



## Recomendaciones para la utilización de la adaptación española del Inventario de Ansiedad de Beck (BAI) en la práctica clínica

Jesús Sanz\*

Universidad Complutense de Madrid, España

### INFORMACIÓN ARTÍCULO

Manuscrito recibido: 24/06/2013  
Revisión recibida: 24/10/2013  
Aceptado: 28/10/2013

*Palabras clave:*

Ansiedad  
Inventario de Ansiedad de Beck  
Baremos  
Significación clínica  
Cribado  
Diagnóstico

*Keywords:*

Anxiety  
Beck Anxiety Inventory  
Norms  
Clinical significance  
Screening  
Diagnosis

### RESUMEN

En sus 25 años de existencia, el Inventario de Ansiedad de Beck (BAI) se ha convertido en el cuestionario autoaplicado para evaluar la gravedad de la ansiedad más utilizado en los países con mayor producción científica en psicología, aunque en España dicho instrumento sigue siendo el Cuestionario de Ansiedad Estado-Rasgo (STAI). En este trabajo se presentan las características básicas del BAI en comparación con la escala de Ansiedad Rasgo del STAI, se describe el proceso de adaptación del BAI a la población española, se resumen las propiedades psicométricas de dicha adaptación y se discute su utilización en la práctica clínica, para lo cual se calculan baremos y puntuaciones de corte para la evaluación de la gravedad de la ansiedad, la evaluación de la significación clínica de los cambios terapéuticos, el cribado de personas con ansiedad clínica y la ayuda en el diagnóstico diferencial de los trastornos de ansiedad.

© 2014 Colegio Oficial de Psicólogos de Madrid. Todos los derechos reservados.

### Recommendations for the use of the Spanish adaptation of the Beck Anxiety Inventory (BAI) in clinical practice

#### ABSTRACT

In its 25 years of existence, the Beck Anxiety Inventory (BAI) has become the most widely used self-report instrument for assessing the severity of anxiety in the countries with higher research productivity in psychology, although in Spain this instrument is still the State-Trait Anxiety Inventory (STAI). This paper presents the basic characteristics of the BAI in comparison with the Trait Anxiety scale of the STAI, describes the process of adaptation of the BAI to the Spanish population, summarizes the psychometric properties of this adaptation, and discusses its use in the clinical practice. To this use, norms and cut-off scores are calculated in order to measure the severity of anxiety, evaluate the clinical significance of therapeutic change, screen for clinical anxiety, and aid in the differential diagnosis of anxiety disorders.

© 2014 Colegio Oficial de Psicólogos de Madrid. All rights reserved.

Este año se cumple el 26° aniversario de la publicación del Inventario de Ansiedad de Beck, internacionalmente conocido por el acrónimo de su nombre original en inglés BAI (Beck Anxiety Inventory; Beck, Epstein, Brown y Steer, 1988). En los 25 años transcurridos desde su primera publicación el BAI se ha convertido en el cuestionario autoaplicado más utilizado para evaluar la gravedad de la ansiedad en los países con mayor producción científica en psicología. Así, un

estudio publicado en 2009 y realizado a partir de la base de datos bibliográficos PsycINFO, identificaba al BAI como el instrumento de evaluación de la ansiedad más utilizado en la investigación psicológica (Piotrowski y Gallant, 2009), avanzando desde la tercera posición que le situaba un estudio similar publicado diez años antes (Piotrowski, 1999) y por encima del Cuestionario de Ansiedad Estado-Rasgo (STAI; Spielberger, Gorsuch y Lushene, 1970) que era el instrumento que entonces ocupaba la primera posición.

En España, el BAI también ha ido ganando en popularidad como instrumento para evaluar la ansiedad tanto en investigación básica y aplicada como en la práctica clínica. De hecho, una búsqueda en PsycINFO realizada el 12 de junio de 2013 localizó 172 estudios espa-

\*La correspondencia sobre este artículo debe enviarse al autor a la Facultad de Psicología. Universidad Complutense de Madrid. 28223 Pozuelo de Alarcón. E-mail: [jsanz@psi.ucm.es](mailto:jsanz@psi.ucm.es)

ños que habían utilizado el BAI para medir la ansiedad. Sin embargo, el STAI sigue siendo todavía el instrumento más utilizado por los psicólogos en España. Así, según un estudio publicado en 2010 y realizado con una muestra de 3126 miembros de los colegios oficiales de psicólogos de España, el STAI era el séptimo test más utilizado por los psicólogos españoles en su práctica diaria y el primero entre los dedicados específicamente a la evaluación de la ansiedad (Muñiz y Fernández-Hermida, 2010), prácticamente las mismas posiciones que había obtenido el instrumento diez años atrás en un estudio similar (sexto test más utilizado y el primero entre los específicos de ansiedad; Muñiz y Fernández-Hermida, 2000). Dado que hasta 2011 no se publicó la primera adaptación española del Inventario de Ansiedad de Beck (Beck y Steer, 2011), cabría aventurar que en los próximos años, al igual que ha ocurrido en otros países (Piotrowski y Gallant, 2009), el BAI podría ir ganando posiciones en las preferencias de los psicólogos españoles hasta alcanzar al STAI.

En este trabajo se presentan las características básicas del BAI, especialmente en comparación con el STAI, se resume su proceso de adaptación a la población española y las propiedades psicométricas de esta adaptación y se discute y se ofrecen baremos y puntuaciones de corte para su utilización en la práctica clínica en la evaluación de la gravedad de la ansiedad, en la evaluación de la significación clínica de los cambios terapéuticos y en el cribado y diagnóstico de los trastornos de ansiedad.

## El BAI frente al STAI-R

Dada la popularidad del STAI entre los psicólogos españoles, cabría plantearse, en primer lugar, cuáles son las diferencias y semejanzas entre el BAI y el STAI, en particular, la subescala de Ansiedad Rasgo del STAI (STAI-R) y, en segundo lugar, si el BAI podría ser un sustituto o un complemento del STAI-R en la evaluación de la ansiedad en el área de la psicología clínica y de la salud. En la tabla 1 se han recogido las principales diferencias y semejanzas entre ambos instrumentos, especialmente en relación con su utilización en psi-

cológia clínica y de la salud. Entre las semejanzas, algunas de las más evidentes son la brevedad de ambos instrumentos (20–21 ítems) y la obtención de índices semejantes y buenos o excelentes ( $\geq .80$  o  $\geq .85$ ; Prieto y Muñiz, 2000) de fiabilidad de consistencia interna. Entre las diferencias, las más notables tienen que ver con: (a) los constructos que pretenden medir, (b) el contenido de sus ítems y los síntomas de ansiedad y depresión que evalúan dichos ítems y (c) sus capacidades para discriminar entre ansiedad y depresión. En concreto, el STAI-R fue construido para medir la ansiedad rasgo definida como “una relativamente estable propensión ansiosa por la que difieren los sujetos en su tendencia a percibir las situaciones como amenazadoras y a elevar, consecuentemente, su ansiedad estado” (Spielberger, Gorsuch y Lushene, 1982, p. 7), mientras que el BAI “se construyó para medir los síntomas de la ansiedad que apenas son compartidos con los de la depresión” (Beck y Steer, 2011, p. 10), es decir, mientras que el STAI-R fue desarrollado para evaluar un rasgo normal de personalidad, el BAI lo fue para evaluar conductas (cognitivas, emocionales, motoras o fisiológicas) anormales o patológicas de ansiedad (ansiedad clínica). Esta diferencia también se ve reflejada, de alguna manera, en sus instrucciones y en el formato de respuesta de sus ítems, puesto que el STAI-R pregunta sobre la frecuencia (casi nunca, a veces, a menudo o casi siempre) de ciertas conductas en la mayoría de las ocasiones, mientras que el BAI pregunta por el grado de molestia (nada, leve, moderado o grave) que le han ocasionado al individuo ciertas conductas anormales durante la última semana. No obstante, en la medida en que la ansiedad estado, conceptualizada como “un estado o condición emocional transitoria del organismo humano, que se caracteriza por sentimientos subjetivos, conscientemente percibidos, de tensión y aprensión, así como por una hiperactividad del sistema nervioso autónomo” (Spielberger et al., 1982, p. 7) puede alcanzar niveles de intensidad, frecuencia y disfuncionalidad anormales o patológicos (ansiedad clínica), el STAI-R podría detectar personas que presentan de forma estable tales niveles y, bajo este supuesto, ha sido utilizado en psicología clínica y de la salud.

**Tabla 1**  
Principales semejanzas y diferencias entre el BAI y el STAI-R

Características	STAI-R (Forma X)	BAI
Año de publicación	1970	1988
Constructo a medir	Ansiedad rasgo	Sintomatología ansiosa
Nº de ítems	20	21
Formato de respuestas de los ítems	Escala Likert de 4 puntos sobre frecuencia	Escala Likert de 4 puntos sobre grado de molestia
Marco temporal de las instrucciones	En general	Última semana
Adaptación española	Seisdedos (Spielberger et al., 1982)	Sanz (Beck y Steer, 2011)
Fiabilidad de consistencia interna en muestras españolas <sup>a</sup>	Adultos = .84/.86/.90 Universitarios = .88 Pacientes médicos = .84	Adultos = .92 Universitarios = .88 Pacientes con trastornos psicológicos = .91
% cubierto de síntomas de TAG según DSM-IV	62.5%	25%
% de ítems que evalúan síntomas de TAG según DSM-IV	70%	9.5%
% cubierto de síntomas de crisis de angustia según DSM-IV	7.1%	78.6%
% de ítems que evalúan síntomas de crisis de angustia según DSM-IV	30%	100%
% cubierto de síntomas de depresión mayor según DSM-IV	66.7%	0%
% de ítems que evalúan síntomas de depresión mayor según DSM-IV	65%	0%
Correlación con medidas de depresión (BDI-II) <sup>b</sup>	Varios países = .66	Varios países = .53 España = .58

Nota. TAG: trastorno de ansiedad generalizada. BDI-II: Inventario de Depresión de Beck-II.

<sup>a</sup>Los datos de fiabilidad de consistencia interna del STAI-R en muestras de adultos han sido tomados de Spielberger et al. (1982) y Guillén-Riquelme y Buela-Casal (2011), en muestras de paciente médicos de Spielberger et al. (1982) y en muestras de estudiantes universitarios de Fonseca-Pedrero, Paino, Sierra-Baigrie, Lemos-Giráldez y Muñiz (2012). Los datos de fiabilidad de consistencia interna del BAI son los recogidos en el manual de la adaptación española (Beck y Steer, 2011).

<sup>b</sup>Todos los resultados obtenidos en muestras de estudiantes universitarios (Beck y Steer, 2011). La correlación del STAI-R con el BDI-II es la correlación media obtenida a partir de 9 estudios con 3955 participantes, ponderada por el número de participantes de cada estudio. La correlación del BAI con el BDI-II en varios países es la correlación media obtenida a partir de 10 estudios con 3099 participantes, ponderada por el número de participantes de cada estudio. La correlación del BAI con el BDI-II en España es la correlación obtenida con la muestra de 727 estudiantes universitarios de la adaptación española del BAI.

Una segunda diferencia importante entre el BAI y el STAI-R reside en el contenido de sus ítems y, en el contexto de su utilización en psicología clínica y de la salud, en el grado en que tales ítems reflejan conductas (cognitivas, emocionales, motoras o fisiológicas) anormales o patológicas de ansiedad (ansiedad clínica). En concreto, un análisis del contenido de sus ítems y de su relevancia para evaluar la ansiedad clínica, en particular, el trastorno de ansiedad generalizada (TAG) y la crisis de angustia o ataque de pánico, muestra que existen importantes diferencias entre el BAI y el STAI-R. En la tabla 2 se presenta la correspondencia entre el contenido de los ítems del BAI y del STAI-R y los síntomas de los criterios diagnósticos del trastorno de ansiedad generalizada según el “Manual diagnóstico y estadístico de los trastornos mentales” (DSM-IV) de la Asociación Americana de Psiquiatría (APA, 1994/1995). Como puede observarse en dicha tabla, el STAI-R es más relevante que el BAI para el contenido de la definición sintomática del TAG del DSM-IV, ya que sus ítems cubren un mayor porcentaje de síntomas del TAG (62.5% frente a 25%) y tiene un mayor porcentaje de ítems que evalúan síntomas del TAG (70% frente a 9.5%). Por el contrario, en la tabla 3 se presenta la correspondencia entre el contenido de los ítems del BAI y del STAI-R y los síntomas de los criterios diagnósticos de la crisis de angustia según el DSM-IV (APA, 1994/1995), pudiéndose inferir de dicha tabla que el BAI es más relevante que el STAI-R para el contenido de la definición sintomática de la crisis de angustia del DSM-IV, ya que sus ítems cubren un mayor porcentaje de síntomas de la crisis de angustia (78.6% frente a 7.1%) y tiene un mayor porcentaje de ítems que evalúan síntomas del TAG (100% frente a 30%).

Por tanto, de los datos de las tablas 2 y 3 (véase también la tabla 1), se podría concluir que, desde la perspectiva de la validez de contenido (Sanz, Izquierdo y García-Vera, en prensa), el STAI-R evalúa de forma razonable el TAG (más del 50% de síntomas del TAG cubiertos) y sería más adecuado para evaluar el TAG que el BAI, mientras que el BAI evalúa de forma razonable la crisis de angustia (más del 50% de síntomas de la crisis de angustia cubiertos) y sería más adecuado para evaluar la crisis de angustia que el STAI-R. Esta última característica del BAI es muy importante para el objetivo general de evaluar la ansiedad clínica, ya que las crisis de angustia, bien sean completas o con sintomatología limitada (o crisis parciales), pueden aparecer en el contexto de la mayoría de los trastornos de ansiedad (p. ej., fobia específica, fobia social, trastorno obsesivo compulsivo, trastor-

no de angustia, agorafobia, trastorno por estrés postraumático, trastorno por estrés agudo).

Por otro lado, desde la perspectiva de la validez de contenido, un instrumento tiene una alta relevancia para evaluar la ansiedad clínica no sólo si sus ítems cubren todos los síntomas que definen los síndromes y trastornos de ansiedad sino, además, si sus ítems no miden síntomas que no formen parte del dominio de la ansiedad clínica, por ejemplo si sus ítems no reflejan síntomas del dominio de la depresión clínica (Sanz et al., en prensa). En la tabla 4 se presenta la correspondencia entre el contenido de los ítems del STAI-R y del BAI y los síntomas de los criterios diagnósticos del episodio depresivo mayor del DSM-IV (APA, 1994/1995). Los datos de la tabla 4 indican que el 65% de los ítems del STAI-R miden también síntomas del episodio depresivo mayor y que tales ítems cubren el 66.7% de los criterios sintomáticos del episodio depresivo mayor del DSM-IV, mientras que el BAI no tiene ítems que, según su contenido, midan los criterios sintomáticos de dicho trastorno. Esta excesiva cobertura de los síntomas depresivos que tiene el STAI-R le plantea un problema de validez discriminante, ya que cabría estimar que una parte muy importante de sus puntuaciones estaría determinada no por la presencia y gravedad de síntomas ansiosos sino por la presencia y gravedad de síntomas depresivos. Por el contrario, la ausencia de contenido depresivo en el BAI sería esperable dado que dicho instrumento “fue desarrollado para abordar la necesidad de un instrumento que discrimine fiablemente la ansiedad de la depresión a la vez que muestra validez convergente” (Beck et al., 1988, p. 893).

Es más, la correlación que guarda el BAI con una medida de depresión tan consolidada y validada como es el Inventario de Depresión de Beck-II o BDI-II (Beck, Steer y Brown, 1996) es significativamente menor que la que presenta el STAI-R, al menos en muestras de estudiantes universitarios de diversos países (véase la tabla 1). Efectivamente, al menos 10 estudios han examinado la relación entre el BAI y el BDI-II en estudiantes universitarios (véase Beck y Steer, 2011) y la correlación media de tales estudios, ponderada por el número de participantes de cada uno de ellos, fue .53, significativamente menor que la correlación media ponderada de .66 encontrada en nueve estudios (véase Beck y Steer, 2011) que habían analizado la relación entre el STAI-R y el BDI-II en estudiantes universitarios ( $z = 8.44, p < .001$ ). Además, la correlación entre las adaptaciones españolas del BAI y del BDI-II que se ha encontrado en muestras de estudiantes

**Tabla 2**

Comparación entre los criterios sintomáticos del DSM-IV para el trastorno de ansiedad generalizada y los ítems del STAI-R y del BAI (número de los ítems entre paréntesis)

Criterio	STAI-R (Forma X)	BAI
1. Ansiedad y preocupación excesivas (expectación aprensiva)	(21) Sentirse bien (24) Querer ser feliz (27) Persona tranquila, serena y sosegada (28) Pensar que las dificultades se amontonan y no se puede con ellas (29) Preocupación excesiva por cosas sin importancia (30) Ser feliz (31) Tomarse las cosas demasiado seriamente (33) Sentirse seguro (37) Rondarle y molestarle pensamientos sin importancia (39) Persona estable (40) Ponerse tenso y agitado al pensar sobre asuntos y preocupaciones actuales	(10) Nerviosismo (4) Incapacidad para relajarse
2. Dificultad para controlar la preocupación	(38) No poder olvidar los desengaños porque afectan mucho	(4) Incapacidad para relajarse
3. Inquietud	(27) Persona tranquila, serena y sosegada (39) Persona estable (40) Ponerse tenso y agitado al pensar sobre asuntos y preocupaciones actuales	
4. Fatigabilidad	(22) Cansarse rápidamente (26) Sentirse descansado	
5. Dificultad para concentrarse		
6. Insomnio o sueño no reparador		
7. Tensión muscular		
8. Irritabilidad	(27) Persona tranquila, serena y sosegada (39) Persona estable	

**Tabla 3**  
Comparación entre los criterios sintomáticos del DSM-IV para la crisis de angustia (ataque de pánico) y los ítems del STAI-R y del BAI (número de los ítems entre paréntesis)

Criterio	STAI-R (Forma X)	BAI
1. Miedo o malestar intensos	(21) Sentirse bien (24) Desear ser feliz (27) Persona tranquila, serena y sosegada (30) Ser feliz (33) Sentirse seguro (39) Persona estable	(4) Incapacidad para relajarse (9) Estar aterrizado (10) Nerviosismo (17) Estar asustado
2. Palpitaciones		(7) Palpitaciones
3. Sudoración		(21) Sudoración
4. Temblores o sacudidas		(12) Temblor de manos (13) Temblor generalizado
5. Sensación de ahogo o falta de aliento		(11) Sensación de ahogo (15) Dificultad para respirar
6. Sensación de atragantarse		
7. Opresión o malestar torácico		
8. Náuseas o molestias abdominales		(18) Molestias abdominales
9. Inestabilidad, mareo o desmayo		(3) Debilidad en piernas (6) Mareos (8) Inestabilidad (19) Desmayo
10. Desrealización o despersonalización		
11. Miedo a perder el control o volverse loco		(5) Miedo a lo peor (14) Miedo a perder el control
12. Miedo a morir		(16) Miedo a morir
13. Parestesias		(1) Hormigueo
14. Escalofríos o sofocaciones		(2) Sensación de calor (20) Rubor facial

**Tabla 4**  
Comparación entre los criterios sintomáticos del DSM-IV para el episodio depresivo mayor y los ítems del STAI-R y del BAI (número de los ítems entre paréntesis)

Criterio	STAI-R (Forma X)	BAI
1. Estado de ánimo deprimido	(21) Sentirse bien (23) Sentir ganas de llorar (24) Desear ser feliz (30) Ser feliz (35) Sentirse triste (melancólico)	
2. Disminución del placer o interés	(36) Estar satisfecho	
3. Aumento/descenso del peso/apetito		
4. Insomnio o hipersomnia		
5. Agitación o entecimiento psicomotor	(27) Persona tranquila, serena y sosegada (39) Persona estable (40) Ponerse tenso y agitado al pensar sobre asuntos y preocupaciones actuales	
6. Fatiga o pérdida de energía	(22) Cansarse rápidamente (26) Sentirse descansado	
7. Sentimientos de inutilidad o culpa	(32) Falta de confianza en sí mismo	
8. Disminución de la capacidad de pensar, concentrarse o tomar decisiones	(25) Perder oportunidades por no decidirse pronto	
9. Ideas de muerte o suicidio		

Nota. Los criterios 4, 5, 6 y 8 del DSM-IV para el episodio depresivo mayor son compartidos parcialmente con los criterios sintomáticos del DSM-IV para el trastorno de ansiedad generalizada (véase la tabla 2).

universitarios (véase la tabla 1) es también significativamente menor que la correlación media ponderada que presentaba el STAI-R con el BDI-II (.58 frente a .66,  $z = 3.22$ ,  $p < .002$ ).

En resumen, el BAI, en concordancia con sus objetivos originales, discrimina mejor entre ansiedad y depresión que el STAI-R, tanto en términos de su validez de contenido como en términos de su validez discriminante respecto al BDI-II. De hecho, los problemas del STAI-R para discriminar ansiedad y depresión han sido señalados repetida-

mente en la literatura científica, de manera que varios estudios empíricos en muestras clínicas y no clínicas, incluyendo muestras españolas, han demostrado que el STAI-R mide tanto ansiedad como depresión (Bados, Gómez-Benito y Balaguer, 2010; Bieling, Antony y Swinson, 1998; Caci, Baylé, Dossios, Robert y Boyer, 2003), incluso cuando se utiliza la Forma Y del STAI-R en la cual se han remplazado 6 ítems de la Forma X para obtener una medida supuestamente más pura de ansiedad frente a la depresión (Spielberger, Gorsuch, Lush-

ne, Vagg y Jacobs, 1983). Estos problemas no deberían sorprender dado que, por ejemplo, el STAI-R incluye siete ítems redactados de forma negativa, es decir, que supuestamente reflejan la ausencia de ansiedad como, por ejemplo, sentirse bien, descansado, tranquilo, sereno, sosegado, seguro, satisfecho o feliz, que, sin embargo, pueden reflejar también la ausencia de otros estados emocionales negativos como, por ejemplo, depresión o ira. De hecho, el carácter inespecífico de tales ítems podría cuestionar la correspondencia con los criterios sintomáticos del TAG y de la crisis de angustia que se recoge en las tablas 2 y 3. Es más, algunos ítems del STAI-R, especialmente en su adaptación española, evalúan, sin duda, síntomas específicos de la depresión más que de la ansiedad (p. ej., los ítems 23 y 35 que rezan, respectivamente, “siento ganas de llorar” y “me siento triste [melancólico]”).

### La adaptación española del BAI

Dadas las características positivas del BAI como instrumento de evaluación de la ansiedad clínica, un grupo de investigación de la Universidad Complutense de Madrid inició en el año 2000 diversos estudios para su validación en España (Magán, Sanz y García-Vera, 2008; Sanz, García-Vera y Fortún, 2012; Sanz y Navarro, 2003), estudios que culminaron con la adaptación española del instrumento realizada en colaboración con *Pearson Clinical & Talent Assessment España* y publicada por esta editorial (Beck y Steer, 2011). La adaptación española del BAI está basada en un análisis exhaustivo de sus propiedades psicométricas de fiabilidad de consistencia interna, validez convergente y discriminante, validez factorial y validez de criterio en dos muestras de pacientes con trastornos psicológicos que totalizaban 392 pacientes, dos muestras de adultos de la población general que totalizaban 348 adultos y dos muestras de estudiantes universitarios que totalizaban 727 estudiantes. Estos análisis sugieren, consistentemente en las diversas muestras, que la adaptación española del BAI muestra buenos índices psicométricos para medir la presencia y gravedad de sintomatología ansiosa, índices que además son similares a los de la versión original o a los de las adaptaciones realizadas en otros países (Beck y Steer, 2011). Para los objetivos del presente trabajo, cabe detallar algunos de esos índices.

Los coeficientes  $\alpha$  de Cronbach de fiabilidad de consistencia interna para los pacientes con trastornos psicológicos, los adultos de la población general y los estudiantes universitarios fueron, respectivamente, .90, .91 y .88, todos ellos excelentes ( $\geq .85$ ; Prieto y Muñiz, 2000) y comparables a los obtenidos en estudios con muestras semejantes de otros países (véase la tabla 5).

En relación con la validez diagnóstica, un análisis de la curva de las características operativas del receptor o curva ROC reveló que el

BAI mostraba un rendimiento diagnóstico (área bajo la curva ROC = .80) adecuado ( $> 0.70$ ; Swets, 1988) para discriminar entre estudiantes universitarios con trastorno de ansiedad y estudiantes sin trastorno de ansiedad evaluados mediante una entrevista diagnóstica estructurada. Este análisis de rendimiento diagnóstico permitió identificar varias puntuaciones de corte del BAI con índices adecuados de validez diagnóstica para detectar personas con trastorno de ansiedad (p. ej., sensibilidad y especificidad  $> 70\%$ ; valores predictivos positivo y negativo  $> 50\%$ ; kappa  $> .40$ ; véase la tabla 6).

**Tabla 6**

Índices de rendimiento diagnóstico de las puntuaciones del BAI para detectar personas con trastornos de ansiedad en una muestra española de estudiantes universitarios (adaptado de Beck y Steer, 2011)

Puntuación del BAI	Sensibilidad	Especificidad	Valor predictivo positivo	Valor predictivo negativo	Kappa
12	81.1	79.8	36.9	96.7	.40
14	76.4	88.2	48.5	96.2	.52
19	42.9	96.5	64.7	92.0	.46

Respecto a la validez de criterio para diferenciar grupos contrastados, la puntuación total media del BAI de los pacientes con trastornos psicológicos fue significativamente mayor que las de los adultos de la población general y los estudiantes universitarios (Beck y Steer, 2011). De hecho, la gravedad global de la ansiedad en los pacientes fue casi el doble de la encontrada en las muestras no clínicas de la población general y de estudiantes universitarios (media = 18.9 frente a 10.1 y 9.4, respectivamente; véanse las tablas 7 y 8), lo que suponía, en términos del tamaño del efecto, unas diferencias estandarizadas de medias ( $d$  de Cohen) iguales a 0.81 y 1.01, respectivamente, que pueden considerarse grandes ( $d > 0.80$ ; Cohen, 1988).

### Recomendaciones para utilizar la adaptación española del BAI en la práctica clínica

La interpretación de las puntuaciones del BAI se suele basar en unas puntuaciones de corte que definen diferentes niveles de gravedad de sintomatología ansiosa. La última edición del manual original del BAI propone las siguientes: 0–7 indica ansiedad mínima, 8–15 ansiedad leve, 16–25 ansiedad moderada y 26–63 ansiedad grave (Beck y Steer, 1993). Estas puntuaciones fueron asumidas en la adaptación española del BAI, aunque sería deseable que futuros estudios confirmaran empíricamente su validez para distinguir niveles de gravedad de la ansiedad en pacientes españoles con trastornos psicológicos.

### Baremos de la adaptación española del BAI

Complementariamente, la interpretación de las puntuaciones en el BAI puede hacerse comparando dichas puntuaciones con los baremos de un grupo de referencia. El manual de la versión original del BAI presenta las medias y desviaciones típicas obtenidas en dos muestras incidentales de 160 y 393 pacientes estadounidenses con trastornos psicológicos y en tres muestras incidentales no clínicas, dos de 65 y 142 estudiantes universitarios ingleses y una de 36 adultos no estudiantes (Beck y Steer, 1993). En la misma línea, en el manual de la adaptación española del BAI se presentan las medias y desviaciones típicas obtenidas en tres muestras incidentales españolas: una de 392 pacientes externos con trastornos psicológicos, otra de 348 adultos de la población general y una tercera de 727 estudiantes universitarios (Beck y Steer, 2011). Sin embargo, y al contrario de lo que ocurre en algunos otros tests psicopatológicos, ni el manual original del BAI ni el manual de su adaptación española presentan tablas de puntuaciones tipificadas, por ejemplo, tablas de puntuaciones  $T$  o de puntuaciones centiles. Estas tablas tienen varias ventajas.

**Tabla 5**

Fiabilidad de consistencia interna ( $\alpha$  de Cronbach) del BAI en diferentes muestras y con diferentes adaptaciones (adaptado de Beck y Steer, 2011)

Población/adaptación del BAI	N	$\alpha$
Pacientes externos con trastornos mentales		
Original o adaptaciones de otros países	3,888	.92
Adaptación española <sup>a</sup>	392	.91
Estudiantes universitarios		
Original o adaptaciones de otros países	3,468	.90
Adaptación española <sup>a</sup>	727	.88
Adultos de la población general		
Original o adaptaciones de otros países	5,009	.89
Adaptación española <sup>a</sup>	348	.92

Nota. <sup>a</sup>Valor medio de los coeficientes de las diferentes muestras de la adaptación española del BAI (Beck y Steer, 2011) tras ponderarlos por el tamaño de cada muestra.

En primer lugar, aunque a partir de las medias y desviaciones típicas de los manuales del BAI el evaluador podría calcular la puntuación típica o la puntuación  $T$  de una persona en concreto (aunque no sus puntuaciones típicas o  $T$  normalizadas), su propio cálculo podría suponer un obstáculo práctico para que dicho evaluador utilizara rutinariamente la información que proporciona esta otra forma de interpretar las puntuaciones del BAI. En segundo lugar, las puntuaciones centiles no pueden calcularse a partir de la media y desviación típica del grupo de referencia y, precisamente, tales puntuaciones destacan por su facilidad de interpretación (p. ej., decir que una persona tiene el centil 75 en el BAI es lo mismo que decir que su ansiedad es más grave que la que tiene el 75% de las personas de su población de referencia) y por su significación universal (p. ej., un centil 75 en el BAI significa lo mismo que un centil 75 en el STAI: la puntuación en cuestión es superior a la alcanzada por el 75% de la población respectiva). Debido a estas ventajas, Yela (1984) aconsejaba la tipificación de un test en normas centiles si éste iba ser destinado al uso práctico, principalmente al clínico.

En las tablas 7 y 8 se presentan puntuaciones centiles del BAI para pacientes españoles con trastornos psicológicos, para adultos de la población general española y para estudiantes universitarios españoles, las cuales permiten evaluar con mayor precisión la gravedad de la ansiedad en ese tipo de muestras. Para obtener dichas tablas, se analizaron los datos de las muestras de participantes de la adaptación española del BAI (Beck y Steer, 2011), las cuales en realidad eran muestras ampliadas de las que habían participado en los estudios previos sobre dicha adaptación (Magán et al., 2008; Sanz et al., 2012; Sanz y Navarro, 2003). En el manual de la adaptación española del BAI (Beck y Steer, 2011) se puede encontrar información más detallada sobre las características sociodemográficas y clínicas de las muestras de participantes en las que están basadas las puntuaciones centiles de las tablas 7 y 8.

Puesto que en cada una de esas tres muestras de participantes la puntuación total media en el BAI de las mujeres fue significativamente mayor que la de los varones,  $t(390) = -2.51, p = .013, d$  de Cohen = 0.27 en la muestra de pacientes,  $t(346) = -2.97, p = .003, d$  de Cohen = 0.32 en la muestra de población general, y  $t(723) = -3.14, p =$

**Tabla 7**

Puntuaciones centiles del BAI para la muestra total de pacientes externos con trastornos psicológicos de la adaptación española del BAI (Beck y Steer, 2011) y para sus submuestras de pacientes varones y mujeres

Centiles	Pacientes externos con trastornos psicológicos		
	Total	Varones	Mujeres
	N = 392	n = 131	n = 261
1	1	0	1
5	3	2	4
10	5	4	5
20	8	6	9
25	9	8	10
30	11	9	12
40	13	13	14
50	16	16	17
60	20	18	22
70	25	22	26
75	27	25	28
80	29	26	30
90	36	33	38
95	40	38	42
99	54	44	56
Media	18.89	16.78	19.95
DT	11.89	10.85	12.26

.002,  $d$  de Cohen = 0.32 en la muestra de estudiantes universitarios, en las tablas 7 y 8 también se presentan las puntuaciones centiles para las correspondientes submuestras de varones y mujeres.

Ambas formas de interpretar las puntuaciones del BAI, la basada en las puntuaciones de corte y la basada en las normas o baremos, son complementarias y deberían combinarse para obtener una visión más adecuada del nivel de sintomatología de ansiedad de una

**Tabla 8**

Puntuaciones centiles del BAI para las muestras totales de adultos de la población general y de estudiantes universitarios de la adaptación española del BAI (Beck y Steer, 2011) y para sus submuestras de varones y mujeres

Centiles	Adultos de la población general			Estudiantes universitarios		
	Total	Varones	Mujeres	Total	Varones	Mujeres
	N = 348	n = 166	n = 182	N = 727	n = 143	n = 582
1	0	0	0	0	0	0
5	0	0	0	1	0	1
10	1	0	2	2	1	2
20	3	2	4	3	3	3
25	4	3	5	4	4	4
30	5	4	6	4	4	5
40	6	5	7	6	5	6
50	7	7	9	7	6	8
60	9	9	11	9	7	9
70	11	10	13	11	9	12
75	13	11	16	13	10	13
80	15	12	19	15	12	16
90	23	19	26	20	16	21
95	29	25	32	24	20	26
99	47	48	48	37	30	38
Media	10.12	8.54	11.56	9.37	7.56	9.81
DT	9.58	8.49	10.28	7.73	5.98	8.04

persona. Así, aunque dos pacientes españoles con trastornos psicológicos pueden puntuar en el BAI en el rango que indica un nivel moderado de ansiedad, por ejemplo, 16 y 25, una consulta a la tabla 7 sugeriría que la primera puntuación indica también que el primer paciente sufre una sintomatología ansiosa más grave que el 50% de los pacientes españoles con trastornos psicológicos, mientras que la segunda puntuación indica que el segundo paciente sufre una sintomatología ansiosa más grave que el 75% de los pacientes españoles con trastornos psicológicos. En este sentido, en la medida que el BAI se ha convertido no sólo en uno de los instrumentos más utilizados para evaluar la sintomatología ansiosa en población clínica, sino también en población no clínica, la posibilidad de contar con normas centiles para adultos españoles de la población general y para estudiantes universitarios españoles permite apreciar mucho mejor las diferencias individuales en ansiedad en las poblaciones no clínicas, ya que en ellas es más difícil encontrar personas que superan las puntuaciones de corte del BAI correspondientes a una ansiedad moderada o grave y, por tanto, los percentiles del grupo de referencia correspondiente permitirían apresar pequeñas diferencias individuales dentro los niveles mínimos y leves de gravedad de la ansiedad.

Por otro lado, a partir de las propiedades psicométricas de la adaptación española del BAI, se pueden proponer pautas de interpretación complementarias para dos objetivos muy importantes en la práctica clínica: evaluar la significación clínica de los cambios terapéuticos e identificar personas con ansiedad clínica.

### Evaluación de la significación clínica de los cambios terapéuticos

Dado que el BAI contiene sólo 21 ítems que se pueden responder en 5-10 minutos y que su corrección es tan fácil y rápida que puede realizarse en 1-2 minutos, es posible aplicar el BAI repetidamente durante el proceso terapéutico para monitorizar los progresos del paciente y evaluar los resultados de la terapia. En este contexto, el BAI puede ayudar a responder a la pregunta de si un paciente está igual, ha mejorado o ha empeorado de su problema de ansiedad o incluso si ya se ha recuperado, es decir, si se ha producido una reducción o aumento clínicamente significativo en su sintomatología ansiosa.

Para ello, se puede utilizar la aproximación estadística a la significación clínica de Jacobson y Truax (1991), que asume que un cambio clínicamente significativo supondría la vuelta a una población funcional de un paciente que antes del tratamiento pertenecía a una población disfuncional, es decir, que dicho cambio supondría que la puntuación de un paciente en el BAI ya no pertenece a la distribución de puntuaciones en el BAI de una población disfuncional (p. ej., los pacientes españoles con trastornos psicológicos) sino a la distribución de una población funcional (p. ej., la población general española).

Para determinar la existencia de un cambio clínicamente significativo en un paciente, el método de Jacobson y Truax (Jacobson y Truax, 1991; McGlinchey, Atkins y Jacobson, 2002) implica, en primer lugar, establecer una puntuación de corte que el paciente debe alcanzar para pasar de una distribución disfuncional a una funcional. Cuando las dos distribuciones se solapan, como ocurre en el BAI (véanse las tablas 7 y 8), la mejor puntuación de corte (C) es el punto medio ponderado entre las medias de las distribuciones funcional y disfuncional:

$$C = \frac{(DT_n \times M_p) + (DT_p \times M_n)}{(DT_n \times M_p)}$$

en la que  $DT_n$  y  $DT_p$  representan las desviaciones típicas en el BAI de la población normal y de pacientes respectivamente y  $M_n$  y  $M_p$  las medias en el BAI de la población normal y de pacientes, respectivamente.

En segundo lugar, el método implica estimar si el cambio que indican las puntuaciones de un instrumento como el BAI no se debe a su error de medida, sino que refleja un cambio fiable, real, en el nivel de sintomatología ansiosa del paciente. Para ello, Jacobson y Truax (1991) proponen un índice de cambio fiable [*reliable change index*, RCI] que tiene en cuenta el error típico de la diferencia entre dos puntuaciones del instrumento ( $s_{dif}$ ), el cual depende de su error típico de medida ( $s_e$ ) que, a su vez, depende de su fiabilidad ( $r_{xx}$ ):

$$RCI = \frac{x_2 - x_1}{s_{dif}}$$

$$s_{dif} = \sqrt{2(s_e)^2} = \sqrt{2(s_x \sqrt{1 - r_{xx}})^2}$$

en las que  $x_2$  sería la puntuación en el BAI de un paciente en un momento dado (p. ej., postratamiento),  $x_1$  la puntuación en el BAI en un momento anterior (p. ej., pretratamiento),  $s_x$  la desviación típica de las puntuaciones del BAI en la población de pacientes y  $r_{xx}$  la fiabilidad de consistencia interna del BAI en dicha población. El error típico de la diferencia entre las dos puntuaciones ( $s_{dif}$ ) describiría la amplitud de la distribución de las puntuaciones de cambio que se esperaría si no ocurriera ningún cambio real, de manera que un RCI mayor que 1.96 sería muy poco probable ( $p < .05$ ) que sucediera sin que ocurriera un cambio real. En consecuencia, el cambio en las puntuaciones en el BAI de un paciente determinado debería superar ese valor del RCI para asegurar que dicho cambio no se debe a los errores de medida del BAI:

$$RCI \Rightarrow 1.96 = \frac{x_2 - x_1}{s_{dif}} > 1.96 \Rightarrow x_2 - x_1 > s_{dif} \times 1.96$$

Basándose en estos dos criterios, el método de Jacobson y Truax (Jacobson y Truax, 1991; McGlinchey et al., 2002) clasifica a un paciente como *recuperado* si su puntuación en un instrumento supone un cambio que supera ese valor de 1.96 del RCI y si dicha puntuación ha superado la puntuación C, *mejorado* si la puntuación supone un cambio que supera el valor de 1.96 del RCI, pero no supera la puntuación C, *sin cambios* si la puntuación no supera ninguno de los dos criterios y *empeorado* si la puntuación supone un cambio que supera el valor de 1.96 del RCI, pero en la dirección de un empeoramiento.

Con los datos de la adaptación española del BAI (véanse las tablas 5, 7 y 8), la puntuación C y el cambio en las puntuaciones del BAI que se corresponde con un valor de 1.96 del RCI serían:

$$C = \frac{(9.58 \times 19.89) + (11.89 \times 10.12)}{(9.58 + 11.89)} = 14.03$$

$$s_{dif} = \sqrt{2(11.89 \sqrt{1 - .91})^2} = 5.04$$

$$x_2 - x_1 > 5.04 \times 1.96 \Rightarrow x_2 - x_1 > 9.87$$

En consecuencia con dicha adaptación, un paciente cuya puntuación en el BAI ha descendido 10 puntos o más y dicha puntuación es menor de 14 se podría considerar *recuperado* de su trastorno de ansiedad, si su puntuación ha descendido 10 puntos o más pero la misma no llega a 14 se podría considerar *mejorado*, si su puntuación no ha descendido 10 puntos y tampoco llega a 14 se podría considerar *sin cambios* y si la puntuación refleja un aumento de 10 puntos o más se podría considerar que ha *empeorado*.

No obstante, es importante recordar que la valoración del estado de un paciente debería tener en cuenta otros criterios sobre sus problemas de ansiedad (p. ej., para hablar de recuperación, habría que considerar si la puntuación < 14 en el BAI se sigue repitiendo duran-

te más semanas) así como tener en cuenta otra información relevante sobre otros problemas y trastornos del paciente, sobre su nivel de funcionamiento en las actividades laborales o sociales habituales o en las relaciones con los demás, etc.

### Identificación de personas con ansiedad clínica

Aunque el BAI no es un instrumento diseñado para diagnosticar trastornos de ansiedad, la información que proporciona puede ayudar a identificar, en la población española no clínica, a personas con un probable trastorno de ansiedad para su posterior confirmación diagnóstica con un instrumento más adecuado (p. ej., una entrevista diagnóstica estructurada) o servir como fuente complementaria de datos que ayude a realizar un diagnóstico de trastorno de ansiedad. Para ello, se pueden utilizar los índices de rendimiento diagnóstico de la tabla 6. Así, cuando se use la adaptación española del BAI como cribado y, por tanto, con confirmación diagnóstica posterior de los casos detectados con el inventario, sería recomendable escoger puntuaciones de corte con una alta sensibilidad, es decir, con una alta capacidad para detectar un trastorno de ansiedad en pacientes que realmente lo están sufriendo, de manera que se asegure que ningún paciente que sufre dicho trastorno quede sin detectar. En este caso es preferible cometer un error debido a un falso positivo, ya que dicho error sería luego fácil de corregir mediante la confirmación diagnóstica posterior, que cometer un error debido a un falso negativo, ya que esta persona con trastorno de ansiedad no sería posteriormente evaluada y pasaría desapercibida. Para la población española no clínica, esa puntuación de corte sería igual o mayor que 12, ya que esta puntuación, obteniendo un adecuado índice de especificidad ( $> 70\%$ ) y de concordancia diagnóstica ( $kappa > .40$ ), podría detectar al 81% de las personas que tienen un trastorno de ansiedad (véase la tabla 6).

Cuando se use la adaptación española del BAI como ayuda para realizar un diagnóstico de trastorno de ansiedad que ya se sospechaba por la información de otros instrumentos, lo recomendable sería escoger puntuaciones de corte con una alta especificidad, es decir, con una alta capacidad para detectar la ausencia de trastorno de ansiedad en personas que realmente no padecen este trastorno, de manera que se asegure que un paciente al que se le asigna ese diagnóstico tiene realmente dicho trastorno (es preferible cometer un error debido a un falso negativo que debido a un falso positivo). De forma relacionada, sería recomendable que dicha puntuación tuviera un valor predictivo positivo mayor de 50%, es decir, que la probabilidad de que una persona que obtenga esa puntuación en el BAI padezca realmente un trastorno de ansiedad esté por encima del 50%. Para la población española no clínica, esa puntuación de corte podría ser igual o mayor que 19, ya que esa puntuación, obteniendo un índice adecuado de concordancia diagnóstica ( $kappa > .40$ ), mostraría una especificidad superior a 90% y un valor predictivo positivo de 65% (véase la tabla 6).

Es importante señalar que las puntuaciones de corte del BAI anteriormente señaladas para los objetivos de cribado y ayuda diagnóstica lo son para la población no clínica española y que sería muy útil que futuras investigaciones pudieran identificar las correspondientes puntuaciones de corte en la población española de pacientes con trastornos psicológicos o en otras poblaciones clínicas (p. ej., pacientes con enfermedades médicas).

### Conclusiones

Con la publicación en 2011 de la adaptación española del BAI y la constatación de sus buenos índices de fiabilidad y validez en muestras clínicas y no clínicas españolas (Beck y Steer, 2011), este inventario se incorpora, con todas las garantías psicométricas de un proceso de validación más sistemático y riguroso en nuestro país, al arsenal de instrumentos específicos para medir la ansiedad (p. ej., el

Inventario de Respuestas y Situaciones de Ansiedad o ISRA) y, en concreto, de instrumentos breves (p. ej., el STAI-R), que forman parte del instrumental actual de los psicólogos clínicos y de la salud españoles, y cabría esperar que en los próximos años, al igual que ha ocurrido en otros países (Piotrowski y Gallant, 2009), el BAI se convierta en uno de los instrumentos más utilizados de ese instrumental.

Frente al STAI-R, el BAI se caracteriza por evaluar peor el trastorno de ansiedad generalizada, pero por evaluar mucho mejor los síntomas que conforman las crisis de angustia (o ataques de pánico), lo cual es muy importante en el área clínica ya que tales crisis pueden aparecer en el contexto de casi todos los trastornos de ansiedad. Así mismo, frente al STAI-R, el BAI se caracteriza por discriminar mejor entre ansiedad y depresión.

La adaptación española del BAI muestra buenos índices psicométricos como instrumento de medida de la presencia y gravedad de la ansiedad, los cuales son similares a los que presenta la versión original o las adaptaciones de otros países, índices que permiten proponer pautas, baremos y puntuaciones de corte útiles para la evaluación de la gravedad de la ansiedad, la valoración de la significación clínica de los cambios terapéuticos, el cribado de personas con ansiedad clínica y la ayuda en el diagnóstico diferencial de los trastornos de ansiedad.

### Extended summary

In its 25 years of existence the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988) has become the most widely used self-report instrument for assessing the severity of anxiety in the countries with higher research productivity in psychology (Piotrowski & Gallant, 2009), overtaking the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) that was in first place ten years before (Piotrowski, 1999). In Spain, the BAI has also grown in popularity, although the STAI is still the most widely used instrument for assessing anxiety (Muñiz & Fernández-Hermida, 2010). This paper presents the basic characteristics of the BAI in comparison with the Trait Anxiety Scale of the STAI (STAI-T), describes the process of adaptation to the Spanish population of the BAI, summarizes the psychometric properties of this adaptation, and discusses its use in the clinical practice for measuring the severity of anxiety, evaluating the clinical significance of therapeutic change, screening for clinical anxiety, and aiding in the differential diagnosis of anxiety disorders.

### BAI versus STAI-T

Table 1 shows the main similarities and differences between the BAI and the STAI-T in the context of their use in health and clinical psychology. Some of their similarities are the brevity of both instruments (20–21 items) and their good or excellent indices of internal consistency reliability ( $\geq .80$  or  $\geq .85$ ; Prieto & Muñiz, 2000). Among their differences, the most remarkable are: (a) the constructs they intend to measure; (b) the content of their items and the anxiety and depression symptoms that such items assess; and (c) their abilities to discriminate between anxiety and depression. The STAI-T was developed to assess a normal personality trait, whereas the BAI was developed to assess anxious symptoms (clinical anxiety). This difference is also evident in their instructions and response formats, since the STAI-T asks respondents about the general frequency of certain behaviors, whereas the BAI asks individuals to rate how much he or she has been bothered by certain behaviors over the past week.

A second difference between the BAI and the STAI-T lies in the content of their items and the extent to which they reflect clinical anxiety. Tables 2 and 3 show, from a content validity perspective, the correspondence between the BAI and STAI-T items and the DSM-IV diagnostic criterion symptoms for generalized anxiety disorder



(GAD) and panic attacks (APA, 1994/1995). As can be seen in those tables, the STAI-T covers more TAG symptoms than the BAI and has more items the content of which is related to TAG, whereas the BAI covers more panic attack symptoms than the STAI-T and has more items the content of which is related to panic attack. This last feature of the BAI is very important since both full and limited-symptom (partial) panic attacks can occur in the context of most anxiety disorders (e.g., specific phobia, social phobia, obsessive-compulsive disorder, panic disorder, agoraphobia, posttraumatic stress disorder, acute stress disorder).

On the other hand, an instrument has high relevance to assess clinical anxiety not only if its items cover all anxious symptoms, but also if, in addition, they do not reflect symptoms of other disorders, for example, depressive symptoms (Sanz et al., in press). Table 4 shows the correspondence between the BAI and STAI-T items and the DSM-IV diagnostic criterion symptoms for the major depressive episode (APA, 1994/1995). The information of Table 4 indicates that the BAI does not have items the content of which is related to major depressive episode symptoms, whereas 65% of STAI-T items assess depressive symptoms. This poses a discriminant validity problem to the STAI-T. Furthermore, the correlation between the BAI and the Beck Depression Inventory-II (BDI-II; Beck, Steer y Brown, 1996) is significantly lower than the correlation between the STAI-T and the BDI-II, at least in university student samples from a variety of countries (see Table 1).

In sum, the BAI show a better ability to discriminate anxiety and depression than the STAI-I in terms of both content validity and discriminant validity with respect to the BDI-II. In fact, several empirical studies in clinical and non-clinical samples, including Spanish samples, have confirmed that the STAI-T measures both anxiety and depression (Bados, Gómez-Benito, & Balaguer, 2010; Bieling, Antony, & Swinson, 1998; Caci, Baylé, Dossios, Robert, & Boyer, 2003).

#### *Spanish adaptation of the BAI*

In 2011, *Pearson Clinical & Talent Assessment España* published the Spanish adaptation of the BAI (Beck & Steer, 2011). This adaptation is based on a thorough analysis of its psychometric properties in 392 patients with mental disorders, 348 adults from the general population, and 727 university students. For the purpose of this article, it is enough to note that Cronbach's alpha coefficients of internal consistency reliability for those three samples were, respectively, .90, .91, and .88, and, therefore, all of them were excellent ( $\geq .85$ ; Prieto & Muñiz, 2000) and similar to those found in studies in other countries (see Table 5). Concerning diagnostic validity, an analysis of receiver operating characteristic (ROC) curve revealed that the BAI exhibited an appropriate diagnostic performance to discriminate between university students with and without anxiety disorder (area under ROC curve = .80, that is,  $> 0.70$ ; Swets, 1988), with several cut-off scores showing appropriate indices of diagnostic validity (see Table 6). Concerning contrasted-groups validity, the mean BAI score of patients with mental disorders was significantly higher than those of adults from the general population and university students (Beck & Steer, 2011), with large effect sizes (Cohen's *d*) of 0.81 and 1.01, respectively.

#### *Recommendations for the use of the Spanish adaptation of the BAI in clinical practice*

Interpretation of BAI scores is usually based on cut-off scores that define different levels of the severity of anxious symptomatology. The current edition of the original BAI proposes the following cut-off scores and definitions: 0–7 indicates minimal anxiety, 8–15 mild, 16–25 moderate, and 26–63 severe (Beck & Steer, 1993). These cut-off scores were adopted for the Spanish adaptation of the BAI.

#### *Norms for the Spanish adaptation of the BAI*

Additionally, interpretation of BAI scores can be made by comparing such scores with norms of a reference group. Tables 7 and 8 display independent tables of BAI centile scores for Spanish patients with mental disorders, adults from the Spanish general population, and Spanish university students. These centile score tables allow one to assess, with greater accuracy, the severity of anxiety in those samples. To calculate Tables 7 and 8, data from the participant samples in the Spanish adaptation of the BAI (Beck & Steer, 2011) were analyzed. In each sample, the women's mean BAI scores were significantly higher than men's, and, therefore, Tables 7 and 8 also display independent tables of BAI centile scores for women and men.

Both ways of interpreting BAI scores, that based on cut-off scores and that based on norms, are complementary and should be combined to obtain a more appropriate view of a person's anxious symptomatology level.

Likewise, on the basis of the psychometric properties of the Spanish adaptation of the BAI, it is possible to suggest complementary interpretation guidelines to achieve two important aims in clinical practice: assessing the clinical significance of therapeutic changes and identifying persons with clinical anxiety.

#### *Assessing the clinical significance of therapeutic changes*

Given its brevity, the BAI can be administered repeatedly during the course of therapy in order to monitor patient's progress and assess therapy outcomes. Thus, the BAI can help to answer the question of whether there is a clinically significant decrease (or increase) in a patient's anxious symptomatology. To do this, Jacobson and Truax's (1991) statistical approach to clinical significance can be used. This approach assumes that a clinically significant change occurs when a patient's BAI score no longer belongs to the BAI score distribution of a dysfunctional population, but it returns to the BAI score distribution of a functional population.

Jacobson and Truax's (1991) method involves, firstly, establishing a cut-off point (C) for each client that must be crossed in moving from the dysfunctional to the functional distribution. Secondly, that method involves determining whether a patient's change in BAI scores (e.g., from pre- to posttest) is reliable, rather than simply an artifact of measurement error. To assess this, Jacobson and Truax (1991) proposed a reliable change index (RCI) that each patient has to pass in order to demonstrate that his or her change in anxious symptomatology is not simply due to measurement error. This RCI takes into account the standard error of the difference between two scores that depends on the standard error of measurement that, in turn, depends on reliability and standard deviation of test scores.

Applying Jacobson and Truax's (1991) method with the data of the Spanish adaptation of the BAI in patients with mental disorders (see Tables 5 and 7), a patient could be classified as *recovered* if his or her BAI score shows a decrease of 10 points or greater and is lower than 14 (i.e., passes both RCI and cut-off criteria), *improved* if his or her BAI score shows a decrease of 10 points or greater but is not lower than 14 (i.e., passes RCI criterion, but not the cut-off criterion), *unchanged* if his or her BAI score does not show a decrease of 10 points or greater and is not lower than 14 (i.e., passes neither criteria), or *deteriorated* if his or her BAI score shows an increase of 10 points or greater (i.e., passes RCI criteria in a worsening direction).

However, it is important to remind that the assessment of a patient's state of health should take into account other criteria on his or her anxiety problems as well as relevant information on other patient's disorders, his or her functioning level in social and occupational areas, etc.

### Identifying persons with clinical anxiety

Although the BAI is not an instrument designed to diagnose anxiety disorders, the information that it provides may help identify persons with a probable anxiety disorder in order to later make a diagnostic confirmation with a more appropriate instrument (e.g., a structured diagnostic interview) or may be useful as a complementary source of data to make a diagnosis of anxiety disorder. To achieve these goals, the diagnostic performance indices shown in Table 6 can be used. Thus, when using the Spanish adaptation of the BAI as a screening instrument in the Spanish non-clinical population, a cut-off score of 12 can be used, since this BAI score can detect 81% of persons who suffer anxiety disorders and shows an adequate specificity index (> 70%) and an adequate index of diagnostic agreement ( $\kappa > .40$ ) (see Table 6). When using the Spanish adaptation of the BAI as an aid to make a diagnosis of anxiety disorder in non-clinical population, a cut-off score of 19 can be used, since this BAI score shows a specificity higher than 90% and a positive predictive value of 65% while obtaining an appropriate index of diagnostic agreement ( $\kappa > .40$ ) (see Table 6).

### Conclusions

With the publication in 2011 of the Spanish adaptation of the BAI and with the confirmation of its good indices of reliability and validity in clinical and non-clinical Spanish samples (Beck & Steer, 2011), this inventory joins the broad arsenal of instruments to measure anxiety that Spanish health and clinical psychologists have to choose from. In comparison with the STAI-T, the BAI is distinguished by a poorer assessment of GAD, but a better assessment of panic attacks. This is very important since panic attacks can occur in the context of most anxiety disorders. In addition, the BAI is distinguished by a better discrimination between anxiety and depression. The good psychometric properties of the Spanish adaptation of the BAI allow us to propose guidelines, norms and cut-off scores that can be useful to assess the severity of anxiety and the clinical significance of therapeutic changes, to screen for anxiety disorders, and to make a diagnosis of anxiety disorder.

### Conflicto de intereses

Los autores de este artículo declaran que no tienen ningún conflicto de intereses.

### Referencias

American Psychiatric Association (1995). *DSM-IV. Manual diagnóstico y estadístico de los trastornos mentales* (orig. 1994). Barcelona: Masson.

- Bados, A., Gómez-Benito, J. y Balaguer, G. (2010). The State-Trait Anxiety Inventory, trait version: Does it really measure anxiety? *Journal of Personality Assessment*, 92, 560–567.
- 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.
- Beck, A. T. y Steer, R. A. (1993). *Manual for the Beck Anxiety Inventory*. San Antonio, TX: The Psychological Corporation.
- Beck, A. T. y Steer, R. A. (2011). *Manual. BAI. Inventario de Ansiedad de Beck* (Adaptación española de Sanz, J.). Madrid: Pearson Educación.
- Beck, A. T., Steer, R. A. y Brown, G. K. (1996). *BDI-II. Beck Depression Inventory-Second Edition manual*. San Antonio, TX: The Psychological Corporation.
- Bieling, P. J., Antony, M. M. y Swinson, R. P. (1998). The State-Trait Anxiety Inventory, trait version: structure and content re-examined. *Behaviour Research and Therapy*, 36, 777–788.
- Caci, H., Baylé, F., Dossios, C. Robert, P. y Boyer, P. (2003). The Spielberger Trait Anxiety Inventory measures more than anxiety. *European Psychiatry*, 18, 394–400.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ª ed.). Hillsdale, NJ: LEA.
- Fonseca-Pedrero, E., Paino, M., Sierra-Baigrie, S., Lemos-Giráldez, S. y Muñiz, J. (2012). Propiedades psicométricas del "Cuestionario de ansiedad estado-rasgo" (STAI) en universitarios. *Behavioral Psychology-Psicología Conductual*, 20, 547–561.
- Guillén-Riquelme, A. y Buela-Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial del ítem en el State Trait Anxiety Inventory (STAI). *Psicothema*, 23, 510–515.
- Jacobson, N. S. y Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19.
- Magán, I., Sanz, J. y García-Vera, M. P. (2008). Psychometric properties of a Spanish version of the Beck Anxiety Inventory (BAI) in general population. *The Spanish Journal of Psychology*, 11, 626–640.
- McGlinchey, J. B., Atkins, D. C. y Jacobson, N. S. (2002). Clinical significance methods: which one to use and how useful are they? *Behavior Therapy*, 33, 529–550.
- Muñiz, J. y Fernández-Hermida, J. R. (2000). La utilización de los test en España. *Papeles del Psicólogo*, 76, 41–49.
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*, 31, 108–121.
- Piotrowski, C. (1999). The status of the Beck Anxiety Inventory in contemporary research. *Psychological Reports*, 85, 261–262.
- Piotrowski, C. y Gallant, N. (2009). Research use of clinical measures for anxiety in the recent psychological literature. *Journal of Instructional Psychology*, 36, 84–86.
- Prieto, G. y Muñiz, J. (2000). Un modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 77, 65–71.
- Sanz, J., García-Vera, M. P. y Fortún, M. (2012). El "Inventario de ansiedad de Beck" (BAI): propiedades psicométricas de la versión española en pacientes con trastornos psicológicos. *Behavioral Psychology-Psicología Conductual*, 20, 563–583.
- Sanz, J., Izquierdo, A. y García-Vera, M. P. (en prensa). Una revisión desde la perspectiva de la validez de contenido de los cuestionarios, escalas e inventarios autoaplicados más utilizados en España para evaluar la depresión clínica en adultos. *Psicopatología Clínica Legal y Forense*.
- Sanz, J. y Navarro, M. E. (2003). Propiedades psicométricas de una versión española del Inventario de Ansiedad de Beck (BAI) en estudiantes universitarios. *Ansiedad y Estrés*, 9, 59–84.
- Spielberger, C. D., Gorsuch, R. L. y Lushene, R. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologist Press.
- Spielberger, C. D., Gorsuch, R. L. y Lushene, R. (1982). *Manual del Cuestionario de Ansiedad Estado-Rasgo (STAI)*. Madrid: TEA Ediciones.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R. y Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory STAI (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.
- Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.
- Yela, M. (1984). *Psicometría y teoría de los tests*. Facultad de Psicología. Universidad Complutense de Madrid.