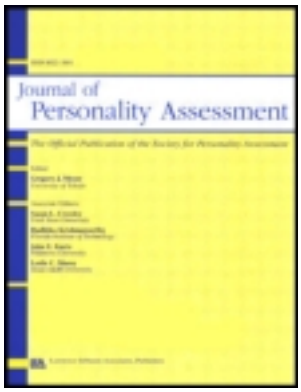

Journal of Personality Assessment
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hjpa20

The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety?

Arturo Bados [a], Juana Gómez-Benito [b c] & Gemma Balaguer [a]

[a] Department of Personality, Assessment and Psychological Treatment, University of Barcelona, Spain

[b] Department of Methodology of the Behavioral Sciences, University of Barcelona, Spain

[c] Institute for Brain, Cognition and Behavior, University of Barcelona, Barcelona, Spain

Version of record first published: 15 Oct 2010.

To cite this article: Arturo Bados, Juana Gómez-Benito & Gemma Balaguer (2010): The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety?, Journal of Personality Assessment, 92:6, 560-567

To link to this article: http://dx.doi.org/10.1080/00223891.2010.513295

Journal of Personality Assessment, *92*(6), 560–567, 2010

ISSN: 0022-3891 print / 1532-7752 online
DOI: 10.1080/00223891.2010.513295

Routledge
Taylor & Francis Group

ARTICLES

The State-Trait Anxiety Inventory, Trait Version: Does It Really Measure Anxiety?

ARTURO BADOS,[1] JUANA GÓMEZ-BENITO,[2,3] AND GEMMA BALAGUER[1]

[1]*Department of Personality, Assessment and Psychological Treatment, University of Barcelona, Spain*
[2]*Department of Methodology of the Behavioral Sciences, University of Barcelona, Spain*
[3]*Institute for Brain, Cognition and Behavior, University of Barcelona, Barcelona, Spain*

To clarify what is actually measured by the trait version of the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970), we conducted a confirmatory factor analysis of various models and evaluated convergent and discriminant validity. The best fit was obtained with both a bifactor model, comprising 2 specific factors plus a general factor, and a 1-construct, 2-method model. The total score and the 2 method subscales of the STAI trait version were more strongly correlated with depression than with anxiety. In the bifactor model with 2 specific factors, the depression subscale showed stronger correlations with measures of depression than with measures of anxiety. The correlation of the hypothetical anxiety subscale with measures of depression was equivalent to or higher than its correlation with measures of anxiety. These results suggest that the questionnaire does not strictly evaluate anxiety but, rather, negative affect.

The State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) is a questionnaire comprising two, 20-item scales designed to measure state and trait anxiety. After a decade of research with the initial version (Form X) of the STAI, new items were drawn up with the aim of developing a supposedly purer measure of state and trait anxiety. The result was Form Y of the STAI (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983).

Both forms of the STAI are widely used measures of anxiety. Although total scores for state and trait anxiety are usually used, factor analysis of the STAI has revealed four correlated factors: presence and absence of both state and trait anxiety (Ponciano et al., 2006; Spielberger, Sydeman, Owen, & Marsh, 1999; Vigneau & Cormier, 2009). The items that refer to the absence of anxiety (negative polarity) are those written in a form that is the opposite of what the scale seeks to measure (e.g., "I feel secure"), and they account for nine of the 20 items (1, 3, 6, 7, 10, 13, 14, 16, 19) in Form Y of the trait version (STAI–T), which measures a person's disposition to respond with anxiety when faced with situations perceived as threatening. Thus, the factors regarding the presence and absence of anxiety correspond to the polarity or direction of the items in relation to the construct measured by the scale.

Bieling, Antony, and Swinson (1998) considered whether the Trait-Anxiety Absent factor was replicable in other samples and whether the items referring to the absence of anxiety were really related to anxiety rather than being possible indicators of depression. An exploratory factor analysis using Form Y of the STAI–T and conducted with North American anxiety disorder patients revealed two factors: Depression and Anxiety. A subsequent confirmatory analysis comparing three models (i.e., one factor [General Anxiety], two uncorrelated factors [Anxiety and Depression], and two factors [Anxiety and Depression] loading on a higher order factor [Negative Affect]) showed that the third model offered the best fit, although the first two of the three criteria used (goodness-of-fit index (GFI) = .84; adjusted goodness-of-fit index = .80; root mean square residual = .06) did not reach the recommended cutoff point (see Clara, Cox, & Enns, 2001; Hu & Bentler, 1999). An examination of the items and of the correlations between the two identified subscales and other measures of anxiety and depression supported the idea that the STAI–T assesses depression in addition to anxiety and, in general, negative affect.

Bieling et al. (1998) correlated the STAI–T with the Beck Depression Inventory (BDI; Beck, Rush, Shaw, & Emery, 1979), the Beck Anxiety Inventory (BAI; Beck & Steer, 1993), and the Depression, Anxiety and Stress Scales, 21-item version (DASS–21; Lovibond & Lovibond, 1995), and they found that the convergent validity of the subscales for Depression (13 items) and Anxiety (7 items) was good. Although Bieling et al. stated that the discriminant validity was also good, we believe that the discriminant validity was acceptable for the Depression subscale but not for the Anxiety subscale. Indeed, the Anxiety subscale of the STAI was more strongly correlated with the BDI ($r = .66$) than with the BAI ($r = .50$) and showed an equivalent correlation with the Depression and Anxiety scales of the DASS–21 ($r = .53$ and .55, respectively).

Bieling et al. (1998) also found that the STAI–T total score was more strongly correlated with measures of depression than with those of anxiety. This finding was replicated by Ponciano

Received January 13, 2009; Revised April 14, 2010.
Address correspondence to Arturo Bados, Department of Personality, Assessment and Psychological Treatment, Faculty of Psychology, University of Barcelona, Paseo Vall d'Hebron, 171, 08035 Barcelona, Spain; Email: abados@ub.edu

et al. (2006) with secondary school pupils and undergraduates as well as by Grös, Antony, Simms, and McCabe (2007) with a clinical sample (85% with a primary diagnosis of anxiety disorder). Likewise, a study of undergraduates by Kohn, Kantor, DeCicco, and Beck (2008) found that the STAI–T correlated similarly with anxiety and depression. In contrast, Caci, Baylé, Dossios, Robert, and Boyer (2003) reported a higher correlation with anxiety in an undergraduate sample. In a study by Fountoulakis et al. (2006), the STAI–T showed equivalent correlations with measures of anxiety, depression, and neuroticism in a sample comprising both healthy and depressed young adults. These data cast doubt on whether the STAI–T only evaluates anxiety.

Caci et al. (2003) conducted another confirmatory factor analysis of the STAI–T with French undergraduates. In line with Bieling et al. (1998), Caci et al. found that the models based on one factor and two uncorrelated factors did not fit the data. Subsequently, they examined the item wording to form three subscales: Anxiety (10 items), Unsuccessfulness (4 items), and Happiness (6 items). They also distinguished three clusters within the Anxiety subscale: restlessness (2 items), worrying (4 items) and self-confidence (4 items). When the three models (anxiety, unsuccessfulness, and happiness) were analyzed separately, the fit was good, but this was not the case for the model that considered jointly the 20 items of the STAI–T distributed across the three previously mentioned subscales.

Vigneau and Cormier (2008) pointed out that a model in which construct and method factors are distinguished may offer an adequate account of the factor structure of the STAI–T. This model combines (a) a one-factor model in which it is assumed that all the items load on the construct that the scale seeks to measure (anxiety) and (b) a polarity-based, two-factor model in which it is assumed that the items load on one of two method factors defined by the item polarity. Keeping the method factors orthogonal to one another and to the construct factor, Vigneau and Cormier (2008) found that the construct-method model offered the best fit to the data, although the first two of the three criteria used in three samples (GFI = .82–85; comparative fit index [CFI] = .84–88; root mean square error of approximation [RMSEA] = .06) did not reach the recommended cutoff point. Furthermore, the analysis was performed with the two STAI scales (Trait and State) in combination; and thus, the model tested was actually a two-construct, two-method model. The study by Vautier and Pohl (2009), which also used both STAI scales but with a different analytic methodology, provided support for the notion that the STAI–T measures a substantive construct together with method effects due to item polarity. Specifically, Vautier and Pohl took a test–retest perspective and found that their structural equation model yielded a good fit when they only specified one latent change variable, considering the temporal change measured on the positive and negative polarity scales of the STAI–T. Vautier and Pohl argued that the model would not have shown a good fit if the two scales had measured two different constructs.

In summary, there are three models of STAI–T factor structure that offer a relatively satisfactory—but not good—fit to the data: two correlated method factors (Spielberger et al., 1999), two factors (Anxiety and Depression) loading on a higher order factor (Negative Affect) (Bieling et al., 1998), and a one-construct, two-method model (Vigneau & Cormier, 2008). As these models have yet to be compared, the first objective of this study was to determine which of them provides the best fit to the data. In addition, we considered a new model: a bifactor model with two specific factors (Anxiety and Depression) and one general factor (Negative Affect). Unlike the model based on two factors plus a higher order factor (in which each item loads on one of two specific factors, which in turn load on a second-order factor), each item in the proposed bifactor model loads on both a specific factor and a general factor, all of which are first order. This model has a number of advantages over the second-order model (Chen, West, & Sousa, 2006), particularly when it comes to analyzing the measurement invariance of the instrument between different samples and when the interest lies in the contribution of the specific factors above and beyond the general/second-order factor.

Given the overall aim of determining what is actually being measured by the STAI–T, the second objective was to study the convergent and discriminant validity of the instrument and its possible subscales. If the STAI–T has good convergent validity, then its total score, its two method subscales, and the Anxiety subscale should show strong correlations with other measures of anxiety. Likewise, if it has good discriminant validity, then the total score and the two method subscales should be more strongly correlated with measures of anxiety than with measures of depression. Additionally, the Depression subscale of the STAI–T should show higher correlations with other measures of depression than with measures of anxiety; and the Anxiety subscale of the STAI–T should show stronger correlations with other measures of anxiety than with measures of depression.

Method

Participants

A total of 346 students from the Faculty of Psychology of the University of Barcelona (Spain) took part in this study; 13 of these failed to respond to one or more items on the STAI, and so the final sample comprised 333 students. Of these, 274 were women and 59 were men. They were all in their 3rd year of studies and were aged between 20 and 25 years. Regarding marital status, 91.6% were single, 7.8% were married or living with a partner, and 0.6% were separated or divorced. In terms of employment, 49.2% had no work apart from their studies, 40.6% worked an average of fewer than 6 hr a day, and 10.2% worked an average of 6 or more hr per day.

Measures

STAI–T (Spielberger et al., 1983). This part of the questionnaire aims to evaluate trait anxiety, that is, the disposition to respond with anxiety to situations perceived as threatening. It comprises 20 items (11 direct worded and 9 reverse worded), and respondents answer according to how they generally feel; responses are given on a frequency scale ranging from 1 (*almost never*) to 4 (*almost always*).

DASS–21 (Lovibond & Lovibond, 1995; Spanish version: Bados, Solanas, & Andrés, 2005). Respondents evaluate ranging from 0 to 3 the intensity/frequency with which they have experienced each of the 21 negative emotional symptoms over the previous week. These symptoms are grouped into three scales (Depression, Anxiety, and Stress), of which the first two we used here. In this sample, the internal consistency of these

two scales was .80 and .65 for Depression and Anxiety, respectively, whereas the correlation between them was .40.

Symptom Checklist–90–R (SCL–90-R; Derogatis, 1983; Spanish version: Derogatis, 2002). This questionnaire, which comprises 90 items scored on a scale of distress ranging from 0 (*not at all*) to 4 (*extremely*), has frequently been used to evaluate aspects of psychopathology. Although we administered the whole questionnaire, we only analyzed two of its nine subscales, those for Depression and Anxiety. In this sample, the internal consistency of these two scales was .88 and .81, respectively, whereas the correlation between them was .64.

BDI (Beck et al., 1979; Spanish version: Sanz & Vázquez, 1998). This inventory comprises 21 items, each of which includes four brief statements that correspond to normal responses and to mild, moderate, and severe depressive symptoms. Respondents choose one statement per item according to their feelings over the previous week. Each item is scored ranging from 0 to 3. The internal consistency in our sample was .81.

BAI (Beck & Steer, 1993; Spanish version: Sanz & Navarro, 2003). This is a specific questionnaire for measuring anxiety that was designed to minimize the presence of those symptoms related to depression. It comprises 21 items or symptoms that respondents score ranging from 0 (*not at all*) to 3 (*severely—I could barely stand it*) according to the degree of distress these symptoms have caused over the previous 7 days. The internal consistency in our sample was .85, whereas its correlation with the BDI was .65.

With the exception of the STAI–T, all the items in the other questionnaires used have a positive polarity. Of these instruments, we specifically chose the DASS–21, the BDI, and the BAI because they were originally designed to differentiate more clearly between anxiety and depression.

Procedure

To date, only Form X of the STAI–T has been adapted for a Spanish-speaking population. Therefore, Form Y of the STAI–T was translated into Spanish by A. Bados, a bilingual psychologist who was also an expert in the field of anxiety. A specialist English translator then produced a back translation that was compared with the original, revealing no relevant differences. In a second phase, we compared the 14 items of this Spanish version of Form Y of the STAI–T that are shared with Form X with those of existing Spanish adaptations of Form X, both those produced in Spain (Spielberger, Gorsuch, & Lushene, 1982) and Latin America (Spielberger & Díaz-Guerrero, 1975). The statements were identical or practically identical, thus lending support to the reliability of the translation of the other six items. Finally, we administered the questionnaire to 35 year-1 students and 43 year-4 students from the Faculty of Psychology of the University of Barcelona. We asked these students to indicate and comment on any item statements whose wording seemed problematic, ambiguous, or difficult to understand. No relevant problems were detected in this process.

Once this process was complete, we asked the 346 students to respond anonymously to the questionnaires during class time. They were told that it formed part of a study being conducted about feelings, complaints, and problems experienced by different groups of people. We determined the order of administration of the measures randomly for each respondent.

Data Analysis

To analyze the structure of the STAI–T by means of a confirmatory factor analysis, we considered five models: (a) one factor (General Anxiety); (b) two correlated method factors; (c) a one-construct, two-method model in which all the items load on the same construct (anxiety) and on one of the two method factors defined by the item polarity; (d) two factors (Anxiety and Depression) loading on a higher order factor (Negative Affect); and (e) a bifactor model with two specific factors (on which the items for anxiety and depression loaded, respectively) and a general one (on which all the items loaded as indicators of negative affect). All three factors in the bifactor model with two specific factors were first order. The anxiety and depression items were the same as in the previous model. In Models c, d, and e, we kept the factors involved orthogonal.

We conducted the confirmatory factor analysis using the statistical package Lisrel 8.8 (Jöreskog & Sörbom, 2006), and we analyzed the covariance matrices using maximum likelihood as the method of estimation. We evaluated the fit of the model by means of the following indexes: (a) χ^2 goodness-of-fit statistic, (b) the ratio χ^2/degrees of freedom (*df*), (c) GFI, d) non-normal fit index (NNFI), (e) CFI, (f) RMSEA, and (g) Akaike's information criterion (AIC). To interpret these indexes, we proposed the following criteria: χ^2/df ratio < 2 (excellent); $\chi^2/df < 3$ (good); $\chi^2/df < 5$ (acceptable); GFI $\geq .90$; CFI $\geq .90$; NNFI $\geq .90$; and RMSEA $\leq .05$ (Clara et al., 2001; Hu & Bentler, 1999). Browne and Cudeck (1992) argued that RMSEA values below .05 indicate a good fit, values between .05 and .08 an acceptable fit, and values between .08 and .10 a marginal fit. Regarding the AIC, values close to zero reflect a good fit; and among two AIC values, the lower one reflects the model with the better fit.

We compared the pairs of correlations in the analysis of discriminant validity by means of the z statistic of Meng, Rosenthal, and Rubin (1992). To minimize the probability of Type I and Type II error, we used a modified Bonferroni procedure (Jaccard & Wan, 1996). According to this procedure, the observed levels of significance are ordered by size, and the lowest value must exceed the traditional Bonferroni level (.05/number of analyses); whereas for each subsequent value, .05 is divided by a number that is 1 fewer [i.e., .05/(number of analyses – 1), .05/(number of analyses – 2), etc.]. This method preserves an overall Type 1 error rate of .05, without inflating Type II error.

RESULTS

Statistics of the STAI–T Form Y Spanish Version

Table 1 presents the statistics of the items, total score, and subscales of the STAI–T. Two items ("I feel inadequate"; "I feel like a failure") had high and positive values of skewness and kurtosis, although, given their content, this is not a surprising finding in an undergraduate population. The corrected item-total correlation for both items was .50. As in the original manual, the corrected item-total correlations were .50 or higher for more than half of the items, whereas the median and range of the corrected item-total correlations were .52 and .41 to .70, respectively (.57 and .38–.69 for the college students in the manual;

TABLE 1.—Statistics of the items, total score, and subscales of the STAI-T.

Abridged Items in Spanish (and English)[a]	*M*	*SD*	Skewness	Kurtosis	Corrected Item-Total Correlation
1. *Bien* (pleasant)	1.74	.78	.81	.01	.70
2. *Nervioso* (nervous)	2.17	.73	.56	.44	.51
3. *Satisfecho de mí mismo* (satisfied with myself)	2.00	.94	.51	−.80	.65
4. *Desearía ser tan feliz como otros* (wish to be as happy as others)	1.95	.94	.71	−.42	.52
5. *Fracasado* (failure)	1.28	.51	1.78	3.14	.50
6. *Descansado* (rested)	2.49	.90	−.06	−.76	.50
7. *Tranquilo* (calm)	2.54	1.04	−.15	−1.14	.41
8. *Dificultades se amontonan, no puedo superarlas* (difficulties piling up, impossible to overcome)	1.69	.72	.83	.43	.51
9. *Preocupación por cosas sin importancia* (worry over unimportant things)	2.00	.90	.68	−.23	.60
10. *Feliz* (happy)	1.83	.88	.70	−.51	.70
11. *Pensamientos perturbadores* (disturbing thoughts)	1.65	.73	.99	.71	.57
12. *Falta de confianza en mí mismo* (lack self-confidence)	2.14	.96	.48	−.69	.62
13. *Seguro* (secure)	2.20	.97	.15	−1.11	.64
14. *Decisiones con facilidad* (decisions easily)	2.42	.97	.07	−.96	.46
15. *Incapaz* (inadequate)	1.25	.52	2.35	6.85	.50
16. *Contento* (content)	1.74	.80	.68	−.55	.70
17. *Molesto por pensamientos sin importancia* (bothered by unimportant thoughts)	1.64	.75	1.09	.91	.49
18. *Desengaños me afectan mucho* (disappointments keenly)	2.10	.92	.53	−.49	.46
19. *Estable* (steady)	2.12	1.02	.48	−.91	.63
20. *Tensión al pensar en preocupaciones recientes* (tension when thinking over recent concerns)	1.84	.85	.76	−.13	.51
STAI–T: Total score	38.79	10.33	.62	−.36	—
STAI–T: Depression	25.70	7.42	.56	−.44	—
STAI–T: Anxiety	13.09	3.85	.76	.22	—
STAI–T: Positive Polarity	19.72	5.40	.76	.10	—
STAI–T: Negative Polarity	19.07	5.77	.45	−.57	—

Note. STAI–T = State-Trait Anxiety Inventory, Trait version.
[a]The complete version of the items in Spanish may be obtained by contacting A. Bados. The full version of the items in English can be consulted in the original manual and in Bieling et al. (1998).

Spielberger et al., 1983). The mean and standard deviation of the total score were similar to those obtained by college students in the original STAI manual ($M = 39.6$, $SD = 9.79$; Spielberger et al., 1983). All these data were also similar to those obtained for the French-Canadian (Gauthier & Bouchard, 1993) and Norwegian adaptations of the STAI–T (Haseth, Hagtvet, & Spielberger, 1990). The mean, standard deviation, skewness, and kurtosis of the items could not be compared with those of the original STAI–T, as these data do not appear in the manual.

Although the means for the subscales of positive and negative polarity are not included in the original manual, the values we obtained in this study are between those reported for the undergraduate sample used in the Dutch adaptation of the STAI–T (18.2 and 20.3, respectively; Mook, van der Ploeg, & Kleijn, 1992) and those of the French-Canadian adaptation of the STAI–T (21.65 and 19.02, respectively; Vigneau & Cormier, 2009). As would be expected, the means of the possible Depression and Anxiety subscales of the STAI–T were clearly lower than those obtained by groups of patients with different anxiety disorders (31.00–38.34 and 16.39–18.59, respectively) but were higher than the means reported by Bieling et al. (1998) for a group of nonclinical volunteers (23.18 and 10.20, respectively). All the data presented so far, together with those we describe in the next section, support the adequacy of the Spanish version of Form Y of the STAI–T.

Analysis of Reliability and Intercorrelations

Table 2 presents the reliabilities (Cronbach's α) of the total score and subscales of the STAI–T (specific subscales: Depression, Anxiety; method subscales: Positive Polarity, Negative Polarity) as well as the correlations between subscales and between these and the total score. The reliability of the total score was practically identical to that given in the original manual (.90; Spielberger et al., 1983), whereas the reliabilities of the subscales of Positive and Negative Polarity were similar to those obtained by Vigneau and Cormier (2009; .88 and .84, respectively). The reliabilities of the depression and anxiety subscales were also similar to those reported by Bieling et al. (1998; .88 and .78, respectively). Finally, the correlations between subscales and between these and the total score were similar to those obtained in previous studies (Bieling et al., 1998; Vigneau & Cormier, 2009).

Confirmatory Factor Analysis

Table 3 presents the fit indexes achieved by the different models tested. The results of the χ^2 statistic showed that the fit was deficient for all the models ($p < .001$). However, the sensitivity

TABLE 2.—Reliabilities (on the bold diagonal) of the total score and subscales of the STAI–T and correlations between subscales and between these and the total score.

	Total Score	Depression	Anxiety	Positive Polarity	Negative Polarity
Total score	**.91**				
Specific subscales					
Depression	.96	**.89**			
Anxiety	.84	.64	**.81**		
Method subscales					
Positive Polarity	.92	.79	.94	**.84**	
Negative Polarity	.93	.98	.61	.71	**.87**

Note. STAI–T = State-Trait Anxiety Inventory, Trait version.

TABLE 3.—Values of the fit indexes for the different models.

Model	χ^2(df)	χ^2/df	GFI	NNFI	CFI	RMSEA (90% CI)	AIC
1. One general factor	833.57 (170)*	4.90	.76	.74	.77	.13 (.12–.13)	1,141.36
2. Two correlated method factors (r = .78)	673.64 (169)*	3.99	.80	.80	.82	.11 (.10–.12)	903.93
3. One-construct, two-method model	414.76 (150)*	2.77	.88	.88	.91	.076 (.067–.084)	554.76
4. Two factors loading on a higher order factor	608.69 (168)*	3.62	.83	.82	.84	.096 (.088–.10)	764.22
5. Bifactor model (two specific factors plus a general factor)	388.98 (150)*	2.59	.89	.89	.92	.073 (.065–.081)	535.49

Note. GFI = goodness-of-fit index; NNFI = non-normal fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; AIC = Akaike's information criterion.
* $p < .0001$.

of the χ^2 statistic to the violation of its assumptions and, in particular, sample size, means that the evaluation of fit should be based mainly on other indexes that are less sensitive to sample size and deviations from normality. According to these other indexes, three of the models showed a poor fit. Two models—the one-construct, two-method model and the bifactor model with two specific factors—showed the best fit. The indexes observed in these two models are considered acceptable: the χ^2/df ratio, the CFI, and the RMSEA achieved good or acceptable values; whereas the GFI and the NNFI were .88 or .89, very close to the criterion value of .90. Most of these indexes were slightly better for the bifactor model with two specific factors, which also yielded a lower AIC.

Figures 1 and 2 show the model structure and the loading/path coefficients for the two best-fitting models. In both models, all the items loaded considerably on the general factor; whereas approximately half of the items showed important loadings on the other two factors, both specific and method factors.

Analysis of Convergent and Discriminant Validity

Table 4 presents the correlations between the STAI–T and the remaining measures. The total score and the two method subscales of the STAI–T yielded significantly stronger correlations with the measures of depression than with the corresponding measures of anxiety in the questionnaires considered (BDI/BAI, DASS–21, and SCL–90–R). The Depression subscale of the STAI–T showed high correlations with the other measures of depression, whereas the correlations between the Anxiety subscale and other anxiety measures were less strong.

The Depression subscale of the STAI–T correlated significantly stronger with the measures of depression than with the corresponding measures of anxiety. The correlations between the Anxiety subscale of the STAI–T and the Anxiety and Depression subscales of the DASS–21 were not significantly different. However, in contrast to what would be expected, the Anxiety subscale correlated significantly higher with depression than with anxiety in the BDI/BAI and SCL–90–R. In contrast to what occurred with the STAI–T, the Anxiety subscale of

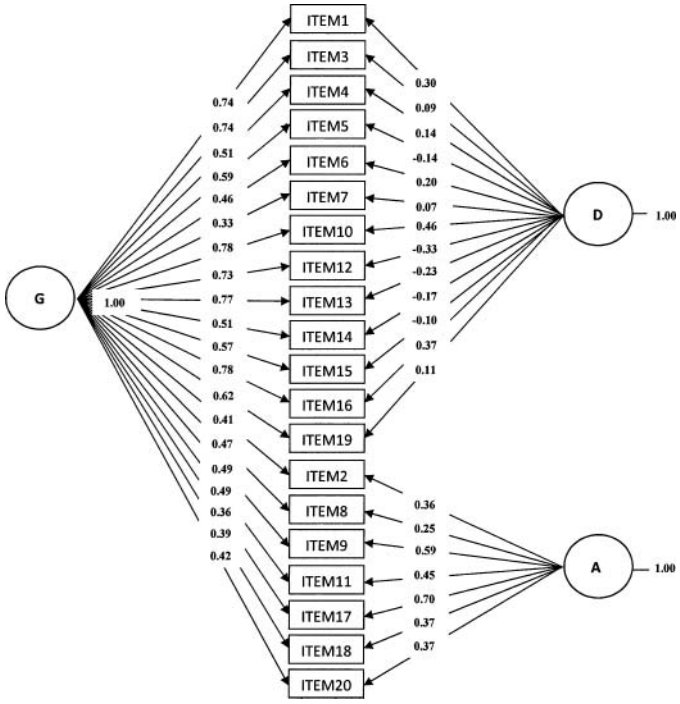

FIGURE 1.—Bifactor model with two specific factors of the State-Trait Anxiety Inventory (Trait version) with standardized parameter estimates. G = General factor; D = Depression factor; A = Anxiety factor.

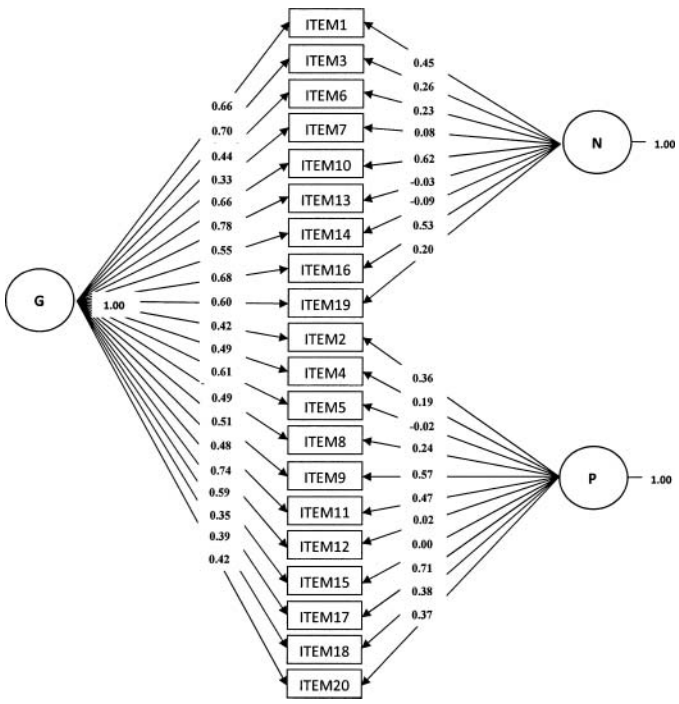

FIGURE 2.—One-construct, two-method model of the State-Trait Anxiety Inventory (Trait version) with standardized parameter estimates. G = General factor; P = Positive Polarity Method factor; N = Negative Polarity Method factor.

TABLE 4.—Correlations between the STAI–T and other questionnaires.

Questionnaire	STAI–T: Depression	STAI–T: Anxiety	STAI–T: Total	STAI–T: Positive Polarity	STAI–T: Negative Polarity
BDI	.70	.61	.73	.70	.65
BAI	.54	.53	.58	.54	.53
Z statistic[a]	4.57*	2.28*	4.51*	4.71*	3.39*
DASS–21: Depression	.61	.49	.62	.58	.57
DASS–21: Anxiety	.42	.43	.46	.45	.41
Z statistic[a]	3.95*	1.03	3.32*	2.66*	3.20*
SCL–90–R: Depression	.70	.67	.75	.74	.66
SCL–90–R: Anxiety	.55	.58	.61	.58	.55
Z statistic[a]	4.68*	2.64*	4.67*	4.97*	3.14*

Note. *N* oscillated between 322 and 333 depending on the questionnaires considered. STAI–T = State-Trait Anxiety Inventory, Trait version; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; DASS–21 = Depression, Anxiety, and Stress Scales, 21-item version; SCL–90–R = Symptom Checklist–90–R; STAI–T = State-Trait Anxiety Inventory, Trait version. All correlations are significant ($p < .000001$).

[a] z statistic of Meng, Roshental, and Rubin (1992) to compare pairs of correlations.

* $p < .025$ (significant value after applying the correction of Jaccard and Wan, 1996).

the other questionnaires showed a higher correlation with other anxiety subscales than with depression subscales (see Table 5).

DISCUSSION

The fit indexes from the confirmatory factor analyses indicate that, of the five models considered, the one-construct, two-method model and the bifactor model with two specific factors were those that showed the best fit, with acceptable fit indexes. The fit of the one-construct, two-method model is consistent with the findings of Vigneau and Cormier (2008) and Vautier and Pohl (2009) and supports the notion that the STAI–T may measure a substantive construct together with method effects due to item polarity. Furthermore, an alternative model that offers a slightly better fit is the bifactor model with two specific factors and one general factor. Thus, there is a general factor (Negative Affect, supposedly) that explains the item communality and two specific domain factors (Depression and Anxiety, supposedly), each of which accounts for the unique influence of the specific domain ahead of the general factor. We point out that there is a considerable overlap between the two models because 7 of the 11 items with positive polarity form the supposed Anxiety subscale; whereas the remaining 4 items, plus the 9 with negative polarity, constitute the supposed Depression subscale.

Indeed, both models are plausible and can be regarded as alternative representations of the structure of the STAI–T. They both show the same pattern whereby all the items load strongly on the general factor, complemented by the presence of method or specific factors on which half of the items load. This raises concerns about the meaning of the general factor and the two specific factors.

TABLE 5.—Correlations between anxiety and depression questionnaires other than the STAI–T.

Questionnaire	BDI	DASS–21: Depression	SCL–90–R: Depression	BAI	DASS–21: Anxiety
DASS–21: Depression	.69				
SCL–90–R: Depression	.81	.76			
BAI	.65	.45	.64		
DASS–21: Anxiety	.47	.4	.48	.65	
SCL–90–R: Anxiety	.64	.44	.64	.79	.61

Note. *N* oscillated between 322 and 333 depending on the questionnaires considered. STAI–T = State-Trait Anxiety Inventory, Trait version; BDI = Beck Depression Inventory; DASS–21 = Depression, Anxiety, and Stress Scales, 21-item version; SCL–90–R = Symptom Checklist–90–R; BAI = Beck Anxiety Inventory.

The data regarding the convergent and discriminant validity of the STAI–T do shed further light on what it actually measures. First, the STAI–T total score was more strongly correlated with measures of depression than with those of anxiety. This is in agreement with the results of previous studies (Bieling et al., 1998; Grös et al., 2007; Ponciano et al., 2006) and suggests that the STAI–T contains items that reflect depression and general negative affect rather than the "absence of anxiety" construct proposed by Spielberger et al. (1983). Indeed, the Negative Polarity subscale (which in theory corresponds to the absence of anxiety) was more strongly correlated with measures of depression than with those of anxiety, and the same occurred with the Positive Polarity subscale.

Second, one might ask whether the two specific factors of the bifactor model correspond to depression and anxiety, as was suggested by Bieling et al. (1998). Examination of the items that make up the two factors of the STAI–T indicates that one group of items evaluates low positive affect and negative self-appraisal, aspects that are typical of depression. Interpreting this second factor as the absence of anxiety, as has been done previously, overlooks the fact that the content of its items does not correspond simply to a lack of fear or worry but is in fact much more specific. Another reason why this factor does not reflect the absence of anxiety is that it showed strong correlations with other measures of depression, and these correlations were higher than those found with other measures of anxiety. These findings are consistent with the results obtained by Bieling et al. in a clinical sample.

The second group of STAI–T items evaluates worry, tension, and disturbing thoughts, which could correspond to the construct of anxiety. However, the discriminant validity of the hypothetical Anxiety subscale is questionable, as its correlation with measures of depression is equivalent to or stronger than the correlation with measures of anxiety. Importantly, this lack of discriminant validity was only observed with the STAI–T, as the other anxiety and depression measures used here showed more satisfactory patterns of convergent and discriminant validity. Bieling et al. (1998) stated that the discriminant validity of the Anxiety subscale of the STAI–T was good because, in comparison with the Depression subscale, it showed higher correlations with other measures of anxiety. However, the anxiety subscale of the STAI was more strongly correlated with the BDI

than with the BAI, and it correlated to the same degree with the Depression and Anxiety scales of the DASS–21. These results coincide with our findings and indicate a lack of discriminant validity.

Furthermore, the convergent validity of the Anxiety subscale was lower than that of the Depression subscale, which is also consistent with the findings of Bieling et al. (1998). This could be explained both by the scale's content and the slightly lower reliability of the anxiety measures used in this study. However, this explanation would not hold for the correlations with the BDI and BAI inventories, which showed similar reliabilities.

A more detailed examination of the items that make up the "Anxiety" factor indicates that some items may be evaluating negative affect more than anxiety in the strict sense; furthermore, there is an absence of items that consider the somatic and behavioral aspects of anxiety. This state of affairs can be considered in light of the tripartite model of anxiety and depression proposed by Clark and Watson (1991). In this model, anxiety, depression and other negative mood states (anger, guilt) present a shared general factor (Negative Affect) and specific factors: Physiological Hyperarousal in the case of anxiety and Anhedonia or the absence of positive affect in the case of depression. According to this model, supported by numerous studies (see Grös et al., 2007), a measure of anxiety should include items that reflect negative affect and physiological hyperarousal but not items related to low positive affect. In contrast, the STAI–T includes several items that reflect positive affect and none that consider physiological hyperarousal.

Thus, the STAI–T would seem to be evaluating negative affect in general more than anxiety specifically. Our results, as well as those of previous studies, have suggested that the questionnaire should be revised. For example, the items that reflect high or low positive affect (e.g. Items 1, 3, 4, 5, 10, and 16) could be deleted, whereas items related to physiological hyperarousal could be added. Another possibility would be to develop a completely new questionnaire, as has been done by Ree, French, MacLeod, and Locke (2008) with the State-Trait Inventory for Cognitive and Somatic Anxiety. This inventory includes both a state and trait form; evaluates anxiety (both cognitive and somatic) in a stricter sense; and, in comparison with the STAI–T, correlates more strongly with the Anxiety scale of the DASS–21 and more weakly with its Depression scale.

Among the limitations of this study, we note that there was an exclusive use of self-report measures. In addition, the process of adapting the STAI–T into Spanish could have been more complete (e.g., by using a larger group of experts; see Hambleton, 2001); and we acknowledge that item bias analyses were not conducted on an item-by-item basis because the corresponding data were not available in the original STAI–T manual. The translation showed a good match for the 14 items that are shared with Form X of the STAI–T, and the data obtained regarding the psychometric properties of the 20 items and the questionnaire as a whole are very similar to those reported in the original manual and by other studies that have used the STAI–T in English. This suggests that the questionnaire adaptation was adequate. Finally, as the undergraduates were all psychology students and most of them were female, it is difficult to know to what extent the results might be generalizable to undergraduates as a whole or to different populations. However, other studies (Bieling et al., 1998; Grös et al., 2007; Ponciano et al., 2006) that have been conducted with clinical and nonclinical samples (secondary school pupils and undergraduates) in other countries have also reported that the STAI–T mainly evaluates depression and, in general, negative affect rather than anxiety alone.

Acknowledgment

This study was partially supported by Grant 2009SGR00822 from the "Generalitat de Catalunya."

References

Bados, A., Solanas, A., & Andrés, R. (2005). Psychometric properties of the Spanish version of Depression, Anxiety and Stress Scales. *Psicothema, 17,* 679–683.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. New York, NY: Guilford.

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

Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The State-Trait Anxiety Inventory, trait version: Structure and content re-examined. *Behaviour Research and Therapy, 36,* 777–788.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21,* 230–258.

Caci, H., Baylé, F., Dossios, C., Robert, P., & Boyer, P. (2003). The Spielberger Trait Anxiety Inventory measures more than anxiety. *European Psychiatry, 18*, 394–400.

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and secondorder models of quality of life. *Multivariate Behavioral Research, 41,* 189–225.

Clara, I. P., Cox, B. J., & Enns, M. W. (2001). Confirmatory factor analysis of the Depression-Anxiety-Stress Scales in depressed and anxious patients. *Journal of Psychopathology and Behavioral Assessment, 23,* 61–67.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100,* 316–336.

Derogatis, L. R. (1983). *SCL–90–R: Administration, scoring and procedures—Manual II* (2nd ed.). Baltimore, MD: Clinical Psychometric Research.

Derogatis, L. R. (2002). *SCL–90–R: Cuestionario de 90 síntomas* [SCL–90–R: Symptom Checklist–90–R]. Madrid, Spain: TEA. (Spanish adaptation by González de Rivera, J. L. et al.)

Fountoulakis, K., Papadopoulou, M., Kleanthous, S., Papadopoulou, A., Bizeli, V., Nimatoudis, I., . . . Kaprinis, G. S. (2006). Reliability and psychometric properties of the Greek translation of the State-Trait Anxiety Inventory form Y: Preliminary data. *Annals of General Psychiatry, 5,* 2–11.

Gauthier, J., & Bouchard, S. (1993). Adaptation canadienne-française de la forme révisée du State-Trait Anxiety Inventory de Spielberger [French-Canadian adaptation of the revised version of the Spielberger's State-Trait Anxiety Inventory]. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 25,* 559–578.

Grös, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA): Comparison to the State-Trait Anxiety Inventory (STAI). *Psychological Assessment, 19,* 369–381.

Hambleton, R. (2001). The next generation of the ITC test translation and adaptation guidelines. *European Journal of Psychological Assessment, 17,* 164–172.

Haseth, K., Hagtvet, K. A., & Spielberger, C. D. (1990). Psychometric properties and research with the Norwegian State-Trait Anxiety Inventory. In C. D. Spielberger, R. Diaz-Guerrero, & J. Strelau (Eds.), *Cross-cultural anxiety* (Vol. 4, pp. 169–181). New York, NY: Hemisphere.

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

Jaccard, J., & Wan, C. K. (1996). *LISREL approaches to interaction effects in multiple regression.* Thousand Oaks, CA: Sage.

Jöreskog, K., & Sörbom, D. (2006). LISREL 8.8 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Kohn, P., Kantor, L., DeCicco, T., & Beck, A. (2008). The Beck Anxiety Inventory-Trait (BAIT): A measure of dispositional anxiety not contaminated by dispositional depression. *Journal of Personality Assessment, 90,* 499–506.

Lovibond, P. F., & Lovibond, S. H. (1995). *Manual for the Depression Anxiety Stress Scales*. Sidney, Australia: Psychology Foundation of Australia.

Meng, X.-L., Roshenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111,* 172–175.

Mook, J., van der Ploeg, H. M., & Kleijn, W. C. (1992). Symptom-positive and symptom-negative items in the State-Trait Anxiety Inventory: A comparison and replication. *Anxiety, Stress and Coping, 5,* 113–123.

Ponciano, E., Daniel, F., Rodrigues, M. J., Medeiros, T., Jardim, H., Cardoso, I., . . . Spielberger, C. (2006, September). *Validity and psychometric properties of STAI–Y on Portuguese high-school and university students*. Poster session presented at the VI International Congress of the Sociedad Española para el Estudio de la Ansiedad y el Estrés (SEAS), Benidorm, Alicante, Spain.

Ree, M., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing cognitive and somatic dimensions of state and trait anxiety: development and validation of the State-Trait Inventory for Cognitive and Somatic Anxiety (STICSA). *Behavioural and Cognitive Psychotherapy, 36,* 313–332.

Sanz, J., & Navarro, M. E. (2003). Propiedades psicométricas de una versión española del Inventario de Ansiedad de Beck (BAI) en estudiantes universitarios [Psychometric properties of a Spanish version of the Beck Anxiety Inventory (BAI) in an undergraduate sample]. *Ansiedad y Estrés, 9,* 59–84.

Sanz, J., & Vázquez, C. (1998). Fiabilidad, validez y datos normativos del Inventario para la Depresion de Beck [Reliability, validity and normative data for the Beck Depression Inventory]. *Psicothema, 10,* 303–318.

Spielberger, C. D., & Díaz-Guerrero, R. (1975). *IDARE: Inventario de ansiedad, rasgo y estado. Manual e instructivo* [IDARE: State-Trait Anxiety Inventory. Manual]. Ciudad de, México: El Manual Moderno.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *STAI: Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1982). *Cuestionario de ansiedad estado-rasgo: Manual* [State-Trait Anxiety Inventory: Manual]. Madrid, Spain: TEA.

Spielberger, C. D., Gorsuch, R. L., Lushene, R. E., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory STAI (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Spielberger, C. D., Sydeman, S. J., Owen, A. E., & Marsh, B. J. (1999). Measuring anxiety and anger with the State-Trait Anxiety Inventory (STAI) and the State-Trait Anger Expression Inventory (STAXI). In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (2nd ed., pp. 993–1021). Mahwah, NJ: Lawrence Erlbaum Associates.

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological Assessment, 21,* 187–193.

Vigneau, F., & Cormier, S. (2008). The factor structure of the State-Trait Anxiety Inventory: An alternative view. *Journal of Personality Assessment, 90,* 280–285.

Vigneau, F., & Cormier, S. (2009). L'inventaire d'anxiété situationnelle et de trait d'anxiété (IASTA-Y): Structure factorielle et biais linguistique [The State-Trait Anxiety Inventory (STAI): Factorial structure and linguistic bias]. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 41,* 115–120.