

Se comprobó si había funcionamiento diferencial del ítem (DIF en adelante). El DIF se define como la diferencia en la probabilidad que tienen los participantes de dos o más grupos, en este caso sexo, de responder de forma similar, tras haber sido igualados en habilidad (Millsap y Meredith, 1992). Según Mellenbergh (1995) se distingue entre DIF uniforme (la probabilidad de responder correctamente al ítem es mayor para uno de los grupos uniformemente para todos los niveles de la variable) y DIF no uniforme (la diferencia de probabilidad para responder correctamente al ítem no es igual a lo largo de los diferentes niveles del atributo).

Para la detección del DIF existen varios procedimientos. En el presente artículo se usará el análisis de regresión logística, propuesto por Miller y Spray (1993). Esta técnica permite detectar el DIF uniforme y no uniforme en ítems politómicos (Hidalgo y López-Pina, 2004). Los análisis, tal como se presentan en Hidalgo y Gómez (2006), consisten en un modelo de regresión logística binaria donde la variable dependiente es el grupo (en este caso

sexo) y las variables independientes son la puntuación total en el test, la respuesta al ítem y la interacción entre ambas. La estrategia para detectar DIF siguiendo las indicaciones de Hidalgo, Gómez y Padilla (2005), consiste en ajustar tres modelos en distintas etapas. En la primera etapa se introduce en la ecuación la puntuación total del sujeto en el test, ajustando así el Modelo 1 (M1) en base a la ausencia de DIF. En la segunda etapa se añade la variable de agrupamiento (Modelo 2, M2). Si la explicación de este modelo respecto al M1 fuese significativa sería indicio de DIF uniforme. En la etapa 3 se introduce la interacción entre el grupo y la puntuación total (Modelo 3, M3). Si la explicación que añade este modelo, respecto a los anteriores, fuese significativa habría DIF no uniforme.

Una vez detectados los ítems con DIF se eliminan sus puntuaciones del total del cuestionario. Tras la purificación se repiten los análisis en todos los ítems con el total corregido. Si aparecen nuevos ítems con DIF se puede proceder a una segunda depuración del total.

Por último, en los ítems con DIF cabe averiguar si se mantiene en todas las escalas de respuesta. Para ello se utiliza el análisis de categorías acumuladas propuesto por Mellenbergh (1995). El análisis consiste en varias dicotomizaciones de las categorías de respuesta. Se repite el análisis de DIF con las diferentes agrupaciones realizadas. En este caso se compara el 0 con el 1, 2 y 3 agrupados, el 0 y 1 contra el 2 y 3, y el 0, 1 y 2 contra el 3.

Para la realización de todos los análisis descritos se utilizó el programa estadístico SPSS 15.0, excepto para el cálculo del tamaño del efecto mediante la *d* de Cohen, realizado con el G*Power 3.1.2.

Resultados

Primeramente se realizó un análisis de las puntuaciones medias y desviación típica de los ítems para detectar puntuaciones extremas. Ninguna de las medias supera 1,78, siendo la menor de 0,45. Este valor, pese a ser bajo, no es de extrañar, puesto que no se trata de una muestra clínica, no esperando por ello puntuaciones elevadas. Las desviaciones típicas se sitúan entre 0,71 y 1,11. Esto indica que los sujetos no tienen un patrón de respuesta con altas frecuencias en los valores centrales o en los extremos.

Seguidamente se realizó un análisis de fiabilidad, mediante el alfa de Cronbach, obteniendo, para los ítems de ansiedad rasgo, un alfa de 0,90. Solo uno de los veinte ítems de la escala mejoraba el alfa (en una milésima) si era eliminado. Por su parte, la ansiedad estado alcanzó un alfa de 0,94, con un solo ítem cuya eliminación mejoraba el alfa en cuatro milésimas.

Una vez establecido el funcionamiento adecuado de los ítems se pasó a la reducción factorial de los mismos para ver la estructura factorial subyacente. En primer lugar, la prueba Kaiser-Meyer-Olkin (KMO= 0,96), así como la prueba de esfericidad de Bartlett ($\chi^2_{,780} = 19,37$; $p < 0,001$), mostraron la adecuación de la muestra para realizar el análisis factorial. Para la elección del número de factores se utilizó un análisis paralelo. Se extrajeron cuatro factores. Sus autovalores son 14,40, 2,60, 2,48 y 1,33, y la varianza total explicada conjuntamente es de 52,04%. Las saturaciones de cada ítem en los diferentes factores se presentan en la tabla 1.

En las saturaciones se observa que el factor 3 corresponde a los ítems de la ansiedad rasgo redactados de forma positiva, mientras que los ítems que saturan en el factor 2 pertenecen a la ansiedad estado redactados de forma positiva. Los otros dos factores, espe-

cialmente el primero, aglutinan ítems de ambas subescalas cuya redacción original es negativa y sus puntuaciones se han invertido. Esta estructura es similar a la hallada en la adaptación original de cuatro factores (ansiedad estado y rasgo, positiva y negativa, respectivamente). El hecho de establecer cuatro factores en función de la presencia o ausencia de ansiedad estado y rasgo se puede considerar un artefacto estadístico, por lo que se interpretará el resto de análisis en función de ansiedad estado y rasgo.

En siguiente lugar se analizaron las medias del total de las subescalas. Se calcularon pruebas t-Student sobre una media, utilizando para el contraste los valores del manual de la adaptación

Tabla 1
Matriz de saturaciones de los ítems de STAI en los factores rotados

	Ítem	Factor 1	Factor 2	Factor 3	Factor 4
Rasgo	1	0,67	0,14	0,31	0,11
	2	0,28	0,09	0,32	-0,01
	3	0,29	0,24	0,47	0,08
	4	0,32	0,15	0,42	0,07
	5	0,13	0,12	0,40	0,17
	6	0,44	0,15	0,22	0,06
	7	0,23	0,25	0,25	0,10
	8	0,30	0,25	0,45	0,16
	9	0,11	0,15	0,64	0,19
	10	0,70	0,11	0,23	0,09
	11	-0,04	0,11	0,46	-0,02
	12	0,20	0,15	0,59	0,51
	13	0,37	0,09	0,38	0,61
	14	0,08	0,15	0,34	0,11
	15	0,40	0,20	0,51	0,06
	16	0,68	0,07	0,30	0,20
	17	0,13	0,20	0,59	0,11
	18	0,19	0,12	0,60	0,06
	19	0,39	0,17	0,30	0,28
	20	0,19	0,26	0,55	0,07
Estado	1	0,40	0,59	0,10	0,24
	2	0,54	0,35	0,16	0,56
	3	0,22	0,75	0,19	0,08
	4	0,28	0,60	0,20	0,07
	5	0,63	0,44	0,07	0,13
	6	0,15	0,79	0,17	0,13
	7	0,20	0,38	0,37	-0,04
	8	0,48	0,37	0,14	0,08
	9	0,30	0,60	0,26	0,08
	10	0,63	0,33	0,09	0,10
	11	0,43	0,20	0,25	0,65
	12	0,12	0,74	0,24	0,13
	13	0,01	0,33	0,15	0,00
	14	0,26	0,49	0,28	0,01
	15	0,52	0,54	0,09	0,12
	16	0,72	0,21	0,14	0,24
	17	0,34	0,58	0,30	0,06
	18	0,13	0,70	0,20	0,11
	19	0,74	0,19	0,17	0,14
	20	0,67	0,37	0,13	0,16

Nota: se ha empleado el método de máxima verosimilitud y la rotación varimax

Sexo	Factor	Media en la adaptación original	Media en el presente estudio	Diferencia de medias	t	g.l. ²	p ³	d de Cohen
Hombres	Rasgo	20,19 (8,89)	18,96 (10,00)	1,23	-2,24	329	0,02	0,20
	Estado	20,54 (10,56)	15,87 (9,92)	6,67	-8,48	322	0,00	0,56
Mujeres	Rasgo	24,99 (10,05)	23,35 (10,60)	1,64	-3,85	617	0,00	0,10
	Estado	23,30 (11,93)	18,20 (11,62)	5,10	-10,58	604	0,00	0,45

¹ Datos obtenidos del manual de la adaptación española. Entre paréntesis se incluye la desviación típica
² Grados de libertad
³ Probabilidad

española. Este análisis se realizó para ambos sexos, ya que son los valores que ofrece la adaptación.

En la tabla 2 se puede observar, además de que todos los valores son significativos, que en las puntuaciones rasgo las diferencias son cercanas a 1,5 puntos, mientras que en la ansiedad estado las diferencias están en torno a los 5 puntos. Los tamaños del efecto son muy bajos en rasgo y elevados en la ansiedad estado. Parece ser que el rasgo no ha cambiado apenas respecto a los valores originales, mientras que la ansiedad estado sí ha sufrido fuertes modificaciones.

En la adaptación española se encontraron diferencias por sexo (Spielberger et al., 1982). Para comprobar si estas diferencias se mantienen en ambas subescalas se utilizó la prueba *t*-Student. En la ansiedad rasgo (asumiendo varianzas iguales; $F= 2,08$, $p= 0,150$) hay diferencias estadísticamente significativas ($t_{(946)} = -6,20$; $p<0,001$). Para la ansiedad estado (asumiendo varianzas distintas; $F= 17,60$, $p<0,001$) también se encuentran diferencias por sexo ($t_{(764,21)} = -3,19$; $p= 0,001$).

Para aportar evidencias de que las diferencias encontradas anteriormente no se deben a un problema de DIF se realizó un análisis del mismo. Los resultados se muestran en la tabla 3.

En la tabla 3 se observa que no se detecta DIF en ninguno de los ítems de ansiedad estado (significación menor a 0,01). Sin embargo, en la escala de ansiedad rasgo hay cuatro ítems (3, 5, 7 y 19) cuyo modelo 2 es significativo al 0,01, por lo que es un primer indicador de la presencia de DIF uniforme. Ningún ítem de ambas subescalas muestra DIF no uniforme. El siguiente paso es comprobar el tamaño del efecto mediante las diferencias en el R^2 del modelo 2 respecto al 1. Siguiendo los criterios propuestos por Jodoin y Gierl (2001), un valor inferior a 0,035 informaría de un DIF inapreciable; de 0,035 a 0,07 un DIF moderado, y mayor a 0,07 correspondería a un DIF elevado. En este caso el ítem 19 mostraría DIF moderado, mientras que, en los otros tres ítems, sería inapreciable. Seguidamente se depuró el total, eliminando estos cuatro ítems para repetir los análisis.

En la segunda fase se realizó de nuevo el procedimiento con el total depurado. Nuevamente los ítems 3, 5, 7 y 19 fueron significativos al 0,01, sin embargo los tamaños del efecto (excluyendo el del ítem 3, con un 0,061) son muy bajos, siendo el mayor de ellos de 0,026. Además de estos ítems, el 13, 14 y el 20 también fueron significativos, no obstante los tamaños del efecto son de 0,012, 0,010 y 0,010, respectivamente (inapreciables). Por todo ello, se puede concluir que únicamente el ítem 3 muestra un DIF medio.

Una vez se detecta el DIF, es necesario analizar si se encuentra en todas las categorías de respuesta o solo en algunas. Para ello, la escala del ítem 3 (único con DIF) se dicotomizó mediante las agrupaciones de Mellenbergh (1995). Con las categorías dicotomizadas se realizan los mismos análisis que anteriormente con los ítems. En ninguna de las alternativas de respuesta se detecta DIF no uniforme. Sin embargo, se observa DIF uniforme al comparar la opción cero contra la uno, dos y tres de forma conjunta. Por tanto, en el ítem 3 únicamente existe DIF al comparar la primera opción («casi nunca») con las otras tres. Es decir, que hombres y mujeres, igualados en nivel de ansiedad, responden con diferente frecuencia en esta opción.

Discusión

En los resultados se ha comprobado que las medias y desviaciones típicas de los ítems son adecuadas (Carretero-Dios y Pérez, 2007). Si bien algunas de las medias son bajas, esto se debe a que es un cuestionario de ansiedad (en su versión original y en muchas adaptaciones tiene baremos clínicos) administrado en población normal. Por ello, las medias algo bajas son esperables. Respecto a la fiabilidad los valores alfa de Cronbach son adecuados dando indicios de una adecuada consistencia interna.

El análisis factorial muestra un porcentaje de varianza explicada superior al 50%. Además, el análisis paralelo indica que se deben extraer cuatro factores, al igual que en la adaptación española (Spielberger et al., 1982). Vera-Villarreal, Buela-Casal y Spielberger (2007) también encuentran cuatro factores al realizar dos extracciones separadamente (una para rasgo y otra para estado), con una versión en español del cuestionario. Cabe indicar que estos factores no replican exactamente los esperados teóricamente, encontrando problemas en los factores que incluyen los ítems invertidos de ambas subescalas. Si bien estos datos muestran cierta confusión, son muy similares a los obtenidos por Bermúdez (1978a, 1978b), así como a los del manual de la adaptación española (Spielberger et al., 1982). Pese a que los datos no son tan claros como en otras revisiones de la versión en inglés, muestran una estructura tetrafactorial similar a las anteriormente halladas, aportando un nuevo indicio de que el test conserva sus propiedades métricas.

Respecto a la comparación de medias respecto al original los resultados son muy interesantes. En primer lugar cabe indicar que los baremos y medias de la adaptación española se realizó con 318 varones y 365 mujeres trabajadoras. El manual no ofrece más da-

Tabla 3
Funcionamiento diferencial del ítem: probabilidades (p) asociadas a los modelos (M) de la regresión y tamaño del efecto

Ítem	χ^2 (M2-M1)	p	χ^2 (M3-M2)	p	R Nagelkerke (M3-M2)	R Nagelkerke (M2-M1)	
Rasgo	1	4,07	0,131	0,09	0,955	0,000	0,006
	2	0,57	0,753	1,24	0,538	0,002	0,001
	3	45,10	0,000	6,30	0,043	0,009	0,061
	4	0,04	0,982	0,90	0,638	0,001	0,000
	5	20,95	0,000	0,07	0,967	0,000	0,029
	6	0,00	1,000	1,79	0,408	0,003	0,000
	7	14,14	0,001	0,05	0,977	0,000	0,020
	8	1,70	0,428	0,29	0,867	0,001	0,002
	9	2,42	0,298	0,03	0,985	0,000	0,003
	10	5,46	0,065	0,42	0,809	0,000	0,008
	11	3,01	0,222	0,74	0,692	0,001	0,004
	12	1,76	0,415	4,69	0,096	0,007	0,002
	13	8,29	0,016	0,05	0,978	0,000	0,012
	14	6,87	0,032	5,14	0,077	0,007	0,010
	15	4,85	0,089	2,58	0,275	0,003	0,007
	16	5,80	0,055	0,02	0,993	0,000	0,008
	17	0,16	0,925	1,19	0,552	0,002	0,000
	18	1,07	0,587	2,51	0,285	0,000	0,001
	19	12,40	0,002	0,05	0,974	0,000	0,017
	20	6,86	0,032	1,78	0,411	0,002	0,010
Estado	1	1,14	0,565	1,74	0,420	0,003	0,001
	2	1,08	0,582	4,94	0,085	0,008	0,001
	3	0,18	0,914	5,08	0,079	0,007	0,000
	4	6,57	0,037	2,00	0,368	0,003	0,009
	5	8,88	0,012	1,68	0,433	0,002	0,013
	6	1,40	0,497	2,00	0,368	0,003	0,002
	7	0,05	0,973	1,35	0,510	0,002	0,000
	8	2,88	0,237	0,83	0,661	0,001	0,004
	9	0,39	0,823	2,23	0,327	0,004	0,000
	10	0,02	0,988	1,23	0,541	0,002	0,000
	11	7,53	0,023	2,10	0,350	0,003	0,011
	12	3,77	0,152	3,90	0,142	0,006	0,005
	13	1,82	0,404	3,14	0,208	0,005	0,002
	14	0,18	0,912	2,54	0,282	0,004	0,000
	15	4,29	0,117	2,47	0,290	0,004	0,006
	16	1,38	0,501	0,00	1,000	0,000	0,002
	17	0,85	0,652	0,76	0,685	0,001	0,001
	18	1,14	0,565	4,10	0,129	0,006	0,001
	19	0,19	0,911	0,70	0,703	0,001	0,000
	20	3,10	0,212	4,24	0,120	0,007	0,004

tos al respecto. En base a ello, se observa que los tamaños muestrales son similares y no hay características diferentes en la presente muestra. Al comparar puntuaciones entre ambas muestras las medias difieren un máximo de 1,64 puntos en la ansiedad rasgo, mientras que alcanzan entre 5 y 6 puntos de diferencia en la ansiedad estado. Este hecho puede ser indicio de que el cuestionario conserva unas buenas propiedades métricas pese al paso del tiempo, ya que la escala de ansiedad rasgo mantiene las puntuaciones (los factores de personalidad son relativamente estables). Al contrario, en los factores ambientales la variación ha sido notable ya que éstos cambian con más facilidad a lo largo del tiempo (Eysenck, 1995).

En el análisis de DIF se encuentra un nuevo indicio sobre la validez del cuestionario. Se observa que únicamente un ítem mues-

tra funcionamiento diferencial por sexo. Para comprender mejor la causa del sesgo en la respuesta de hombres y mujeres cabe realizar un análisis del contenido del ítem. La cuestión planteada es «Siento ganas de llorar», en este caso parece lógico pensar que, pese a que se hayan igualado hombres y mujeres en nivel de ansiedad, los hombres serán más propensos a puntuar que casi nunca sienten ganas de llorar. Mientras, las mujeres tendrán una respuesta más equilibrada en las diferentes alternativas, lo que explica el DIF hallado.

En conclusión, que la escala se mantenga estable en la medida de la ansiedad rasgo pero haya variado en la ansiedad estado, además de mantener las medidas psicométricas y no verse afectado por DIF, son buenos indicios para considerar adecuada la utiliza-

ción de la escala. No obstante, cabe indicar que el estudio cuenta con la limitación de ser un muestreo incidental, lo que puede influir en las medidas. Para investigaciones futuras se debería replicar el

experimento en varias muestras e incluir medidas de la validez de contenido, criterio y predictiva. Además, habría que rehacer los baremos, especialmente para la subescala de ansiedad estado.

Referencias

- Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., et al. (2004a). 12-Month comorbidity patterns and associated factors in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109 (Suppl. 420), 28-37.
- Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., et al. (2004b). Prevalence of mental disorders in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109 (Suppl. 420), 21-27.
- Alonso, J., Angermeyer, M.C., Bernert, S., Bruffaerts, R., Brugha, T.S., Bryson, H., et al. (2004c). Use of mental health services in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. *Acta Psychiatrica Scandinavica*, 109 (Suppl. 420), 47-54.
- Audin-Sobocki, P., y Wittchen, H.U. (2005). Cost of anxiety disorders in Europe. *European Journal of Neurology*, 12 (Suppl. 1), 39-44.
- Barnes, L., Harp, D., y Jung, W.S. (2002). Reliability generalization of scores on the Spielberger State-Trait Anxiety Inventory. *Educational and Psychological Measurement*, 62, 603-618.
- Bermúdez, J. (1978a). Anxiety and performance. *Revista de Psicología General y Aplicada*, 151, 183-207.
- Bermúdez, J. (1978b). Functional analysis of anxiety. *Revista de Psicología General y Aplicada*, 153, 617-634.
- Carretero-Dios, H., y Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology*, 7, 863-882.
- Eysenck, H.J. (1995). Un modelo de personalidad: rasgos generales. En M.D. Avia y M.L. Sánchez-Bernados (Comps.), *Personalidad: aspectos cognitivos y sociales*. Madrid: Pirámide.
- Hidalgo, M.D., y Gómez, J. (2006). Nonuniform DIF detection using discriminant logistic analysis and multinomial logistic regression: A comparison for polytomous items. *Quality & Quantity*, 40, 805-823.
- Hidalgo, M.D., Gómez, J., y Padilla, J.L. (2005). Regresión logística: alternativas de análisis en la detección del funcionamiento diferencial del ítem. *Psicothema*, 17, 509-515.
- Hidalgo, M.D., y López-Pina, J.A. (2004). Differential item functioning detection and effect size: A comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement*, 64, 903-915.
- Hishinuma, E.S., Miyamoto, R.H., Nishimura, S.T., y Nahulu, L.B. (2000). Differences in State-Trait Anxiety Inventory Scores for ethnically diverse adolescents in Hawaii. *Cultural Diversity and Ethnic Minority*, 6, 73-83.
- Jodoin, M.G., y Gierl, M.J. (2001). Evaluating type I Error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14, 329-349.
- Mellenbergh, G.J. (1995). Conceptual notes on models for discrete polytomous item responses. *Applied Psychological Measurement*, 19, 91-100.
- Miller, T.R., y Spray, J.A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement*, 30, 107-122.
- Millsap, R.E., y Meredith, W. (1992). Inferential conditions in the statistical detection of measurement bias. *Applied Psychological Measurement*, 16, 389-402.
- Montero, I., y León, O.G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7, 847-862.
- Muñiz, J., y Fernández-Hermida, J.R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.
- Spielberger, C.D. (1989). *State-Trait Anxiety Inventory: A comprehensive bibliography* (2ª ed.). Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C.D., y Díaz-Guerrero, R. (Eds.) (1976). *Cross-cultural anxiety*. Washington, DC: Hemisphere/Wiley.
- Spielberger, C.D., y Díaz-Guerrero, R. (Eds.) (1983). *Cross-cultural anxiety* (Vol. 2). Washington, DC: Hemisphere/Wiley.
- Spielberger, C.D., y Díaz-Guerrero, R. (Eds.) (1986). *Cross-cultural anxiety* (Vol. 3). Washington, DC: Hemisphere/Wiley.
- Spielberger, C.D., Díaz-Guerrero, R., y Strelau, J. (Eds.) (1990). *Cross-cultural anxiety* (Vol. 4). Washington, DC: Hemisphere/Wiley.
- Spielberger, C.D., Gorsuch, R., y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C.D., Gorsuch, R.L., y Lushene, R. (1982). *Manual del Cuestionario de Ansiedad Estado/Rasgo (STAI)*. Madrid, España: TEA Ediciones.
- Suzuki, T., Tsukamoto, K., y Abe, K. (2000). Characteristic factor structures of the Japanese version of the State-Trait Anxiety Inventory: Co-existence of positive-negative and State-Trait Factor Structures. *Journal of Personality Assessment*, 74, 447-458.
- Vera-Villaruel, P., Buéla-Casal, G., y Spielberger, C.D. (2007). Preliminary analysis and normative data of the State-Trait Anxiety Inventory (STAI) in adolescent and adults of Santiago, Chile. *Terapia Psicológica*, 25, 155-162.
- Vigneau, F., y Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment*, 90, 280-285.
- Wang, P.S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M.C., Borges, G., Bromet, E.J., et al. (2007). Use of mental health services for anxiety, mood and substance disorders in 17 countries in the WHO world mental health surveys. *The Lancet*, 370, 841-850.

Versión Breve del STAI en Adolescentes y Universitarios Españoles

STAI Short Version in Adolescents and College Spanish

Alejandro Guillén-Riquelme
Gualberto Buela-Casal
Universidad de Granada, España

(Rec: 3 de julio de 2013 / Acept: 16 de agosto de 2013)

Resumen

El STAI es uno de los instrumentos de evaluación psicológica más empleados, generándose diversas versiones breves del mismo. No se han hallado versiones breves en las que se haya comprobado el ajuste para muestras de adolescentes ni universitarios españoles. El objetivo de este artículo es comprobar cuál de las versiones cortas propuestas en la literatura tiene un mejor ajuste en adolescentes y universitarios. Para ello, se aplicó el STAI a 482 adolescentes escolarizados y 510 estudiantes universitarios, de diversas ciudades españolas. Mediante AFC se evaluó el ajuste de cinco versiones breves del STAI. En los universitarios hay dos modelos en los que se observa un buen ajuste. En los adolescentes no ajusta ningún modelo. Puede que el proceso de diferenciación emocional con la edad explique que en los adolescentes las versiones breves halladas (siempre con adultos) no serían válidas, si pudiendo emplearse con universitarios.

Palabras clave: STAI, versión breve, ansiedad, adolescencia.

Abstract

The STAI is an instrument very used, generating different short versions. We found no short versions in which it is tests the fit for teenagers and college samples Spanish. The aim of this paper is to test which of the short versions proposed in the literature have a better fit in adolescents and college students. To do this, we applied the STAI to 482 adolescent students and 510 university students from different Spanish cities. Through AFC we evaluated five short versions of the STAI. At the university there are two models in which there is a good fit. In adolescents not fit any model. Maybe the emotional differentiation process explains that the found short versions (always with adults) are not valid in teenagers, and can be used with university students.

Key words: STAI, short version, anxiety, adolescence.

Introducción

El State-Trait Anxiety Inventory es uno de los instrumentos clásicos para evaluar ansiedad, uno de los trastornos psicológicos con una mayor prevalencia en población general (Beutel, Bleichner, von Heumann, Tritt y Hardt, 2011) así como en universitarios (Micin y Bagladi, 2011). Desarrollado por Spielberger, Gorsuch y Lushene (1970) este cuestionario ha sido citado en más de 14.000 estudios de archivo, siendo uno de los instrumentos más empleados por psicólogos (Spielberger y Reheiser, 2009). Su aplicación

permite obtener dos escalas: una de ansiedad rasgo (tendencia estable de percibir estímulos como más amenazantes) y ansiedad estado (presencia de estímulos ansiosos en el entorno del evaluado en los días cercanos a la evaluación). Además, el STAI ha contado con más de 60 adaptaciones o traducciones (Spielberger y Reheiser, 2009), entre ellas, la española (Buela-Casal, Guillén-Riquelme y Seisdedos, 2011), donde es el sexto cuestionario más utilizado por psicólogos clínicos (Muñiz y Fernández-Hermida, 2010). Esta adaptación cuenta con adecuadas propiedades psicométricas en población general (Guillén-Riquelme y Buela-Casal, 2011).

En 1983, se realiza una doble modificación del STAI: a) cambia la dirección de algunos ítems, para equiparar el número de elementos positivos y negativos; b) varía el contenido de otros, para facilitar su comprensión (Spielberger, 1983). Estos cambios dan respuesta a diversas críticas realizadas sobre el STAI. A la primera versión se le denomina como forma X y a la versión con los cambios descritos, como forma Y. Este hecho es importante, porque el contenido de algunos ítems varía según la versión (forma X o Y), pese a que las evaluaciones realizadas con una u otra no difieran (Spielberger, 1983). La adaptación española se realizó a partir de la forma X (Buela-Casal et al., 2011).

Diversos autores realizan una versión reducida de los ítems del cuestionario que permite evaluar ansiedad de forma fiable. La versión breve realizada por van Knippenberg, Duivenvoorden, Bonkey Passchier (1990), aplica, por primera vez, una metodología científica y unas pruebas estadísticas adecuadas; requisitos que no cumplen las propuestas anteriores. Estos autores realizan nueve versiones breves, a partir de las correlaciones ítem-total corregidas. Las versiones varían en el número de ítems (combinaciones de 4, 6 y 8 ítems). Además, utilizan dos muestras de pacientes médicos y una de estudiantes universitarios. En este trabajo, se concluye que es posible establecer versiones breves del STAI, con adecuadas propiedades psicométricas. También, se establece que las versiones de seis y ocho ítems son las adecuadas para la evaluación de la ansiedad estado y rasgo; mientras que las de cuatro ítems resultan insuficientes. En la muestra de universitarios, la versión breve está formada por los ítems número 2, 4, 8, 11, 12, 15, 17 y 18 para la ansiedad estado, y por los ítems 7, 8, 14, 15, 16, 17, 18 y 20 para la ansiedad rasgo. Hay que precisar que las versiones de 6 ítems se obtienen, eliminando, en la escala de ansiedad estado, los ítems 8 y 12, y, en la de ansiedad rasgo, los ítems 8 y 20.

Recientemente, se han propuesto nuevas versiones breves del STAI. Kaipper, Chachamovich, Hidalgo, Torres y Caumo (2010) realizaron una versión breve de ambas escalas. El banco de ítems estaba formado por todos los ítems de la forma X, más los ítems con diferente contenido de la forma Y. Estos autores establecen que los ítems de la versión breve de la escala de ansiedad estado son: 1, 2, 5, 6, 7, 8, 11, 13, 14, 16, 17, 18 y 19, además de otros ítems de la forma Y. En el caso de la ansiedad rasgo los ítems seleccionados son: 1, 2, 5, 7, 8, 9, 10, 12, 16, 17, 18 y 20, además de otros ítems de la forma Y. Esta versión breve tiene 25 ítems pero, en la actualidad, también se defienden versiones de seis ítems por cada una de las escalas. De esta forma, Fioravanti-Bastos, Cheniaux y Landeira-Fernandez (2011) afirman que la selección de los ítems 1, 3, 5, 12, 15 y 17, para la ansiedad estado y los

ítems 7, 9, 11, 13, 16 y 20 para la ansiedad rasgo, resultan adecuados para evaluar estos factores.

En España se realizó una versión breve mediante una muestra de población general, empleando análisis discriminante (Buela-Casal y Guillén-Riquelme, en prensa). Esta versión está compuesta por los ítems 3, 8, 16, 18 y 20 para la ansiedad estado y los ítems 3, 8, 10, 12 y 18 para la ansiedad rasgo. Por otra parte, se comparaba, mediante análisis factorial confirmatorio, si el ajuste de esa versión corta era mejor que el de otras reducciones anteriores, concluyendo que la versión breve con un mejor ajuste era la presentada por van Knippenberg et al. (1990). Además del trabajo de Buela-Casal y Guillén-Riquelme (en prensa), se ha realizado otra versión breve con muestra española (Perpiñá-Galvañ, Richart-Martínez y Cabañero-Martínez, 2011; Perpiñá-Galvañ, Richart-Martínez, Cabañero-Martínez y Martínez-Durá, 2011). Sin embargo, esta versión muestra un par de problemas: a) únicamente emplea muestra de pacientes con intubación respiratoria; b) se realiza la confirmación de la versión propuesta por Marteau y Bekker (1992), derivada de la forma Y del STAI, cuyos ítems sólo coinciden parcialmente con los de la adaptación española, proveniente de la forma X (Buela-Casal et al., 2011).

El STAI se caracteriza por la calidad de sus propiedades psicométricas a la hora de evaluar ansiedad en diversas poblaciones, entre ellas la de adolescentes. Así pues, la adaptación española resulta fiable y válida, presentando baremos específicos para este segmento de la población (Buela-Casal et al., 2011). Además, la adaptación chilena (desarrollada a partir de la versión española) también posee una consistencia interna elevada al ser aplicada en una muestra de adolescentes (Vera-Villarreal, Celis-Atenas, Córdova-Rubio, Buela-Casal y Spielberger, 2007).

Entre las numerosas versiones breves del STAI ninguna sobresale por un mayor apoyo empírico ni ha sido validada en población adolescente. Por ello, el objetivo del presente trabajo es comprobar si las diferentes versiones breves realizadas, a partir de la forma X del STAI, ajustan adecuadamente en: a) una muestra de adolescentes españoles y b) una muestra de universitarios. Al utilizar ambas muestras, se podrá comparar, de forma más precisa, el ajuste de la versión corta en la adolescencia y al inicio de la edad adulta. La hipótesis inicial es el adecuado ajuste de los diferentes modelos, siendo los hallados por van Knippenberg et al. (1990) y por Buela-Casal y Guillén-Riquelme (en prensa), los que ajusten de forma más precisa en adolescentes y universitarios, puesto que en ellos se ha observado mejores propiedades psicométricas en una muestra de adultos españoles.

Método

Participantes

Para la consecución del objetivo planteado, se utilizaron dos muestras independientes. La primera de ellas formada por 482 adolescentes de diversos institutos (públicos y privados) de cuatro ciudades españolas. Del total, 186 (42.2%) son varones. La media de edad es de 15.48 ($DT = 1.58$), siendo el mínimo de 12 y el máximo de 19. En el caso de los varones la media es de 15.31 años ($DT = 1.76$). En las mujeres la media es de 15.61 años ($DT = 1.59$).

La segunda muestra está formada por 510 universitarios matriculados en seis universidades españolas (todas ellas públicas), procedentes de diversas carreras. La media de edad es de 23.15 años ($DT = 5.51$) y el rango de edad abarca desde 18 a 54 años, con más de un 90% de los participantes menores de 29 años). En el caso de los varones ($n = 186$; 36.5%) la media de edad es de 25.05 años ($DT = 7.71$). En las mujeres la edad media es de 22.06 años ($DT = 3.96$).

Instrumentos

Tras realizar preguntas relativas al lugar de residencia, sexo y edad, se procedió a aplicar la adaptación española del *State-Trait Anxiety Inventory* (Buela-Casal et al., 2011). Este cuestionario permite evaluar ansiedad rasgo (tendencia a percibir los estímulos como ansiógenos o no) y estado (grado en el que hay elementos ansiosos en el entorno de la persona en el tiempo cercano a la evaluación). La escala de respuesta es de tipo Likert con 4 alternativas (en la adaptación española desde 0 hasta 3). Tanto la ansiedad estado como la ansiedad rasgo cuentan con 20 ítems para su evaluación. Por ello, el rango de respuestas va desde 0 a 60, correspondiéndose mayor nivel de ansiedad con una mayor puntuación. La adaptación española tiene adecuadas propiedades psicométricas (Guillén-Riquelme y Buela-Casal, 2011; Buela-Casal et al., 2011). Además, la escala se puede aplicar en población adolescente así como universitarios, teniendo evidencias de fiabilidad y validez.

Procedimiento

El procedimiento es diferente para cada una de las muestras obtenidas. En el caso de la muestra de adolescentes, en primer lugar se contactó con diversos centros de cuatro ciudades españolas. La aplicación de los cuadernillos transcurrió en la propia aula, durante las horas de tutorías a grupos

completos (entre 25 y 35 alumnos de forma general). El experimentador informaba de la confidencialidad del estudio así como del consentimiento informado. Seguidamente se repartieron los cuadernillos al grupo y el experimentador permaneció en el aula durante la aplicación, para resolver posibles dudas de cumplimentación.

En el caso de la muestra de universitarios, se contactó con los centros para la colaboración. El experimentador se personó en las aulas e informó sobre la confidencialidad y el consentimiento informado. De nuevo se aplicó la prueba a grupos completos, pese a que en este caso eran, de media, tres veces más numerosos que los de los adolescentes. Durante la prueba, el experimentador permaneció en el aula para contestar posibles dudas.

Análisis de datos

El análisis factorial confirmatorio fue seleccionado como la técnica analítica idónea para el ajuste de los modelos. Esta aproximación permite facilitar la comparaciones de invarianza y equivalencia para futuros estudios, garantizando la utilización del test en diversas muestras (Muñiz, Elosua y Hambleton, 2013). Se pusieron a prueba cinco modelos diferentes: el modelo 1 se corresponde al presentado por Buela-Casal y Guillén-Riquelme (en prensa); el modelo 2 es el propuesto por van Knippenberg et al. (1990) de 6 ítems; el modelo 3 se corresponde con el anterior pero en la versión de 8 ítems; el modelo 4 es el propuesto por Kaipper et al. (2010) y el modelo 5 por Fioravanti-Bastos et al. (2011). La naturaleza categórica de los ítems (escala de cuatro alternativas) hizo necesario aplicar una prueba que solventase la falta de normalidad, resultante del uso de este tipo de variables. Así pues, se seleccionó el método robusto a partir del *Weighted Least Squares* (WLSMV; Brown, 2006). La comprobación del ajuste de los modelos a los datos se realizó mediante cuatro estadísticos: el χ^2 dividido por los grados de libertad, el *Comparative Fit Index* (CFI), el *Tucker-Lewis Index* (TLI) y el *Root Mean Square Error of Approximation* (RMSEA). Los dos primeros deben tener valores superiores al 0.95 para que se considere que el ajuste es adecuado. En el RMSEA se exigen valores inferiores a 0.08. A la hora de establecer el modelo se incluyeron, de forma conjunta, ambas escalas (estado y rasgo) y su covariación, ya que estudios previos indicaban una fuerte relación entre ellas (Buela-Casal y Guillén-Riquelme, en prensa). Para los análisis se utilizó el paquete lavaan (Rosseel, 2012) del programa R. Tras ello, la redacción del artículo se realizó siguiendo las recomendaciones propuestas por Hartley (2012).

Resultados

Análisis factorial confirmatorio en adolescentes

En la muestra de adolescentes se comprobó el ajuste de los cinco modelos teóricos, expuestos anteriormente. En cada uno de ellos y, en función de los índices de modificación, se incluyeron diversas covarianzas de los errores, respetando que: 1) los errores fuesen de ítems de un mismo factor, 2) el índice de modificación tuviese un valor superior a 20 puntos. En el modelo número 4 al tener un mayor número de ítems, para evitar una sobrecarga de covariaciones de errores, se amplió este criterio hasta un valor de 30 puntos. En la Tabla 1, se recoge un resumen de los índices de ajuste de cada uno de los modelos teóricos. Ninguno de ellos alcanza niveles adecuados de ajuste, aunque, en el modelo 5 se observan valores más cercanos a un ajuste adecuado.

Análisis factorial confirmatorio en universitarios

En el caso de los universitarios, hay un correcto ajuste de diferentes modelos. El modelo 1 tiene índices próximos a los criterios establecidos. El modelo 2, conformado por 6 ítems, tiene un mejor ajuste que el resto, en la muestra de universitarios. El modelo 3 tiene un ajuste similar al

anterior, siendo ligeramente inferior, y contando con dos ítems más por cada una de las escalas. El modelo 4 es, junto con el 1, el que cuenta con índices de ajuste más bajos. El modelo 5, no llega a ajustar adecuadamente, aunque el CFI resulte superior a 0.95. Un resumen se recoge en la Tabla 2.

Tras analizar el ajuste se comprobó la covariación entre factores. En el modelo 2 (el que tiene mejores índices) la covariación es de 0.89, lo que indica una fuerte relación entre ambos factores. En el resto de modelos también se encuentra valores similares entre la covariación de la ansiedad estado y rasgo.

Análisis de fiabilidad en el modelo con mejor ajuste en universitarios

A partir de los ítems del modelo que mejor ajuste obtiene, en la muestra de universitarios, se realizó un análisis de la fiabilidad para cada una de las escalas. En el caso de la ansiedad estado, el alfa de Cronbach es de 0.82. En la correlación ítem total corregida los valores oscilan desde 0.52 a 0.73, por lo que todos los ítems están relacionados con el total. En el caso de la ansiedad rasgo el alfa de Cronbach es de 0.69 (la eliminación de un ítem mejoraría el alfa en dos centésimas). En la correlación ítem total corregida se observan resultados dispares ya que la correlación de los

Tabla 1. *Ajuste de los modelos propuestos en el análisis factorial confirmatorio en adolescentes*

Estadístico	Modelo				
	m1	m2	m3	m4	m5
χ^2	231.65	346.20	583.29	1,482.25	230.51
Grados de libertad	28	47	95	255	44
χ^2 /Grados de libertad	8.27	7.37	6.14	5.81	5.24
p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Comparative Fit Index (CFI)	0.887	0.906	0.889	0.936	0.891
Tucker-Lewis Index (TLI)	0.818	0.868	0.86	0.925	0.837
Root Mean Square Error of Approximation (RMSEA)	0.122	0.114	0.103	0.099	0.093
I.C. inf. RMSEA	0.108	0.103	0.095	0.094	0.082
I.C. sup. RMSEA	0.137	0.126	0.111	0.104	0.105

Nota. Se empleó como método de ajuste el *Weighted Least Squares*. m = modelo. El modelo 1: datos de Buela-Casal y Guillén-Riquelme (en prensa); modelo 2: propuesto por van Knippenberg et al. (1990) de 6 ítems; modelo 3: propuesto por van Knippenberg et al. (1990) de 8 ítems; modelo 4: propuesto por Kaipper et al. (2010); modelo 5: Fioravanti-Bastos et al. (2011) modificado. p = probabilidad; I.C. = intervalo de confianza al 95%.

ítems 15, 16, 17 y 18 oscila de 0.42 a 0.55, mientras en el ítem 7, es de 0.27 y en el ítem 14 es de 0.38.

Discusión

En el prolegómeno de este estudio se propuso como objetivo “comprobar si las diferentes versiones breves realizadas, a partir de la forma X del STAI, ajustan adecuadamente en: a) una muestra de adolescentes españoles y b) una muestra de universitarios”, también se partió de la hipótesis de que se obtendría un adecuado ajuste de los modelos en ambas muestras. En los universitarios la escala breve que mejor ajuste ha mostrado, coincide con la establecida por Buela-Casal y Guillén-Riquelme (en prensa) en población general española. Se confirma que la versión breve propuesta para población general (Buela-Casal y Guillén-Riquelme, en prensa), a partir de la versión de seis ítems de van Knippenberg et al. (1990), se puede utilizar para evaluar ansiedad en universitarios. El hecho de contar con una versión breve de este cuestionario tan empleado resulta necesario, no solo en los grupos analizados, si no también en pacientes con problemas físicos y que requieran varias evaluaciones (Martínez-López, Durá Ferrandis, Andreu Vaillo, Galón Garrido, Murgui Pérez y Ibáñez Guerra, 2012).

Ninguna de las versiones breves del STAI, incluida la seleccionada para los universitarios, muestra un correcto ajuste en el grupo de adolescentes. La nula incidencia de un buen ajuste en el grupo de adolescentes, de las cinco versiones breves teóricas propuestas, resulta extraña ya que la versión completa del STAI tiene adecuadas propiedades psicométricas en esta población (Buela-Casal et al., 2011). El hecho de que en el grupo de adolescentes no se ajuste ninguna de las versiones breves teóricas tiene varias explicaciones. En primer lugar, cabe indicar que la mayoría de las versiones breves se realizaron a partir de muestras de adultos (con o sin trastornos de ansiedad). Por ello, puede que ninguna de estas versiones se ajuste a una muestra de adolescentes. Sin embargo, este argumento no justifica el hecho de que la versión corta para adultos no se pueda aplicar a adolescentes, mientras que en el STAI completo se observan adecuadas propiedades psicométricas en esta muestra (Buela-Casal et al., 2011; Vera-Villaroel et al., 2007). Una posible explicación a este hecho se puede hallar en la teoría general del afecto, donde se defiende que la evolución de las emociones se produce de un modo similar a la de la inteligencia, según la hipótesis de la diferenciación. Se considera que se parte de afectos muy generales, desde los cuales, a partir de un proceso de diferenciación, se desarrollan las emociones más específicas, permitiendo diferenciarlas entre sí. Así pues, Yik, Russell y Feldman-Barrett (1999) proponen la existencia

Tabla 2. Ajuste de los modelos propuestos en el análisis factorial confirmatorio en universitarios

Estadístico	Modelo				
	m1	m2	m3	m4	m5
χ^2	160.54	143.97	358.58	1,487.54	202.18
Grados de libertad	29	44	94	259	45
χ^2 /Grados de libertad	5.54	3.27	3.81	5.74	4.49
p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Comparative Fit Index (CFI)	0.942	0.976	0.96	0.964	0.959
Tucker-Lewis Index (TLI)	0.909	0.964	0.949	0.959	0.939
Root Mean Square Error of Approximation (RMSEA)	0.094	0.067	0.074	0.096	0.083
I.C. inf RMSEA	0.08	0.055	0.066	0.092	0.071
I.C. sup RMSEA	0.109	0.079	0.083	0.101	0.095

Nota. Se empleó como método de ajuste el *Weighted Least Squares*. m = modelo. El modelo 1: datos de este artículo; modelo 2: propuesto por van Knippenberg et al. (1990) de 6 ítems; modelo 3: propuesto por van Knippenberg et al. (1990) de 8 ítems; modelo 4: propuesto por Kaipper et al. (2010); modelo 5: Fioravanti-Bastos et al. (2011) modificado. p = probabilidad; I.C. = intervalo de confianza al 95%.

de dos continuos en los que se enmarcarían las diferentes emociones: el primero de ellos el *arousal*, oscilante desde la activación a la desactivación; el segundo el de la valencia del afecto, indicativo del carácter positivo o negativo de los sentimientos. La ansiedad se vincularía exclusivamente con afecto negativo (Watson, Clark y Stasik, 2011). De acuerdo con este modelo, las emociones son más generales en la adolescencia que en el inicio de la edad adulta (tramo en el que se encuentra la muestra de universitarios), ya que la diferenciación se incrementa con el desarrollo psicológico del individuo. Por tanto, en adolescentes de muy tempranas edades puede ser recomendable emplear versiones breves de cuestionarios más generales sobre sensibilidad ansiosa (por ejemplo: Kämpfe et al., 2012).

En los universitarios el modelo que mejor ajusta es el presentado por van Knippenberg et al. (1990) lo que confirma la hipótesis inicial, donde se enunciaba que dicha versión tendría el mejor ajuste, puesto que esta propuesta fue la que se seleccionó como versión breve en población general española (Buela-Casal y Guillén-Riquelme, en prensa). Este hecho, confirma que se puede emplear una versión breve en universitarios, pese a que tengan mayores puntuaciones medias en ansiedad que la población general (Buela-Casal et al., 2011), pueden ser evaluados empleando una versión reducida de la misma. Además, al considerar el grupo de universitarios como una parte de la población general, se aporta una nueva evidencia de que la versión breve seleccionada en la investigación de Buela-Casal y Guillén-Riquelme (en prensa), es la más apropiada. Observando otras versiones breves de cuestionarios de ansiedad que ajustan bien en este tipo de muestras (por ejemplo: Antúnez y Vinet, 2012). Así pues, la versión de seis ítems de van Knippenberg et al. (1990) permite evaluar ansiedad estado y rasgo en población general (incluyendo al grupo de universitarios).

En resumen, la versión propuesta por van Knippenberg et al. (1990) y confirmada por Buela-Casal y Guillén-Riquelme (en prensa), compuesta por los ítems 2, 4, 11, 15, 17 y 18, para la ansiedad estado y por los ítems 7, 14, 15, 16, 17 y 18 para la ansiedad rasgo, permite evaluar de forma fiable ambos factores en universitarios. En el caso de los adolescentes, pese a que la versión completa sea válida, las versiones breves utilizadas hasta la fecha no proporcionan un correcto ajuste. Es imprescindible hallar una versión breve que permita evaluar ansiedad en adolescentes de forma válida y fiable. Una de las limitaciones del presente estudio es que las muestras no están seleccionadas de forma completamente aleatoria. Por ello, cabría plantearse en futuros estudios el usar este tipo de selección muestral, así como comprobar la validez de contenido del cuestionario

(Delgado-Rico, Carretero-Dios y Ruch, 2012). A pesar de ello, al establecer que la versión breve puede ser empleada para el análisis de la ansiedad rasgo y estado en universitarios, se aporta una nueva evidencia de su correcto funcionamiento en población general. Este hallazgo puede tener grandes aplicaciones favoreciendo la investigación y la evaluación de este colectivo, con un menor número de ítems, pero una fiabilidad similar a la de la versión completa.

Referencias

- Antúnez, Z., y Vinet, E. V. (2012). Escalas de depresión, ansiedad y Estrés (DASS-21): Validación de la Versión abreviada en Estudiantes Universitarios Chilenos. *Terapia Psicológica*, 30, 49-55.
- Beutel, M. E., Bleichner, F., von Heymann, F., Tritt, K., y Hardt, J. (2011). Inpatient psychosomatic treatment of anxiety disorders: Comorbidities, predictors, and outcomes. *International Journal of Clinical and Health Psychology*, 11, 443-457.
- Brown, T. A. (2006). Data Issues in CFA: Missing, Non-Normal, and Categorical Data. En Kenny, D.A. (Ed.), *Confirmatory Factor Analysis for Applied Research* (pp. 363-411). New York: The Guilford Press.
- Buela-Casal, G., y Guillén-Riquelme, A. (En prensa). Short version of Spanish adaptation of the State-Trait Anxiety Inventory.
- Buela-Casal, G., Guillén-Riquelme, A., y Seisdedos, N. (2011). STAI: Cuestionario de Ansiedad Estado-Rasgo. Adaptación española (8ª ed.). Madrid: TEA Ediciones.
- Delgado-Rico, E., Carretero-Dios, H., y Ruch, W. (2012). Content validity evidences in test development: An applied perspective. *International Journal of Clinical and Health Psychology*, 12, 449-460.
- Fioravanti-Bastos, A. C. M., Cheniaux, E., y Landeira-Fernandez, J. (2011). Development and Validation of a Short-Form Version of the Brazilian State-Trait Anxiety Inventory. *Psicologia: Reflexão e Crítica*, 24, 485-494.
- Guillén-Riquelme, A., y Buela-Casal, A. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23, 510-515.
- Hartley, J. (2012). New ways of making academic articles easier to read. *International Journal of Clinical and Health Psychology*, 12, 143-160.
- Kaapper, M. B., Chachamovich, E., Hidalgo, M. P. L., Torres, I. L. S., y Caumo, W. (2010). Evaluation of the structures of Brazilian State-Trait Anxiety Inventory using a Rasch psychometric approach. *Journal of Psychosomatic Research*, 68, 223-233.
- Kämpfe, C. K., Gloster, A. T., Wittchen, H. U., Helbig-Lang, S., Lang, T., ... Gerlach, A. L. (2012). Experiential avoidance and anxiety sensitivity in patients with panic disorder and agoraphobia: Do both constructs measure the same? *International Journal of Clinical and Health Psychology*, 12, 5-22.
- Martínez-López, P., Durá Ferrandis, E., Andreu Vaillo, Y., Galón Garrido, M. J., Murgui Pérez, S., y Ibáñez Guerra, E. (2012). Structural validity and distress screening potential of the Hospital Anxiety and Depression Scale in cancer. *International Journal of Clinical and Health Psychology*, 12, 435-447.
- Marteau, T. M., y Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31, 301-306.
- Micín, S., y Bagladi, V. (2011). Salud Mental en Estudiantes Universitarios: Incidencia de Psicopatología y Antecedentes de Conducta Suicida en Población que Acude a un Servicio de Salud Estudiantil. *Terapia Psicológica*, 29, 53-64.
- Muñiz, J., Elosua, P., y Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25, 151-157.

- Muñiz, J., y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31, 108-121.
- Perpiñá-Galvañ, J., Richart-Martínez, M., y Cabañero-Martínez, M. J. (2011). Fiabilidad y validez de una versión corta de la escala de medida de la ansiedad STAI en pacientes respiratorios. *Archivos de Bronconeumología*, 47, 184-189.
- Perpiñá-Galvañ, J., Richart-Martínez, M., Cabañero-Martínez, M. J., y Martínez-Durá, I. (2011). Validez de contenido de versión corta de la subescala del Cuestionario State-Trait Anxiety Inventory (STAI). *Revista Latino-Americana de Enfermagem*, 19, A04:1-A04:6.
- RosseeL, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48, 1-36.
- Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory: STAI(Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Gorsuch, R. L., y Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, California: Consulting Psychologist Press.
- Spielberger, C. D., y Reheier, E. C. (2009). Assessment of Emotions: Anxiety, Anger, Depression, and Curiosity. *Applied Psychology: Health and Well-Being*, 1, 271-302. doi:10.1111/j.1758-0854.2009.01017.x.
- van Knippenberg, F. C. E., Duivenvoorden, H. J., Bonke, B., y Passchier, J. (1990). Shortening the State-Trait Anxiety Inventory. *Journal of Clinical Epidemiology*, 43, 995-1000.
- Vera-Villaruel, P., Celis-Atenas, K., Córdova-Rubio, N., Buela-Casal, G., y Spielberger, C. D. (2007). Preliminary Analysis and Normative Data of the State-Trait Anxiety Inventory (STAI) in Adolescent and Adults of Santiago, Chile. *Terapia Psicológica*, 25, 155-162.
- Watson, D., Clark, L. A., y Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, 11, 429-442.
- Yik, M., Russell, J., y Feldman-Barrett, L. (1999). Structure of self-reported current affect: integration and beyond. *Journal of Personality and Social Psychology*, 77, 600-619.

