Psychometric Properties of a Spanish Version of the Beck Anxiety Inventory (BAI) in General Population

Inés Magán, Jesús Sanz, and María Paz García-Vera
Universidad Complutense (Spain)

This is the first study that provides normative, reliability, factor validity and discriminant validity data of the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988) in the Spanish general population. Sanz and Navarro's (2003) Spanish version of the BAI was administered to 249 adults. Factor analyses suggested that the BAI taps a general anxiety dimension comprising two related factors (somatic and affective-cognitive symptoms), but these factors hardly explained any additional variance and, therefore, little information is lost in considering only full-scale scores. Internal consistency estimate for the BAI was high (α = .93). The BAI was correlated .63 with the BDI-II and .32 with the Trait-Anger scale of the STAXI 2, but a factor analysis of their items revealed three factors, suggesting that the correlations between the instruments may be better accounted for by relationships between anxiety, depression, and anger, than by problems of discriminant validity. The mean BAI total score and the distribution of BAI scores were similar to those found in other countries. BAI norm scores for the community sample were provided from the total sample and from the male and female subsamples, as females scored higher than males. The utility of these scores for assessing clinical significance of treatment outcomes for anxiety is discussed.
Keywords: anxiety, BAI, reliability, validity, norms

Se presentan por primera vez datos normativos, de fiabilidad, validez factorial y validez discriminante del Inventario de Ansiedad de Beck (BAI; Beck, Epstein, Brown y Steer, 1988) en la población general española. La versión española del BAI de Sanz y Navarro (2003) fue administrada a 249 adultos. Los análisis factoriales indicaron que el BAI mide una dimensión general de ansiedad compuesta de dos factores relacionados (somático y afectivo cognitivo), pero estos factores apenas explicaban varianza adicional por lo que no se pierde mucha información al considerar únicamente la puntuación global. La fiabilidad de consistencia interna del BAI fue elevada (alfa = 0,93). El BAI correlacionó 0,63 con el BDI-II y 0,32 con la escala de Ira Rasgo del STAXI 2, pero el análisis factorial de los tres instrumentos reveló que sus ítems formaban tres factores, sugiriendo que las correlaciones entre instrumentos se deben más a la relación entre ansiedad, depresión e ira que a un problema de validez discriminante. La puntuación media en el BAI y la distribución de sus puntuaciones fueron similares a las encontradas en otros países. Se ofrecen baremos para la muestra total y dividida por el sexo, ya que las mujeres puntuaron más alto que los varones, y se discute su utilidad para evaluar la significación clínica de los resultados de los tratamientos para la ansiedad.
Palabras clave: ansiedad, BAI, fiabilidad, validez, normas

Correspondence concerning this article can be addressed to Jesús Sanz, Departamento de Personalidad, Evaluación y Psicología Clínica, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid. E-mail: jsanz@psi.ucm.es

How to cite the authors of this article: Magán, I., Sanz, J. and García-Vera, M.P.

The Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988) is one of the most frequently used self-report instruments to assess anxious symptomatology in patients with psychological disorders and in the normal population, both in clinical practice and research (Piotrowski, 1999; Sanz & Navarro, 2003). In Spain, in the last few years, the BAI has become increasingly popular as an instrument to assess anxiety both in basic and applied research (see the list of studies collected in Sanz & Navarro, 2003). In fact, various translations to Spanish of the BAI have been published (i.e., Botella & Ballester, 1997; Comeche, Díaz, & Vallejo, 1995; Echeburúa, 1993). However, after reviewing the most important bibliographic databases in psychology and similar disciplines, both in Spanish (PSICODOC, ISOC) and in other languages (PsycINFO, MEDLINE, Web of Science and SCOPUS), up to January 2008, we have not found any study that specifically analyzes the psychometric properties of any of these translations in any Spanish samples (i.e., psychopathological patients, general population, university students, medical patients).

Sanz and Navarro (2003) elaborated a new Spanish version of the BAI and analyzed its psychometric properties (reliability, content validity, criterion validity, factor validity, discriminant validity, and standardization) in a sample of Spanish university students. This study continues the process of adaptation of this Spanish version of the BAI, contributing normative values and data about its reliability, factor validity, and discriminant validity in a sample of adults extracted from the general Spanish population. A search in the above-mentioned bibliographic databases indicates that this is the first study to analyze the psychometric properties in the general Spanish population of Sanz and Navarro's version of the BAI or of any other Spanish translation of the BAI.

This information will not only allow the use of the BAI as an assessment instrument of anxious symptomatology in the general Spanish population, but also to establish the specific criteria to assess the clinical significance of the results obtained with therapies for anxiety.

The BAI has been used in a multitude of studies to assess response to treatment of groups of patients with anxiety disorders, especially panic disorder and generalized anxiety disorder, and to determine the most effective treatment (see, for example, the reviews of the literature of Ayers, Sorrell, Thorp, & Wetherell, 2007; Chambless & Gillis, 1993; McEvoy & Nathan, 2007; Siev & Chambless, 2007). The BAI is habitually administered before and after treatment and, among other parameters, a statistically significant reduction in the mean BAI score is considered a positive response to treatment, at least as far as reduction of anxious symptomatology is concerned. However, it is obvious that the goal of any therapeutic intervention is not to achieve a statistically significant improvement in the average problem of a group of patients—or at least not only that—but instead to achieve a clinically significant improvement; that is, an improvement of a clinically relevant magnitude, and with practical effects in the patients' lives that leads to a recovery of their normal functioning.

Investigators dedicated to the study of the assessment of treatments have developed diverse criteria to assess the degree to which a treatment produces a clinically significant improvement in a group of patients or whether a particular patient has improved in a clinically significant way or has recovered (see Kazdin, 1992). A much used procedure, sometimes known as *normative comparisons*, is to take into account whether, after completing treatment, the patient's score approaches the mean score of a "normal" reference group, that is, whether or not the patient, after treatment, is different from normal people regarding their symptoms and main complaints (Jacobson & Truax, 1991; Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999).

In order to value this criterion, the patient's posttreatment score can be analyzed to see if it is equal to or lower than the mean or the median of the norms of a sufficiently large and representative sample of the general population (Hollon & Flick, 1988), or whether it falls near this mean or median, at least within the interval of one standard deviation above the mean (Kendall & Grove, 1988; Kendall et al., 1999). At the group level, treatment efficacy would be observed, for example, in the percentage of patients that, at posttreatment, obtain a score equal to or lower than the mean (or median), or a score within the interval of one standard deviation around that mean (or median). Therefore, in order to assess this criterion of recovery or clinically significant improvement, information is required about the distribution of the scores (central tendency statistics or dispersion statistics) obtained in a "normal" population with the instrument. Ideally, these normative values should be obtained from a sample of the stratified general population as a function of diverse sociodemographic variables and, of course, only if the instrument presents an acceptable level of reliability that allows one to assume that such values are generalizable to the population from which the sample was extracted.

Summing up, the goal of the present work is to obtain normative values and data about the reliability, factor validity, and discriminant validity of Sanz and Navarro's (2003) Spanish version of the BAI in a sample of the general Spanish population in order to offer to Spanish researchers and professionals working in the area of anxiety an instrument that serves to measure anxious symptomatology in this population, and to establish clinically significant recovery or improvement criteria based on normative comparisons when assessing the efficacy and effectiveness of treatments for anxiety disorders and problems.

Method

Participants

In this study, 262 adults participated initially, selected from the general population of the Region of Madrid by means of the "snow ball" technique: we asked a group of

university psychology students to invite their relatives and friends to participate in a study about personality and anger-hostility, according to certain criteria that assured some heterogeneity of the sample as far as participants' age and sex is concerned. The data of 13 persons who returned an incomplete questionnaire or who did not indicate their sex or age were discarded, so the final sample comprised 249 people, 131 female and 118 male, with ages between 18 and 78 years ($M = 37.8$, $SD = 16.2$). Despite the fact that a sample obtained this way is not random, we achieved reasonable heterogeneity in sex and age and, in fact, for some levels defined by these variables (i.e., the groups of males and females from 35 to 54 years), their profile was very similar to that of the Spanish population (see Table 1). More information about the sociodemographic characteristics of this sample are displayed in Table 2.

Instruments

Beck Anxiety Inventory (BAI; Beck et al., 1988). The BAI is a 21-item self-report with an inventory format designed to assess the severity of clinical anxiety symptomatology. Each BAI item reflects an anxiety symptom and for each one, respondents rate the degree to which they were affected by it during the past week, on a 4-point Likert-type scale, ranging from 0 (*not at all*) to 3 (*severely; I could barely stand it*). Regarding scoring, each item is assigned 0 to 3 points, depending on the individual's response and, after directly adding the score of each item, a total score, ranging from 0 to 63, can be obtained. Various psychometric studies guarantee the reliability and validity of the BAI in very diverse samples (psychiatric patients, patients with anxiety disorders, adolescents with mental disorders, elders, medical patients, university students; see the studies cited

Table 1
Comparison of the Characteristics (Sex and Age) of the Sample of Participants in this Study with the Spanish Population

	Study Sample		Spanish Population*	
Age	Male ($n = 118$)	Female ($n = 131$)	Male ($N = 16,243,472$)	Female ($N = 17,262,495$)
From 18 to 34 years	22.9 %	27.7 %	16.9 %	16.2 %
From 35 to 54 years	15.3 %	16.1 %	16.9 %	17.0 %
55 years or over	9.2 %	8.8 %	14.6 %	18.3 %
Subtotals	47.4 %	52.6 %	48.5 %	51.5 %

Note. * INE (Instituto Nacional de Estadística [National Statistics Institute], 2004).

Table 2
Sociodemographic Characteristics of the Sample (in Percentages)

	Total ($N = 249$)
Educational level	
No studies	3.6
Primary. Compulsory or equivalent	17.6
Secondary. High school, professional training, or equivalent	20.5
University or specialized	50.5
Other unofficial studies	1.2
No reply	6.4
Profession	
Qualified worker	12.4
Services sector	11.6
Administrative personnel	6
Entrepreneur	0.8
Professional or technician	24.5
Housewife	10.8
Student	24.9
Retired or pensioner	4.4
Unemployed	0.8
Other occupation	0.8
No reply	2.8

in Sanz & Navarro, 2003), including adults from the general population (Gillis, Haaga, & Ford, 1995; Jylhä & Isometsä, 2006; Marai, 2004; Nordhagen, 2001; Osman, Barrios, Aukes, Osman, & Markway, 1993). In this study, we used the Spanish version of the BAI developed by Sanz and Navarro (2003).

Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996). The BDI is a 21-item self-report designed to assess the severity of depressive symptomatology. In each item, respondents choose, from a set of four alternatives arranged from lower to higher severity, the sentences that best describe the way they had been feeling during the past two weeks, including the day they complete the instrument. Each item is assigned 0 to 3 points, depending on the individual's response and, after directly adding the score of each item, a total score, ranging from 0 to 63, can be obtained. In this study, we used the Spanish version of the BDI-II developed by Sanz, Navarro, and Vázquez (2003), whose psychometric properties have been examined both in samples of university students (Sanz, Navarro, et al., 2003), patients with psychological disorders (Sanz, García-Vera, Espinosa, Fortún, & Vázquez, 2005), and in a sample of adults from the general Spanish population (Sanz, Perdigón, & Vázquez, 2003), obtaining in all cases adequate reliability and validity indexes that are comparable to those found in previous studies carried out in other countries.

Trait-State Anger Expression Inventory-2 (STAXI-2; Spielberger, 1999). The STAXI-2 is a 49-item self-report instrument with a 4-point Likert-type format (from 0 to 3) that measures various facets of the construct of anger (anger as state and trait, and different styles of expression and control of anger). In the present study, we only used the scale of Trait-Anger, which assesses, by means of 10 items, the tendency to perceive many different situations as annoying or frustrating and to respond to them with a high state of anger. The Spanish adaptation of the STAXI-2 has adequate reliability and validity indexes that are similar to those of the original version (Spielberger, Miguel-Tobal, Casado, & Cano-Vindel, 2001). Specifically, in the Spanish normative sample, the Trait-Anger Scale obtained a Cronbach's alpha of .82 (Spielberger et al., 2001), which was subsequently replicated in another sample of the general Spanish population (Sanz, Magán, & García-Vera, 2006) and that generally suggests appropriate levels of reliability and internal consistency for this scale.

Procedure

We administered individually to all participants the BAI, together with the BDI-II, the STAXI-2, and two other questionnaires about thoughts and attitudes related to anger and hostility which were the goals of another investigation (the Anger-Hostility-related Thoughts Inventory [in Spanish,

the "Inventario de Pensamientos Relacionados con la Ira-Hostilidad" or IPRI] and the Anger-Hostility-related Attitudes and Beliefs Inventory [in Spanish, the "Inventario de Actitudes y Creencias Relacionadas con la Ira-Hostilidad" or IACRI]; Magán, Sanz, & García-Vera, 2008a, 2008b). After handing in a signed informed consent form, the participants filled in a notebook that included the inventories in the following order: IPRI, IACRI, STAXI-2, BAI and BDI-II. The notebooks were administered by the psychology students who had invited the participants to collaborate in this study, as part of the activities of a voluntary seminar. The students' training and supervision when administering the notebooks was carried out by the first author of this work, during the voluntary seminar.

Results and Discussion

Factor Validity

Along the lines of previous literature (Beck et al., 1988; Hewitt & Norton, 1993; Steer, Rissmiller, Ranieri, & Beck, 1993), we decided to conduct a principal axis analysis of the correlation matrix of the 21 items of the BAI, because both the Kaiser-Meyer-Olkin test of sample adequacy (KMO = .92) and Bartlett's sphericity test ($\chi^2 = 2670.2$, $p < .001$) indicated that the factor model was adequate for the data of our present sample. Velicer's minimum average partial correlation (MAP) test, performed with the SPSS instructions program of O'Connor (2000), indicated a one-factor solution. Simulation studies show that the MAP is one of the best methods to assess the dimensionality of a data matrix (Zwick & Velicer, 1986). This single factor, which explained 42.7% of the variance, reflects the construct of anxiety. The plausibility of this one-factor solution is also supported by the fact that all the items of the BAI, with the sole exception of Item 20, had a factor loading of more than .40 on the one-factor solution, and the loading was also relatively high for Item 20 on that factor (.38). In Table 3 are displayed the factor loadings of the BAI items in the one-factor solution.

As some previous factor studies, both in samples from the general population (Freeston, Ladouceur, Thibodeau, Gagnon, & Rhéaume, 1994), in university students (Creamer, Foran, & Bell, 1995), or psychiatric patients (Beck et al., 1988; Hewitt & Norton, 1993; Steer et al., 1993) have obtained a bi-factor solution with one somatic factor and one affective-cognitive factor, we performed a second factor analysis in which we retained the first two factors extracted by means of principal axes and subsequently rotated them with an oblique procedure (*promax*). The factor loadings for this bi-factor solution are also presented in Table 3, in which it can be observed that the two factors obtained correspond to the somatic and affective-cognitive factors

Table 3
Factor Analysis of the BAI

Item	Principal Axes Analysis			Schmid–Leiman Transformation		
	1 Factor	2 Factors		2nd-order-Factor	1st-order-Factors	
		1	2		1	2
1. Hormigueo o entumecimiento [numbness or tingling]	**.65**	**.40**	.29	**.60**	.21	.15
2. Sensación de calor [feeling hot]	**.46**	**.37**	.13	**.42**	.19	.07
3. Debilidad en las piernas [wobbliness in legs]	**.61**	**.39**	.27	**.56**	.20	.14
4. Incapacidad para relajarme [unable to relax]	**.62**	.27	**.39**	**.56**	.14	.20
5. Miedo a que suceda lo peor [fear of the worst happening]	**.64**	−.17	**.87**	**.60**	−.09	**.45**
6. Mareos o vertigos [dizzy or lightheaded]	**.66**	**.61**	.10	**.61**	.32	.05
7. Palpitaciones o taquicardia [heart pounding or racing]	**.69**	**.68**	.07	**.64**	**.35**	.03
8. Sensación de inestabilidad [unsteady]	**.77**	**.36**	**.47**	**.70**	.18	.24
9. Sensación de estar aterrorizado [terrified]	**.75**	−.01	**.82**	**.70**	−.00	**.43**
1. Nerviosismo [nervous]	**.66**	.20	**.51**	**.61**	.10	.27
11. Sensación de ahogo [feelings of choking]	**.71**	**.70**	.06	**.65**	**.36**	.03
12. Temblor de manos [hands trembling]	**.68**	**.37**	.35	**.62**	.19	.18
13. Temblor generalizado [shaky]	**.76**	.30	**.51**	**.70**	.16	.27
14. Miedo a perder el control [fear of losing control]	**.62**	.10	**.57**	**.58**	.05	.30
15. Dificultad para respirar [difficulty breathing]	**.71**	**.93**	-.16	**.66**	**.48**	−.08
16. Miedo a morir [fear of dying]	**.52**	.09	**.47**	**.48**	.05	.24
17. Estar asustado [scared]	**.67**	−.15	**.88**	**.63**	−.08	**.46**
18. Indigestión o molestia abdominal [indigestion or discomfort in abdomen]	**.51**	**.48**	.06	**.46**	.25	.03
19. Sensación de desmayarse [faint]	**.53**	**.63**	−.06	**.48**	.32	−.03
2. Rubor facial [face flushed]	**.38**	.34	.07	**.35**	.17	.03
21. Sudoración [sweating (not due to heat)]	**.52**	**.61**	−.05	**.47**	.31	−.02

Note. $N = 249$. In the principal axes analysis, we present the factor matrix for the 1-factor solution and the configuration matrix (after *promax* rotation) for the 2-factor solution. In all cases, factor loadings ≥ .5 are in boldface. [Translator's note: the English translations of the items are in brackets after the Spanish version.]

described in the literature. In effect, the first factor, which explained 42.7% of the variance of the data, was defined by 10 items of the BAI, which presented significant loadings (> .35) on that factor and negligible loadings (< .35) on the second factor, and whose content was clearly somatic or vegetative: "difficulty breathing," "feelings of choking," "heart pounding or racing," "faint," "dizzy or lightheaded," "sweating," "abdominal discomfort," "numbness- tingling," "wobbliness in legs," and "feeling hot." The second factor, which explained 6.8%% of the variance, was made up of 8 items with significant loadings (> .35) on that factor, and negligible loadings (< .35) on the first factor: "scared," "fear of the worst happening," "terrified," "fear of losing control," "nervous," "shaky," "fear of dying," and "unable to relax." Except for "shaky," these items reflect purely affective-cognitive symptoms. The only items that have no clear location in the bi-factor solution were "face flushed," "hands trembling," and "unsteady." The first item loaded higher on the somatic factor, almost but not quite reaching (.34) the

significant loading criterion (.35), whereas the two latter items had loadings higher than .35 on both the first and second factor; however, "hands trembling" loaded higher on the first or somatic factor (.37 vs. .35), whereas "unsteady" loaded higher on the second or affective-cognitive factor (.47 vs. .36).

In accordance with prior literature, both factors, the somatic and the affective-cognitive, were highly correlated ($r = .73$). In fact, a comparative analysis with the bi-factor solutions of the BAI found in previous studies that also performed exploratory factor analysis, suggests that the two factors found in the present study correspond reasonably well to the somatic and cognitive-affective factors found in such literature (see the review of bi-factor solutions of Sanz & Navarro, 2003). Thus, for example, we could quantify the degree of convergence between the bi-factor solution found in the present sample of the general Spanish population and the one found by Sanz and Navarro (2003) in university students by calculating Pearson's coefficient

correlation between the factor loadings. The correlation between the factor loadings of the somatic factors of the two samples was .80, and the correlation between the factor loadings of the two affective-cognitive factors was .84, both of them exceeding the standard value of .75 proposed by Cliff (1966) to allow one to state that two factors have a similar interpretation.

To sum up, in the present sample of the Spanish population, the BAI shows two factor solutions with adequate plausibility indexes, a one-factor solution and a two-factor solution, with highly correlated factors. However, these results are not necessarily contradictory, as they could reflect that the BAI measures a general dimension (or common second-order factor) of anxiety that would comprise two symptomatic, highly correlated dimensions (or specific first-order factors), a somatic dimension and an affective-cognitive one. Nevertheless, the issue remains: what are the relative contributions of the general factor and the specific factors in the BAI? Therefore, a principal axis factor analysis based on the correlation between the somatic and affective-cognitive factors was carried out using the SPSS instructions of Wolff and Preising (2005). The Schmid-Leiman (Gorsuch, 1983) transformation of the factor loading matrixes was carried out, both on this second-order factor analysis and on the rotated two-factor first-order solution of the BAI. Thus, we could estimate the relative quantity of independent variance explained by the specific or first-order factors in comparison to the quantity of variance explained by the common or second-order factor, as well as the factor loadings of the BAI items, both on the second-order factor and on each of the first-order factors. These factor loadings are presented in Table 3. The results of the Schmid-Leiman transformation revealed that the common or second-order factor explained 76.2% of the variance, whereas the two first-order factors only accounted for 12.3% and 11.5%, respectively, of the variance. In fact, all 21 items of the BAI presented loadings equal to or higher than .35 on the second-order factor and, for 15 items, these loadings were equal to or higher than .50, whereas only 6 items had loadings equal to or higher than .35 on the first-order factors (3 items on each factor) and none of these 6 items loaded higher on the first-order factor than on the second-order factor. Moreover, for 8 of the 21 BAI items, the loadings on the first-order factors did not exceed .25 (see Table 3).

Therefore, these results support the one-dimensionality of the BAI and suggest that it doesn't make much sense to create two BAI subscales (a somatic subscale and an affective-cognitive one), but to use instead a global score, as hardly any information is lost by not considering the specific variance explained by the somatic and affective-cognitive factors.

Internal Consistency

The analysis of the internal consistency of the BAI yielded an alpha coefficient of .93, which indicates very good internal consistency and replicates the coefficients found in prior literature with similar samples (see Table 4). In fact, De Ayala, Vonderharr-Carlson, and Kim (2005) recently carried out a meta-analysis of reliability coefficients for the BAI and obtained a mean alpha coefficient of .88 from 11 studies carried out with nonpsychiatric samples without university students and a mean alpha coefficient of .89 from 12 studies with university students, both coefficients lower than the one found in the present study.

Table 4

Psychometric Characteristics of the BAI in Diverse Studies with Samples from the General Population

	Sample Characteristics				Psychometric Properties of the BAI		
Study	Country	*N*	% of women	Mean age	*M*	*SD*	α
Nordhagen (2001)	Norway	869	51.1	45.8	5	5.7	0.88
Gillis et al. (1995)	USA	242	51	—	6.6	8.1	—
Jylhä & Isometsä (2006)	Finland	436	51.2	45	6.3	7.7	0.91
Marai (2004)	Fiji	45	57.8	30.3	9.4	7.5	0.94
Osman et al. (1993)	USA	225	70.7	37.1	11.5	10.3	0.92
Freeston et al. (1994)	Canada	474	66	39.1	9.9	10.5	0.92
Yim & Mahalingam (2006)	India	399	50	32.8	10.3	10.9	0.91
Robles et al. (2001, Study 2)	Mexico	1000	58.7	30.4	12.0	9.3	0.83
Robles et al. (2001, Study 3)	Mexico	188	51.9	38.4	9.9	9.3	—
Previous studies [a]	—	3878	56.0	38.1	8.9	8.6	0.88
This study (2007)	Spain	249	52.6	37.8	11.2	10.3	0.93

Note. [a] *N* = sum of participants from previous studies. For the remaining statistics, we present the mean value weighted by the number of participants in each study. - = no data.

The mean of the inter-item correlations of the BAI was .39, with a minimum of .11 and a maximum of .77. The correlations between the scores of each one of the items and the total corrected BAI score (that is, the total score without taking into account the item itself) are presented in Table 5. The correlation coefficients found ranged between .37 for the item "face flushed" and .74 for the item "unsteady," with all of them being statistically significant (with $N = 249$, a correlation coefficient > .21 is statistically significant at $p < .001$) and higher than the minimum of .30 proposed by Nunnally and Bernstein (1994).

Nevertheless, the item "face flushed" presented a corrected item-total correlation coefficient that was lower than that of the remaining items (.37 vs. the range of .49-.74 of the remaining items) and, in fact, it was the only item whose elimination from the BAI did not decrease the alpha coefficient of the instrument's internal consistency (see Table 5). In fact, the item "face flushed" was also the item that presented the lowest loadings on the factor solutions of the BAI (see Table 3). It is hard to determine why the psychometric properties of this item were so much poorer than those of the rest of the items. It is possible that the symptom of face flushed does not form part of the anxiety construct or that the item in its original form does not measure that symptom well. Nonetheless, previous studies with the original version of the BAI report good psychometric indexes for the item "face flushed" (Beck & Steer, 1990; Creamer et al., 1995; Osman, Kopper, Barrios, Osman, & Wade, 1997). Another possibility is that this symptom is specific to anxiety in the USA or in English-speaking countries, but not in other countries. However, although the item "face flushed" in the Norwegian version of the BAI presents corrected item-total correlations of .30 (Nordhagen, 2001), in the French-Canadian version, the indexes of homogeneity and factor validity were adequate (Freeston et al., 1994). This possibility may refer only to Spain or to the Spanish version used in the present study but, in a sample of Spanish university students, this version presented a corrected item-total correlation of .43 and a factor loading of .43 (Sanz & Navarro, 2003). Lastly, it is possible that the expression "face flushed" [*rubor facial*] measures the corresponding symptom well in Spanish university students, but it is an uncommon phrase in people from the general Spanish population with a lower educational level. In this sense, future investigations with samples of the general Spanish population could examine the usefulness of alternative expressions such as, for example, *sonrojarse* or *enrojecimiento de la cara* [Translator's note: both phrases translate to *face flushed* or *getting red in the face*] or add some words to the item to clarify its content such as, for example, *rubor facial* (*ponerse colorado = getting red in the face*).

Table 5

Means, Standard Deviations, Corrected Item-Total Correlations (r_{tot}), Internal Consistency of the Scale if the Item is eliminated (α) and Mantel-Haenszel Statistic (α_{MH}) for the BAI Items

Item	M	SD	r_{tot}	α	α_{MH}
1. Hormigueo o entumecimiento [numbness or tingling]	0.35	0.61	.62	.926	0.68
2. Sensación de calor [feeling hot]	0.75	0.75	.46	.929	1.04
3. Debilidad en las piernas [wobbliness in legs]	0.47	0.76	.59	.926	1.30
4. Incapacidad para relajarme [unable to relax]	0.88	0.88	.60	.926	1.44
5. Miedo a que suceda lo peor [fear of the worst happening]	0.79	0.93	.61	.926	0.89
6. Mareos o vertigos [dizzy or lightheaded]	0.35	0.73	.64	.925	2.41*
7. Palpitaciones o taquicardia [heart pounding or racing]	0.56	0.79	.67	.925	1.31
8. Sensación de inestabilidad [unsteady]	0.51	0.79	.74	.923	0.99
9. Sensación de estar aterrorizado [terrified]	0.35	0.77	.71	.924	0.89
1. Nerviosismo [nervous]	1.26	0.82	.65	.925	1.14
11. Sensación de ahogo [feelings of choking]	0.39	0.77	.67	.925	1.11
12. Temblor de manos [hands trembling]	0.45	0.77	.65	.925	0.84
13. Temblor generalizado [shaky]	0.24	0.61	.72	.924	1.34
14. Miedo a perder el control [fear of losing control]	0.45	0.78	.60	.926	0.48*
15. Dificultad para respirar [difficulty breathing]	0.39	0.73	.68	.925	1.33
16. Miedo a morir [fear of dying]	0.32	0.70	.49	.928	1.18
17. Estar asustado [scared]	0.59	0.80	.64	.925	1.95*
18. Indigestión o molestia abdominal [indigestion or discomfort in abdomen]	0.66	0.84	.49	.928	0.53*
19. Sensación de desmayarse [faint]	0.20	0.55	.50	.928	1.49
2. Rubor facial [face flushed]	0.60	0.75	.37	.930	1.37
21. Sudoración [sweating (not due to heat)]	0.67	0.80	.51	.928	0.81

*Note. N = 249. * Statistically significant at p < .05.*

Discriminant Validity: Relation with Depression and Trait-Anger

In Table 6 are presented the correlation coefficients between the BAI, the BDI-II, and the Trait-Anger scale of the STAXI-2. All these correlations were statistically significant and, according to the conventional values of Cohen (1988) for effect sizes of correlation coefficients, the correlation between the BAI and the BDI-II ($r = .63$) can be considered high ($> .50$), and the correlation between the BAI and the Trait-Anger scale of the STAXI-2 ($r = .32$) can be considered medium ($> .30$). In general, these results replicate the correlations found in former literature. Thus, for example, previous studies with samples from the general population have found correlations between the BAI and the BDI-II that range between .61 (Nordhagen, 2001) and .73 (Jylhä & Isometsä, 2006), whereas, with samples of university students, correlations that range between .56 (Osman et al., 1997; Steer & Clark, 1997) and .62 (Aasen, 2001) have been found, and a correlation of .58 was reported in a sample of Spanish university students (Sanz & Navarro, 2003). Regarding the relations between the BAI and the Trait-Anger scale of the STAXI-2, a recent study carried out with a sample of 324 adults of the general population of ages between 50 and 70 years reported a correlation of .19 (Stewart, Janicki, Muldoon, Sutton-Tyrrell, & Kamarck, 2007), whereas, in a sample of 60 Vietnam veterans, a correlation of .64 was found (Taft, Street, Marshall, Dowdall, & Riggs, 2007).

The correlations found in the present study could, in theory, suggest a lack of discriminant validity of the BAI. Nevertheless, given the overlapping of the definitions of anxiety, depression, and anger, and especially of the first two constructs (Sanz & Navarro, 2003), some correlation among their measures is to be expected, especially between anxiety and depression. Therefore, the question that should be posed is whether the BAI, despite its high or moderate correlation with measurements of depression and anger such as the BDI-II and the STAXI-2, allows us to measure an affective symptomatology different from depression and anger.

To answer this question, we conducted a factor analysis with all the items of the BAI, the BDI-II, and the Trait-Anger scale of the STAXI-2, because both the Kaiser-Meyer-Olkin test of sample adequacy (KMO = .90) and Bartlett's sphericity test ($\chi^2 = 6932.3$, $p < .001$) indicated that the factor model was adequate with the data of our present sample. As we wished to examine whether the BAI measures affective symptoms different from depression and anger, in the factor analysis, we decided to extract three factors by principal axes and to rotate them with *promax*. As can be seen in Table 7, the resulting configuration matrix clearly reflected the existence of a factor of anxiety, another of depression, and a third one of trait-anger, which correlated highly or moderately (see Table 6), but were clearly defined by the items of their corresponding instruments[1]. In effect, considering loadings lower than .25 as negligible, all the items that defined the first factor (anxiety)

Table 6

Correlations between the Instruments and Factors of Anxiety, Depression, and Anger

Factors – Instrument	1. Anxiety – BAI	2. Depression – BDI-II	3. Anger – STAXI-2
1. Anxiety – BAI	—	.63	.32
2. Depression – BDI-II	.61	—	.37
3. Anger – STAXI-2	.34	.41	—

Note. $N = 249$. The upper semi-matrix represents the correlations between instruments. The lower semi-matrix represents the correlations between the three rotated factors that were extracted from the factor analysis of these instruments.

All the correlation coefficients were significant at $p < .001$.

---

[1] Although Velicer's MAP test indicated the extraction of four factors, in order to validate the BAI, the four-factor solution reached with the same extraction procedure (principal axes) and factor rotation (oblimin) was practically the same as the three-factor solution. Specifically, the configuration matrix of the four-factor solution reflected the clear existence of an anxiety factor defined by the BAI items, as 19 of its 21 items presented loadings equal to or higher than .40 on this factor and 16 of these items presented loadings of less than .25 on the remaining three factors. Moreover, in this anxiety factor, no BDI-II or STAXI-2 item had a loading higher than .25. The BDI-II defined a second factor (depression) as 17 of its 21 items had loadings equal to or higher than .40 on this factor, and loadings of less than .25 on the remaining three factors. In fact, no BAI or STAXI-2 item had a loading higher than .25 on the depression factor. The two remaining factors were defined, respectively, by the two 5-item subscales into which the Trait-Anger scale of the STAXI-2 is usually divided (Spielberger et al., 2001): the Angry Temperament subscale, which measures the tendency to experience and express anger without specific provocation, and the Angry Reaction subscale, which assesses the tendency to express anger when one is criticized or treated unfairly by others. Thus, the respective items of each of these subscales showed loadings equal to or higher than .40 on the corresponding factor and, except for one item, loadings lower than .25 on the remaining three factors. Moreover, no BDI-II or BAI item loaded higher than .40 on these two anger factors and, in fact, most of the items from the BDI-II (20 out of 21) and the BAI (18 out of 21) loaded lower than .25 on them.

Table 7

Conjoint Factor Analysis of the Items of the BAI, the BDI-II, and the Trait-Anger Scale of the STAXI-2

Item	Instrument	Factor 1	Factor 2	Factor 3
1. Hormigueo o entumecimiento [numbness or tingling]	BAI	**.73**	—	—
2. Sensación de calor [feeling hot]	BAI	**.42**	—	—
3. Debilidad en las piernas [wobbliness in legs]	BAI	**.53**	—	—
4. Incapacidad para relajarme [unable to relax]	BAI	**.53**	—	—
5. Miedo a que suceda lo peor [fear of the worst happening]	BAI	**.61**	—	—
6. Mareos o vertigos [dizzy or lightheaded]	BAI	**.63**	—	—
7. Palpitaciones o taquicardia [heart pounding or racing]	BAI	**.65**	—	—
8. Sensación de inestabilidad [unsteady]	BAI	**.67**	—	—
9. Sensación de estar aterrorizado [terrified]	BAI	**.75**	—	—
1. Nerviosismo [nervous]	BAI	**.60**	—	—
11. Sensación de ahogo [feelings of choking]	BAI	**.70**	—	—
12. Temblor de manos [hands trembling]	BAI	**.69**	—	—
13. Temblor generalizado [shaky]	BAI	**.84**	—	—
14. Miedo a perder el control [fear of losing control]	BAI	**.63**	—	—
15. Dificultad para respirar [difficulty breathing]	BAI	**.64**	—	—
16. Miedo a morir [fear of dying]	BAI	**.60**	—	—
17. Estar asustado [scared]	BAI	**.71**	—	—
18. Indigestión o molestia abdominal [indigestion or discomfort in abdomen]	BAI	.37	—	—
19. Sensación de desmayarse [faint]	BAI	**.52**	—	—
2. Rubor facial [face flushed]	BAI	.31	—	—
21. Sudoración [sweating (not due to heat)]	BAI	**.52**	—	—
1. Sadness	BDI—II	—	**.57**	—
2. Pessimism	BDI—II	—	**.52**	—
3. Feelings of failure	BDI—II	—	**.71**	—
4. Loss of pleasure	BDI—II	—	**.76**	—
5. Feelings of guilt	BDI—II	—	.35	—
6. Feelings of being punished	BDI—II	—	**.48**	—
7. Dissatisfaction with self	BDI—II	—	**.73**	—
8. Self—blame	BDI—II	—	**.41**	—
9. Thoughts about or desire to commit suicide	BDI—II	—	**.59**	—
10. Crying	BDI—II	—	**.57**	—
11. Agitation	BDI—II	—	—	.28
12. Loss of interest	BDI—II	—	**.64**	—
13. Indecision	BDI—II	—	**.66**	—
14. Useless	BDI—II	—	**.72**	—
15. Loss of energy	BDI—II	—	**.68**	—
16. Changes in sleep pattern	BDI—II	—	**.46**	—
17. Irritability	BDI—II	—	.30	.27
18. Change of appetite	BDI—II	—	**.43**	—
19. Difficulty concentrating	BDI—II	—	**.69**	—
20. Tiredness or fatigue	BDI—II	—	**.62**	—
21. Loss of interest in sex	BDI—II	—	**.61**	—
1. Easily angered	STAXI—2	—	—	**.75**
2. Irritable	STAXI—2	—	—	**.70**
3. Exalted	STAXI—2	—	—	**.59**
4. Feel annoyed by lack of acknowledgement	STAXI—2	—	—	**.53**
5. Tendency to lose control	STAXI—2	—	—	**.63**
6. Feel furious when criticized	STAXI—2	—	—	**.53**
7. Feel furious when undervalued	STAXI—2	—	—	**.55**
8. Easily annoyed	STAXI—2	—	—	**.72**
9. Angry when things don't turn out as planned	STAXI—2	—	—	**.58**
10. Angry when treated unjustly	STAXI—2	—	—	**.59**

Note. $N = 249$. Loadings < .25 are not shown. Loadings ≥ .40 are in boldface.

belonged to the BAI (with loadings ranging between .84 and .31); in fact, 19 out of the 21 BAI items presented loadings over .40 on this first factor (see Table 7). Likewise, all the items that defined the second factor (depression) belonged to the BDI-II (with loadings between .76 and .30), and 18 of the 21 BDI-II items had loadings over .40 on the second factor (see Table 7). Lastly, all the items that defined the third factor (anger) belonged to the Trait-Anger scale of the STAXI-2 (with loadings between .75 and .53), except for two, the items "agitation" and "irritability" from the BDI-II, which also had loadings over .25 in this third factor; These exceptions seem logical as they refer to symptoms shared by both the constructs of anger and depression. Nonetheless, the loadings of these two BDI-II items were lower than .30, whereas all the items of the STAXI-2 presented loadings over .40 in the third factor (see Table 7).

Distribution of the BAI Scores

The total BAI scores ranged between 0 and 57, with a mean of 11.2 (SD = 10.3). These values are similar to those found in many of the prior studies carried out in other countries, both with adults from the general population and with university students. Specifically, the mean of the BAI in the present sample was approximately 2 points higher than the mean found after analyzing conjointly the 9 previous studies with the general population found in the literature (11.2 vs. 8.9, see Table 4), and is only 1.5 points higher than the mean found in previous studies with samples of university students, both Spanish and from other countries (9.6 and 9.7, respectively, Sanz & Navarro, 2003).

A more detailed analysis of the previous results with samples of the general population (see Table 4) suggests that the mean BAI scores with this kind of samples seem to be distributed around the two ranges, so that three studies with samples from Norway, Finland, and the USA found means ranging between 5 and 6.6, whereas six studies with samples from Fiji, Canada, USA, India, and Mexico obtained means ranging between 9.4 and 12, a range in which the Spanish sample of the present study is included. With all due prudence regarding the variations in the composition of the samples and the possible existence of sampling errors, this pattern of results could indicate the presence of genuine differences among countries as far as levels of anxious symptomatology measured by the BAI are concerned, and would be consistent with the international differences in the prevalence of anxiety disorders (The WHO World Mental Health Survey Consortium, 2004). Thus, recent studies indicate that the prevalence of anxiety disorders at 12 months differs significantly even among European countries (European Commission, 2004), and may differ even more if countries from different continents are compared (The WHO World Mental Health Survey Consortium).

As is usual in nonclinical samples, including samples from the general population (Jylhä & Isometsä, 2006;

Nordhagen, 2001; Stewart et al., 2007), the distribution curve of the BAI scores in the sample of the present study was skewed toward higher values (skewness index = 1.8) and peaked at the lowest values (kurtosis = 3.4) so that both the median (8) and the mode (7) were lower than the mean (11.2). Consistently with this, most of the people of the sample obtained low scores in the BAI, so that, using as referents the ranges of the scores proposed by Beck and Steer (1990) to distinguish different levels of severity of anxious symptomatology, 57% of the sample obtained scores between 0 and 9 ("normal anxiety"), 24.1% between 10 and 18 ("mild anxiety"), 12.9% between 19 and 29 ("moderate anxiety"), and only 6% scored between 30 and 63 ("severe anxiety").

As can be seen in Table 4, the mean score of the items was 0.53 (range = 0.20 – 1.26), with the following items receiving higher scores: "nervous," "unable to relax," "fear of the worst happening," and "feeling hot." In contrast, the symptoms that were present in their mildest form were "faint," "shaky," and "fear of dying." These results are partially similar to those found in other samples of the general population and in samples of university students. For example, both in adults of the general population from Norway (Nordhagen, 2001) or Canada (Freeston et al., 1994), and in university students from the USA (Osman et al., 1997), Australia (Creamer et al., 1995) or Spain (Sanz & Navarro, 2003), the three items with the highest intensity coincide, even in their order in the case of the Canadian sample and the three university samples, with those found in the sample of the general Spanish population: "nervous," "unable to relax," and "fear of the worst happening." Moreover, in the Canadian sample of the general population, the three items with the lowest score also coincided with those found in the present sample; in fact, in the three previous studies with university students, it was found that one of the three items with the lowest score was "fear of dying" and, in the sample of Spanish university students, also "shaky."

Demographic Differences and Norms for the General Spanish Population

The correlations of the BAI scores with age and educational level were almost null and statistically nonsignificant (r = .03 and –.06, both ns). However, a statistically significant bi-serial point correlation was found between sex, coded as 0 (male) and 1 (female), and the BAI scores (r = .13, p < .035), although the magnitude of this correlation was small. In fact, an independent measures t-test revealed that women scored significantly higher than men (12.5 vs. 9.8), $t(247)$ = –2.12, p < .035, although this difference was only 2.7 points, that is, 0.26 standard deviation units in terms of the d statistic of effect size, a difference that may be considered small in terms of Cohen's (1988) conventional values for effect sizes. This higher score of women compared to men is consistent with the results obtained in previous studies carried out both with adults from the general population (Freeston et al., 1994; Osman et al.,

1993) and with psychiatric patients (Hewitt & Norton, 1993), or university students (Borden, Peterson, & Jackson, 1991; Creamer et al., 1995; Nitschke, Heller, Imig, McDonald, & Miller, 2001), including Spanish university students (Sanz & Navarro, 2003).

Nonetheless, in order for these sex differences in the BAI to be established as genuine, it would have to be determined that there is no differential functioning of any of its items in that direction, that is, we would have to verify that there are no items in which there are differences between men and women despite the fact that groups of men and women with the same level of anxiety are being compared. As in previous research (i.e., Hidalgo Montesinos, Galindo Garre, Inglés Saura, Campoy Menéndez, & Ortiz Soria, 1999), in order to verify the differential functioning of the items, we applied the Mantel-Haenszel method, for which we dichotomized the score of each item as a function of the median, and we used as the comparison variable the total BAI score, divided into five levels of anxiety so that each level would have a similar percentage of people. In this context, the method allows one to calculate a statistic, the Mantel-Haenszel (α_{MH}) statistic, which compares the number of men and women who score higher and lower than the median in each item for each one of the five levels of anxiety.

In Table 5 are displayed the values of the α_{MH} for each one of the items. Only in 4 of the 21 BAI items, was the value of the α_{MH} statistic significant at $p < .05$ (see Table 5). Specifically, the items "dizzy or lightheaded," "scared," "fear of losing control," and "indigestion or discomfort in abdomen" were suspected of differential functioning, because the response to them depended both on the level of anxiety and on sex. Nevertheless, when calculating the total BAI

score without using these four items and again comparing the men and women from the present sample with an independent measures t-test, the results revealed that the women scored statistically significantly higher than the men in the BAI (10.3 vs. 7.9), $t(247) = -2.23$, $p < .026$, and that this difference was small ($d = .28$).

Concluding, given that we found significant sex differences in the BAI and we did not find any significant relation of the BAI with age or educational level, in Table 8 are presented the normative scores (in percentiles) obtained both from the total sample of adults from the general population and, separately, from the subsamples of men and women.

Clinically Significant Recovery or Improvement Criteria Based on Norms of the General Population (Normative Comparisons)

Defining recovery from anxiety as a score equal to or lower than the mean of the norms from a sample of the general population (Hollon & Flick, 1988) or as a score equal to or lower than one standard deviation above the mean (Kendall & Grove, 1988), and taking into account the descriptive statistics obtained with the present sample, we could estimate that the clinically significant recovery or improvement criterion for Spanish patients with anxious symptomatology could either be a score equal to or lower than 11 (mean) or a score equal to or lower than 21 ($M + SD$). Which of the two criteria is the most appropriate?

In principle, the cut-off points that define the different categories of severity of anxiety proposed by Beck and Steer (1990) suggest that a score of 21 as a recovery criterion is too high because, according to this proposal, the presence of

Table 8

Centile Scores of the BAI for the Total Sample of the General Population and for the Subsamples of Males and Females

Centiles	Total $N = 249$	Males $n = 118$	Females $n = 131$
1	0	0	0
5	0	0	1
10	2	1	2
20	4	3	5
25	5	4	5
30	5	5	6
40	7	6	7
50	8	7	9
60	10	9	11
70	12	11	14
75	14	12	17
80	17	14	22
90	26	22	28
95	32	29	40
99	49	54	50
M	11.2	9.8	12.5
SD	10.3	9.3	11.1

mild anxiety symptomatology is observed at a score over 9, and the presence of moderate anxiety symptomatology at a score over 18. In fact, in the second edition of the original manual of the BAI, Beck and Steer (1993) propose even lower cut-off points to define the different categories of severity of anxiety (0-7 = "normal," 8-15 = "mild," 16-25 = "moderate," and 26-63 = "severe"), which confirms the idea that a score of 21 is too high to be used as a recovery criterion. In contrast, the experience accumulated with the BAI when assessing clinically significant therapeutic results also seems to favor 11 (the mean) instead of 21 ($M + SD$) as the most adequate criterion for the BAI, because in order to define an improvement or clinically significant improvement with this instrument, many studies have adopted the cut-off points proposed by Beck and Steer (1993) or Beck and Steer (1990) to delimit mild anxious symptomatology from normality (8 or 10, respectively; i.e., Butler, Fennell, Robson, & Gelder, 1991; McEvoy & Nathan, 2007; Westbrook & Kirk, 2005).

To conclude, and tentatively, we propose a score equal to or lower than 11 in the BAI (the mean of the general Spanish population) to estimate that a Spanish adult patient has recovered from his or her anxious symptomatology or has improved in a clinically significant way.

Conclusions

The BAI is an instrument designed to measure clinical anxiety and discriminate this construct from depression (Beck et al., 1988). The main goal of this study was to analyze the psychometric properties in a sample of the general Spanish population of the Spanish version of the BAI elaborated by Sanz and Navarro (2003), a version that had already showed satisfactory psychometric properties in a sample of Spanish university students. A secondary goal of this study was to establish criteria to assess the clinical significance of the results obtained with treatments for anxiety, obtaining for this purpose reference scores in a sample of the general population with which to perform *normative comparisons.* The results obtained allow us to conclude the following:

1. The BAI allows the establishment of individual differences among people from the general Spanish population insofar as the presence and intensity of anxiety symptoms is concerned, although the *distribution of the BAI scores* in the general population does not match a normal curve, but instead, most people obtain low scores This distribution is similar to the one found in previous studies carried out in other countries and, in fact, the Spanish sample of this study presented a mean BAI score similar to the one found in a large part of the studies carried out with adults from the general population in other countries (i.e., Mexico, Canada, India). However, we also found studies with samples from the general population of Finland or Norway that obtained notably lower scores in the BAI and even two studies from the USA that differed in mean score obtained in the BAI, in one case, coming close to the means found in Mexico, Canada, or India and, in the other case, to the means found in Finland or Norway. Given the small number of studies carried out to date, it is not possible to determine whether these differences between countries are due to variations in the composition of the samples, sampling errors, or cultural differences, but it would certainly be appropriate for future research to examine this last possibility, because this would be consistent with the international differences in the prevalence of anxiety disorders. Nevertheless, it is important to note that, independently of differences in the global BAI score, the symptoms or items of the BAI that obtained the highest scores in the Spanish sample were the same as those of most of the samples of other countries.

2. Among the adults of the general Spanish population, the women scored higher in the BAI than the men. In contrast, we did not find any relation of the BAI scores with age or the educational level. Although the magnitude of the difference between men and women in the BAI was small, this result replicates the *sex differences* found in the prior literature with the BAI, it is coherent with the higher prevalence of anxiety disorders among women (i.e., Kessler et al., 1994), and cannot be explained by the presence in the BAI of items with a differential functioning according to sex, because only 4 of the 21 items presented signs of differential functioning, and the sex differences appeared even when these 4 items were eliminated to calculate the global BAI score.

3. Reliability in terms of internal consistency of the Spanish version of the BAI was high and similar to the levels found in other countries. Therefore, it is not surprising that all the items of the BAI presented acceptable homogeneity indexes. However, the item "face flushed" showed lower homogeneity and factor validity indexes than those of the remaining BAI items, so that future research with the Spanish version of the BAI should examine in detail the psychometric behavior of this item, in particular the possibility that the expression used to name the corresponding symptom is uncommon among people with a lower educational level and, therefore, consider the need to use alternative expressions or to clarify its content.

4. The results of the factor analysis carried out with the Spanish version of the BAI indicate that, in adult samples from the general Spanish population, this instrument seems to measure a dimension of general anxiety that comprises two specific symptomatic dimensions that are highly related, a somatic dimension and an affective-cognitive dimension, and it replicates the results found in the previous literature and,

specifically, the results found in Spanish university students. However, the two specific dimensions barely explain the variance beyond that accounted for by the general anxiety dimension, so it can be concluded that, in samples from the general Spanish population, the BAI is configured as a one-dimensional instrument in which it doesn't make much sense to create subscales to measure the specific dimensions because not much information is lost when considering only the global score.

5. The BAI correlates highly with the BDI-II and moderately with the Trait-Anger scale of the STAXI-2, but the factor analyses carried out indicate that the items from all three scales are distinguishable from each other and suggest that the correlations among the three instruments may be due more to the relation between the constructs of anxiety, depression, and anger than to a problem of lack of *discriminant validity* of the BAI.

6. As the BAI seems to reliably and validly measure anxious symptomatology in adults from the general Spanish population, it is possible to propose a score equal to or less than 11 in this instrument (the mean of the general Spanish population) as a criterion to assess whether an adult patient, after receiving a treatment for anxiety, is any different from normal people regarding their symptoms and main complaints, that is, whether the patient presents *clinically significant recovery or improvement.* Nonetheless, it is important to note the tentative nature of this proposal and the need to replicate this mean in future studies with samples from the general Spanish population and, especially, to base the validity of this proposal on studies that compare, for example, patients with anxiety disorders who, after treatment, score higher than or lower than this criterion with regard to the appraisal of symptomatological improvement provided by clinicians, or with regard to the appraisal of quality of life made by people from their environment, or referring to their long-term relapse indexes. Likewise, its validity would receive more support if new studies of samples from the general population compared the level of quality of life or the psychosocial functioning of people who score higher or lower than this mean.

7. The above conclusions about psychometric properties of the BAI in the general Spanish population should be considered taking into account the limitations of the present study, the most important of which refers to the possibility of generalizing the results due to the sampling system used (circumstantial, with a "captive" population of relatives and friends of university students), to the procedure of data collecting (by psychology university students as part of their academic practice), and to the reduced number of participants assessed. Therefore, it seems necessary that future research should replicate the results obtained in this study and, if possible, the replication should be made with a larger sample size and a random sampling procedure.

8. To sum up, with the above-mentioned caveats, it can be concluded that the Spanish version of the BAI seems to have acceptable psychometric properties as an assessment instrument for anxious symptomatology in adults from the general Spanish population that make it recommendable to use in this kind of population and for this purpose. However, the data of this study do not justify the use of the BAI as a diagnostic instrument, that is, to make a diagnostic judgment about whether or not a person from the general population suffers from an anxiety disorder. Doing this would lead to confounding the levels of analysis: symptom-syndrome-disorder. The BAI is adequate to identify anxiety symptoms and to quantify their intensity, which is obviously very important to assess, for example, a person's general status, to identify people with symptoms or syndromes of anxiety, or to assess therapeutic progress. However, the diagnosis of an anxiety disorder is effected taking into account not only the type and number of symptoms present, but also certain duration criteria (i.e., in the generalized anxiety disorder, anxiety symptoms and excessive concern should last for more than 6 months), severity, (i.e., in posttraumatic stress disorder, there should be at least 6 symptoms and, as in all anxiety disorders, the symptoms should cause clinically significant distress), course (i.e., in acute stress disorder, the symptoms should appear in the first month after a traumatic event), impairment (i.e., the symptoms should cause social, academic, or labor deterioration), and absence of certain possible causes or of certain concurrent diagnoses (i.e., the symptoms should not be due to the direct physiological effects of drugs of pharmaceuticals or to a medical illness and should not appear exclusively during a psychotic disorder).

Nevertheless, the BAI could be a good screening instrument to detect people with an anxiety disorder in samples of the general population (and whose possible diagnosis should be subsequently confirmed by means of some kind of diagnostic interview), until future studies expressly assess its psychometric properties for this purpose (sensitivity, specificity, degree of agreement with a well established diagnosis, etc.) and determine empirically the most appropriate cut-off score in the general Spanish population to identify each of the anxiety disorders (i.e., generalized anxiety disorder, panic disorder, etc.). In this sense, the analysis of the content validity of the BAI carried out by Sanz and Navarro (2003) indicated that, although the instrument seemed to tap 45% of the symptoms included as criteria of the anxiety disorders in the *Diagnostic and Statistical Manual of Mental Disorders* (4th edition, American Psychiatric Association, 1994) and that are

specific compared to depressive disorders, many other symptoms that define clinical anxiety and distinguish this construct from depression are not tapped by the BAI, and that these deficiencies do not affect all anxiety disorders equally because, for example, in the BAI, there is a predominance of items that explore physiological symptoms of anxiety in detriment of motor and cognitive symptoms. This suggests that, as a screening instrument, the BAI may function better in anxiety disorders with a high physiological component such as panic disorder, and worse in disorders with a stronger motor or cognitive component such as, for example, social phobia or obsessive-compulsive disorder.

References

Aasen, H. (2001). *An empirical investigation of depression symptoms: Norms, psychometric characteristics and factor structure of the Beck Depression Inventory-II*. University of Bergen. Retrieved February 14, 2003, from: http://www.ub.uib.no/elpub/2001/h/308003/Hovedoppgave.pdf.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.).Washington, DC: Author.

Ayers, C. R., Sorrell, J. T., Thorp, S. R., & Wetherell, J. L. (2007). Evidence-based psychological treatments for late-life anxiety. *Psychology and Aging, 22*, 8-17.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56*, 893-897.

Beck, A. T., & Steer, R. (1990). *Beck Anxiety Inventory manual*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., & Steer, R. (1993). *Beck Anxiety Inventory manual*. San Antonio, TX: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI-II. Beck Depression Inventory-Second Edition manual*. San Antonio, TX: The Psychological Corporation.

Borden, J. W., Peterson, D. R., & Jackson, E. A. (1991). The Beck Anxiety Inventory in nonclinical samples: Initial psychometric properties. *Journal of Psychopathology and Behavioral Assessment, 13*, 345-356.

Botella, C., & Ballester, R. (1997). *Trastorno de pánico: Evaluación y tratamiento*. Barcelona: Martínez Roca.

Butler, G., Fennell, M., Robson, P., & Gelder, M. (1991). Comparison of behavior therapy and cognitive behavior therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 59*, 167-175.

Chambless, D. L., & Gillis, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology, 61*, 248-260.

Cliff, J. (1966). Orthogonal rotation to congruence. *Psychometrika, 31*, 33-42.

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

Comeche, M. I., Díaz, M., & Vallejo, M. (1995). *Cuestionarios, inventarios y escalas. Ansiedad, depresión y habilidades sociales*. Madrid: Fundación Universidad-Empresa.

Creamer, M., Foran, J., & Bell, R. (1995). The Beck Anxiety Inventory in a non-clinical sample. *Behaviour Research and Therapy, 33*, 477-485.

De Ayala, R. J., Vonderharr-Carlson, D. J., & Kim, D. (2005). Assessing the reliability of the Beck Anxiety Inventory scores. *Educational and Psychological Measurement, 65*, 742-756.

Echeburúa, E. (1993). *Ansiedad crónica. Evaluación y tratamiento*. Madrid: Eudema.

European Commission (2004). *The state of mental health in the European Union*. Brussels: Author.

Freeston, M. H., Ladouceur, R., Thibodeau, N., Gagnon, F., & Rhéaume, J. (1994). L'inventaire d'anxiété de Beck: Propriétés psychométriques d'une traduction française. *L'Encéphale, 20*, 47-55.

Gillis, M. M., Haaga, D. A. F., & Ford, G. T. (1995). Normative values for the Beck Anxiety Inventory, Fear Questionnaire, Penn State Worry Questionnaire, and Social Phobia Anxiety Inventory. *Psychological Assessment, 7*, 450-455.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Hewitt, P. L., & Norton, G. R. (1993). The Beck Anxiety Inventory: A psychometric analysis. *Psychological Assessment, 5*, 408-412.

Hidalgo Montesinos, M. D., Galindo Garre, F., Inglés Saura, C. J., Campoy Menéndez G., & Ortiz Soria, B. (1999). Estudio del funcionamiento diferencial de los ítems en una Escala de Habilidades Sociales para Adolescentes. *Anales de Psicología, 15*, 333-345.

Hollon, S. D., & Flick, S. N. (1988). On the meaning and methods of clinical significance. *Behavioral Assessment, 10*, 197-206.

Instituto Nacional de Estadística (2004). *Censos de población y viviendas 2001. Resultados definitivos*. Retrieved July 25, 2005, from: http://www.ine.es.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.

Jylhä, P., & Isometsä, E. (2006). The relationship of neuroticism and extraversion to symptoms of anxiety and depression in the general population. *Depression and Anxiety, 23*, 281-289.

Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). Boston: Allyn & Bacon.

Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. *Behavioral Assessment, 10*, 147-158.

Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67*, 285-299.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51*, 8-19.

Magán, I., Sanz, J., & García-Vera, M. P. (2008a). *Development and validation of a measure of thoughts specific to anger and hostility: the Angry/Hostile Thoughts Inventory*. Manuscript submitted for publication.

Magán, I., Sanz, J., & García-Vera, M. P. (2008b). *The Angry/Hostile Attitudes Inventory: a new tool to measure dysfunctional attitudes specifically associated with anger and hostility*. Manuscript submitted for publication.

Marai, L. (2004). Anxiety and hopelessness in two South Pacific countries: Exploratory studies. *Social Behavior and Personality: An International Journal, 32*, 723-730.

McEvoy, P. M., & Nathan, P. (2007). Effectiveness of cognitive behavior therapy for diagnostically heterogeneous groups: a benchmarking study. *Journal of Consulting and Clinical Psychology, 75*, 344-350.

Nitschke, J. B., Heller, W., Imig, J. C., McDonald, R. P., & Miller, G.A. (2001). Distinguishing dimensions of anxiety and depression. *Cognitive Therapy and Research, 25*, 1-22.

Nordhagen, T. (2001). *Beck Anxiety Inventory: translation and validation of a Norwegian version*. Master's thesis. University of Bergen. Retrieved October 11, 2005. from http://www.ub.uib.no/elpub/2001/h/308004/Hovedoppgave.pdf.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill [Spanish translation: Teoría psicométrica. Mexico: McGraw-Hill, 1995].

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, and Computers, 32*, 396-402.

Osman, A., Barrios, F. X., Aukes, D., Osman, J. R., & Markway, K. (1993). The Beck Anxiety Inventory: Psychometric properties in a community population. *Journal of Psychopathology and Behavioral Assessment, 15*, 287-297.

Osman, A., Kopper, B. A., Barrios, F. X., Osman, J. R., & Wade, T. (1997). The Beck Anxiety Inventory: Reexamination of factor structure and psychometric properties. *Journal of Clinical Psychology, 53*, 7-14.

Piotrowski, C. (1999). The status of the Beck Anxiety Inventory in contemporary research. *Psychological Reports, 85*, 261-262.

Robles, R., Varela, R., Jurado, S. & Páez, F. (2001). Versión mexicana del Inventario de Ansiedad de Beck: propiedades psicométricas. *Revista Mexicana de Psicología, 18*, 211-218.

Sanz, J., García-Vera, M. P., Espinosa, R., Fortún, M., & Vázquez, C. (2005). Adaptación española del Inventario para la Depresión de BeckCII (BDI-II): 3. Propiedades psicométricas en pacientes con trastornos psicológicos. *Clínica y Salud, 16*, 121-142.

Sanz, J., Magán, I., & García-Vera, M. P. (2006). Personalidad y el síndrome AHI (agresión-hostilidad-ira): relación de los Cinco Grandes con ira y hostilidad. *Psicopatología Clínica, Legal y Forense, 6*, 153-176.

Sanz, J., & 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.

Sanz, J., Navarro, M. E., & Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de Beck-II (BDI-II): 1. Propiedades psicométricas en estudiantes universitarios. *Análisis y Modificación de Conducta, 29*, 239-288.

Sanz, J., Perdigón, L. A., & Vázquez, C. (2003). Adaptación española del Inventario para la Depresión de BeckCII (BDI-II): 2. Propiedades psicométricas en población general. *Clínica y Salud, 14*, 249-280.

Siev, J., & Chambless, D. L. (2007). Specificity of treatment effects: Cognitive therapy and relaxation for generalized anxiety and panic disorders. *Journal of Consulting and Clinical Psychology, 75*, 513-522.

Spielberger, C. D. (1999). *Manual for the State-Trait Anger Expression Inventory, STAXI-2*. Odessa, FL: Psychological Assessment Resources.

Spielberger, C. D., Miguel-Tobal, J. J., Casado, M. I., & Cano-Vindel, A. (2001). *Inventario de Expresión de Ira Estado-Rasgo: STAXI-2*. Madrid: TEA.

Steer, R. A., & Clark, D. A. (1997). Psychometric characteristics of the Beck Depression Inventory-II with college students. *Measurement and Evaluation in Counseling and Development, 30*, 128-136.

Steer, R. A., Rissmiller, D. J., Ranieri, W. F., & Beck, A. T. (1993). Structure of the computer-assisted Beck Anxiety Inventory with psychiatric inpatients. *Journal of Personality Assessment, 60*, 532-542.

Stewart, J. C., Janicki, D. L., Muldoon, M. F., Sutton-Tyrrell, K., & Kamarck, T. W. (2007). Negative emotions and 3-year progression of subclinical atherosclerosis. *Archives of General Psychiatry, 64*, 225-233.

Taft, C. T., Street, A. E., Marshall, A. D., Dowdall, D. J., & Riggs, D. S. (2007). Posttraumatic stress disorder, anger, and partner abuse among Vietnam combat veterans. *Journal of Family Psychology, 21*, 270-277.

The WHO World Mental Health Survey Consortium (2004). Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA, 291*, 2581-2590.

Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy, 43*, 1243-1261.

Wolff, H., & Preising, K. (2005). Exploring item and higher order factor structure with the Schmid–Leiman solution: Syntax codes for SPSS and SAS. *Behavior Research Methods, 37*, 48-58.

Yim, J. Y., & Mahalingam, R. (2006). Culture, masculinity, and psychological well-being in Punjab, India. *Sex Roles, 55*, 715-724.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 3*, 432-442.