

Robust Dimensions of Anxiety Sensitivity: Development and Initial Validation of the Anxiety Sensitivity Index—3

Steven Taylor
University of British Columbia

Michael J. Zvolensky
University of Vermont

Brian J. Cox
University of Manitoba

Brett Deacon
University of Wyoming

Richard G. Heimberg and Deborah Roth Ledley
Temple University

Jonathan S. Abramowitz
University of North Carolina at Chapel Hill

Robert M. Holaway
Temple University

Bonifacio Sandin
Universidad Nacional de Educación a Distancia

Sherry H. Stewart
Dalhousie University

Meredith Coles
State University of New York at Binghamton

Winnie Eng
City University of New York

Erin S. Daly
Boston University and Veterans Affairs Boston Healthcare System

Willem A. Arrindell
University of Groningen

Martine Bouvard
Université de Savoie

Samuel Jurado Cardenas
University of Mexico

Accumulating evidence suggests that anxiety sensitivity (fear of arousal-related sensations) plays an important role in many clinical conditions, particularly anxiety disorders. Research has increasingly focused on how the basic dimensions of anxiety sensitivity are related to various forms of psychopathology. Such work has been hampered because the original measure—the Anxiety Sensitivity Index (ASI)—was not designed to be multidimensional. Subsequently developed multidimensional measures have unstable factor structures or measure only a subset of the most widely replicated factors. Therefore, the authors developed, via factor analysis of responses from U.S. and Canadian nonclinical participants ($n = 2,361$), an 18-item measure, the ASI-3, which assesses the 3 factors best replicated in previous research: Physical, Cognitive, and Social Concerns. Factorial validity of the ASI-3 was supported by confirmatory factor analyses of 6 replication samples, including nonclinical samples from the United States and Canada, France, Mexico, the Netherlands, and Spain ($n = 4,494$) and a clinical sample from the United States and Canada ($n = 390$). The ASI-3 displayed generally good performance on other indices of reliability and validity, along with evidence of improved psychometric properties over the original ASI.

Keywords: anxiety sensitivity, Anxiety Sensitivity Index, anxiety disorders

Supplemental material: <http://dx.doi.org/10.1037/1040-3590.19.2.176.supp>

Steven Taylor, Department of Psychiatry, University of British Columbia, Vancouver, Canada; Michael J. Zvolensky, Department of Psychology, University of Vermont; Brian J. Cox, Department of Psychiatry, University of Manitoba, Canada; Brett Deacon, Department of Psychology, University of Wyoming; Richard G. Heimberg, Deborah Roth Ledley, and Robert M. Holaway, Department of Psychology, Temple University; Jonathan S. Abramowitz, Department of Psychology, University of North Carolina at Chapel Hill; Bonifacio Sandin, Department of Psychology, Universidad Nacional de Educación a Distancia, Madrid, Spain; Sherry H. Stewart, Departments of Psychiatry and Psychology, Dalhousie University, Halifax, Nova Scotia, Canada; Meredith Coles, Department of Psychology, State

University of New York at Binghamton; Winnie Eng, Department of Psychology, City University of New York; Erin S. Daly, Department of Psychiatry, Boston University, and Veterans Affairs Boston Healthcare System; Willem A. Arrindell, Department of Clinical Psychology, University of Groningen, the Netherlands; Martine Bouvard, Department of Psychology, Université de Savoie, Chambéry, France; Samuel Jurado Cardenas, Department of Psychology, University of Mexico, Mexico City.

Correspondence concerning this article, including requests for obtaining copies of the ASI-3, should be addressed to Steven Taylor, Department of Psychiatry, University of British Columbia, Vancouver, British Columbia V6T 2A1, Canada. E-mail: taylor@unixg.ubc.ca

Anxiety sensitivity (AS) is the fear of arousal-related sensations, arising from beliefs that the sensations have adverse consequences such as death, insanity, or social rejection (Reiss & McNally, 1985). AS is conceptualized as a contributor to individual differences in general fearfulness and as a diathesis for various types of anxiety disorders, including panic disorder, social anxiety disorder, specific phobia, and posttraumatic stress disorder (Reiss & McNally, 1985; Taylor, 1999). This is because AS is an anxiety amplifier; when highly anxiety-sensitive people become anxious, they become alarmed about their arousal-related sensations, which further intensifies their anxiety. Consistent with this formulation is evidence showing that AS is elevated in people with various types of anxiety disorders compared with control participants, and a person's current level of AS predicts the risk of future anxiety symptoms (see Taylor, 1999, for a review).

AS was originally conceived as a unidimensional construct, as measured by the Anxiety Sensitivity Index (ASI; Peterson & Reiss, 1992). Many factor analyses of the ASI have been conducted, with solutions ranging from one to four factors (Taylor, 1999). The inconsistent factor solutions may have arisen for various reasons, including differences across studies in factor selection criteria (e.g., use of the eigenvalue > 1 rule in some studies, which can lead to factor overextraction) and the use of small samples in some studies, which can yield unreliable findings (Taylor, 1999). The fact that the ASI was not constructed to be multidimensional may also have contributed to instability (lack of replicability) in the factor structure, because some of the domains of AS (as described below) were measured by only a few items, thereby reducing the odds that a given factor would be reliably obtained across different samples.

Despite the inconsistencies, the most commonly obtained factor solution consists of three correlated factors labeled Physical Concerns, Cognitive Concerns, and Social Concerns (Taylor, 1999). To illustrate these factors, high scores on Physical Concerns are associated with the belief that palpitations lead to cardiac arrest. High scores on Cognitive Concerns are associated with the belief that concentration difficulties lead to insanity. High scores on Social Concerns are associated with the belief that publicly observable anxiety reactions (e.g., trembling) will elicit social rejection or ridicule. The three factors load on a single higher order factor (Global AS). The factor structure does not appear to vary as a function of gender or age, at least for the age ranges that have been studied (i.e., adults and children as young as 7 years; e.g., Dehon, Weems, Stickle, Costa, & Berman, 2005; Stewart, Taylor, & Baker, 1997).

There are several problems with the ASI for the multidimensional assessment of AS. The ASI Physical Concerns subscale (constructed on the basis of factor analysis) has eight items, whereas the Cognitive and Social Concerns subscales each contain only four items. Therefore, the latter two subscales might not have adequate reliability, especially if they do not have strong content validity. Related to the issue of content validity, particularly for the purpose of measuring specific AS dimensions, is the fact that some ASI items do not target specific dimensions. For example, the item "It scares me when I am nauseous" could measure physical concerns, or it could measure social concerns (i.e., concerns about the social consequences of vomiting in front of others).

In attempts to overcome the problems with the ASI, Taylor and Cox (1998a, 1998b) developed a 36-item revised scale—the ASI—

Revised (ASI-R)—and a 60-item Anxiety Sensitivity Profile. Unfortunately, both scales have unstable factor structures, with different studies obtaining different solutions (e.g., Deacon, Abramowitz, Woods, & Tolin, 2003; Zvolensky et al., 2003). Further research is needed to develop a psychometrically sound, multidimensional measure of AS.¹ Accordingly, in the present series of investigations we selected items from the ASI-R, which is the most widely used multidimensional measure of AS, and constructed and evaluated a revised scale, the ASI-3, which measures the three most commonly replicated dimensions: physical, cognitive, and social concerns. This scale, like the original ASI, was intended for use in clinical and nonclinical samples. Although it is theoretically important to develop a scale that can reliably and validly assess AS in clinical samples, it is equally important that such a scale should be able to assess AS in nonclinical samples in order to identify people who are theoretically at risk for developing anxiety disorders or related problems. Accordingly, we evaluated the ASI-3 with both types of samples.

Study 1 concerned the development of the ASI-3. Study 2 evaluated the factorial validity of the ASI-3 in six replication samples and compared the results with those of the ASI. Study 3 examined reliability as internal consistency of each of the ASI-3 subscales, which were compared with those of the ASI subscales. Although only one index of reliability was examined (internal consistency), it provides a good estimate of reliability in general because the sampling of item content is usually the major source of measurement error for traitlike constructs (Nunnally & Bernstein, 1994). Study 4 examined convergent, discriminant, and criterion-related (known groups) validities. The studies described in this article were based on data collected, with informed consent, in previous research (Bernstein et al., 2006; Deacon et al., 2003; Roth, Coles, & Heimberg, 2002; Taylor & Cox, 1998a; Zvolensky et al., 2003). None of those studies addressed the aims of the present article.

Study 1: Construction of the ASI-3

A goal of this study was to construct the ASI-3 by selecting ASI-R items that each measured only one of the domains of physical, cognitive, or social concerns. A further goal was to determine whether the items formed subscales corresponding to separate but correlated factors. Item selection for the ASI-3 was done by balancing two opposing goals: to have an overall scale that is as efficient to administer as the original (16-item) ASI, while also having subscales that each contain a sufficient number of items to ensure adequate subscale reliability. Therefore, it was decided to create three 6- to 7-item subscales.

Method

Participants. A sample of 4,720 young adults were recruited from universities across the United States and Canada: Dalhousie

¹ Blais et al. (2001) developed a modified version of the ASI by deleting items that were not highly correlated with the total score. This resulted in the deletion of all social concerns items, and so the scale measured only two of the three most widely replicated factors (physical and cognitive concerns). A comprehensive assessment of AS requires that all of the most widely replicated factors be assessed. Given these concerns, we did not specifically examine the Blais et al. ASI in the present study.

University ($n = 155$), University of British Columbia ($n = 156$), University of Manitoba ($n = 167$), University of Vermont ($n = 347$), Temple University ($n = 1,912$), and Northern Illinois University ($n = 1,983$). Participants were all undergraduates, except for those recruited from the University of Vermont site, who were a mix of 70% college students and 30% nonstudent young adults recruited from the community. The overall sample of 4,720 participants was randomly split in two. The U.S.–Canadian Subsample 1 ($n = 2,361$) was used for the construction of the ASI–3 in the present study. Subsample 2 was one of the samples used in subsequent studies described in this article.² The mean age of Subsample 1 was 19.6 years ($SD = 3.4$), and 66% were women. The majority of participants were White (61%), with the remainder being African American (19%), Asian (9%), Hispanic (3%), or other (8%).

Measures and procedure. All participants completed a short questionnaire assessing demographic features. U.S. participants completed the 36-item ASI–R. Canadian participants completed a 42-item scale, consisting of the original 16-item ASI and the 36-item ASI–R (the ASI and ASI–R have 10 items in common). Participants completed the measures in classroom or individual settings for either course credit or an honorarium.

Scale construction and statistical methods. Items were initially selected, from the pool of 36 ASI–R items, with an emphasis on content validity. An item was selected if its content unambiguously corresponded to only one of the domains of physical, cognitive, or social concerns. Twelve items were eliminated because it was unclear which domain they measured. For example, “It scares me when I feel faint” could measure physical concerns (e.g., fear of fainting and injuring oneself) or social concerns (e.g., fear of making a spectacle of oneself by collapsing in front of others). Similarly, the item “When my head is pounding, I worry I could have a stroke” was eliminated because it could measure cognitive concerns (fear of cognitive impairment resulting from a stroke) or physical concerns (fear of stroke-related death or physical impairment). Of the remaining 24 ASI–R items, 3 were eliminated because their wording was very similar to that of other items. For each set of such redundant items, we retained the item with the lowest Flesch–Kincaid reading level. The resulting pool consisted of seven items for each of the three subscales, in which each subscale assessed a latent factor (Physical, Cognitive, or Social Concerns).³

The degree of fit of the items to the three-factor model was tested with LISREL 8.72 (Jöreskog & Sörbom, 2005), by means of polychoric correlations because the data were ordinal, and robust weighted least-squares. As such, all of the resulting fit indices were robust estimates. A model-fitting (confirmatory factor analytic) approach was used instead of a purely exploratory factor analytic approach because we hypothesized, on the basis of previous research (e.g., Zinbarg, Barlow, & Brown, 1997), that a three-factor solution would provide a good fit to the data.⁴ However, model modification indices were used in this study, and so confirmatory factor analysis was used in an exploratory mode for the purpose of item selection.

Factors were allowed to covary with one another because previous AS research has consistently shown that such factors are correlated (e.g., Taylor & Cox, 1998a, 1998b; Taylor, Koch, Woody, & McLean, 1996). Error terms (residuals) were not permitted to be correlated. The model was simultaneously fitted to

two groups: women and men from Subsample 1. This was done to test whether the factor structure was invariant across gender, as suggested by previous studies of AS, and to make any model modifications (item deletions) that equally emphasized model fit in the samples of women and men.

A stepwise approach was adopted for the multigroup analyses: (a) An individual three-factor confirmatory analysis was conducted for each subgroup (i.e., separate analyses for women and men in the present study), (b) tests of equal form across subgroups (i.e., fitting the same number of factors to each gender), (c) tests of invariance of factor loadings across subgroups (i.e., equal factor loadings for the three-factor model across gender), (d) tests of invariance of factor correlations, with factor loadings constrained to be invariant, and finally, (e) tests of invariance of error variances (with loadings, error variances, and factor correlations constrained). We tested whether model fit deteriorated with successive constraints, as assessed by changes in fit indices and, in the case of nested models, by Satorra–Bentler corrected differences in chi-square (i.e., χ^2_{diff}). Only the main analyses for this stepwise approach are reported in this article. Here, we report the results for the most stringent evaluation of the ASI–3, that is, the results for Step (e). The results for the remaining steps are available in an appendix (available online), which reports these and other results pertinent to, but not essential for, an adequate psychometric evaluation of the ASI–3.

The selection of fit indices was based on the findings and recommendations of Hu and Bentler (1998). They recommended

² We chose to combine the U.S.–Canadian samples and then randomly split the combined group in two because that should improve the generalizability (external validity) of the results. That is, the scale construction sample (Subsample 1) consisted of respondents from several different regions and universities. This meant that the final 18 items of the ASI–3 were selected to be broadly applicable to many samples instead of being fitted to one particular sample. The latter can lead to model overfitting (i.e., fitting the model to the specific characteristics of a given sample; Brown, 2006), which undermines the generalizability of the results. Accordingly, Subsamples 1 and 2 were deliberately heterogeneous (in terms of our U.S.–Canadian recruitment sites), with the goal of creating a scale that would have a stable factor structure across other samples. As it turned out, the results from these combined groups were no different from the findings from other single-site nonclinical groups, that is, the Dutch, French, Mexican, and Spanish groups. In other words, the homogeneity versus heterogeneity of the samples had no effect on the pattern of results.

³ Statistical analyses supported this approach to item selection. Here, all 36 ASI–R items were assigned, on the basis of their content, to Physical, Cognitive, or Social Concerns subscales. If an item was relevant to more than one domain (e.g., physical and social concerns), then it was assigned to more than one subscale. Each subscale was individually factor analyzed with principal axis factor analysis, and the first factor was extracted. Thus, three single-factor solutions were obtained, each representing one of the three domains of physical, cognitive, or social concerns. The 18 items that were ultimately retained in the ASI–3 all had strong ($>.50$) loadings on their respective factors. Collectively, the loadings for items included in the ASI–3 were one standard deviation higher than those of items that were excluded; respectively, the means (and standard deviations) were .64 (.07) versus .56 (.09).

⁴ The main model that was tested, consisting of three correlated factors, yields identical goodness-of-fit results as a hierarchical model in which the three factors load on a single higher order factor (Brown, 2006).

using at least two indices, one of which is the standardized-root-mean-square residual (SRMR). Of the other recommended indices, we selected the root-mean-square error of approximation (RMSEA), the comparative fit index (CFI) and the Tucker–Lewis Index (TLI). The SRMR was used because it is among the most sensitive to misspecified factor correlations, and the other fit indices were selected because they are among the most sensitive to misspecified factor loadings (Hu & Bentler, 1998). To interpret whether a given factor model provided a good fit to the data, we used Hu and Bentler's (1999) empirically derived cutoff values. These values minimize errors in deciding whether a model provides a good fit to the data. Good fit is indicated by $SRMR \leq .08$, $RMSEA \leq .06$, $CFI \geq .95$, or $TLI \geq .95$. For descriptive purposes, chi-square values are also reported for each confirmatory factor analysis.

For analyses in which two or more competing models were tested, the relative goodness of fit was tested in several ways. If the models were nested, then Satorra–Bentler corrected differences in chi-square values were computed. The 90% confidence interval (CI) for each RMSEA value was also computed to compare competing factor models.

Results and Discussion

For the 21 items, the percentage of missing data for a given item ranged from 0% to 0.3%. The three-factor model for the 21 items—fitted simultaneously to women and men according to Step (e) described above—yielded two fit indices indicating a good fit: $CFI = .972$ and $TLI = .972$. The values of the other indices fell outside the threshold for good fit: $SRMR = .085$, $RMSEA = .077$ ($RMSEA$ 90th percentile CI: .074, .079), $\chi^2(417, N = 2,361) = 3,305.44, p < .001$. To improve model fit, we used LISREL model modification indices to identify the worst-fitting item from each subscale, that is, the item for which there would be the greatest drop in chi-square if that item was allowed to cross-load on one or both of the other factors. This method enabled us to identify items that measured more than one latent dimension (i.e., items that were not pure measures of a given factor). Three items were identified and deleted, and the three-factor model was tested on the remaining 18 items, in the manner described above. With this minor modification, all four fit indices indicated that the three-factor (18-item) model had a good fit to the data: $CFI = .986$, $TLI = .986$, $SRMR = .060$, and $RMSEA = .058$ (90th percentile CI: .055, .061), $\chi^2(303, N = 2,361) = 1,487.79, p < .001$. The items and their loadings appear in Table 1, which represents the model simultaneously fitted to the groups of women and men. The three factors were correlated .70 to .82. When subscale scores were computed on the basis of the unit-weighted sum of item scores for a given factor, the subscale correlations ranged from .53 to .62. In the remaining studies reported in this article, subscale scores were computed on the basis of unit-weighted sum of items.

A good fit to the data was also obtained for the three-factor (18-item) model when data from women and men were combined to form a single group: $CFI = .986$, $TLI = .984$, $SRMR = .051$, $RMSEA = .058$ (90th percentile CI: .055, .061), $\chi^2(132, N = 2,361) = 1,163.62, p < .001$. The three factors were correlated .70 to .81. The matrix of polychoric correlations for Subsample 1 (combining women and men) is shown for descriptive purposes in Table 2. The mean interitem polychoric correlations for each subscale were as follows: Physical Concerns (.48), Cognitive

Concerns (.63), and Social Concerns (.48). For each subscale for this combined sample, coefficient alpha and the range of corrected item-total Pearson product-moment correlations (r_{it}) were as follows: Physical Concerns $\alpha = .79$, r_{it} range = .48 to .60; Cognitive Concerns $\alpha = .84$, r_{it} range = .48 to .67; Social Concerns $\alpha = .79$, r_{it} range = .46 to .72.

As a further test of whether the three-factor model provided an optimal fit to the data, we compared it to the one- and two-factor models, using the two-group (women vs. men) methods described above. These one- and two-factor models have been obtained in some previous factor analyses (e.g., Zvolensky et al., 2003), although they have not been as widely replicated as the three-factor model. For the one-factor model, all 18 items loaded on a single dimension. The fit values were as follows: $CFI = .967$, $TLI = .967$, $SRMR = .083$, and $RMSEA = .088$ (90th percentile CI: .085, .091), $\chi^2(306, N = 2,361) = 3,094.31, p < .001$. The two-factor model consisted of two correlated factors: a Physical Concerns factor and a combined Cognitive–Social Concerns factor. The fit values were $CFI = .971$, $TLI = .971$, $SRMR = .076$, and $RMSEA = .081$ (90th percentile CI: .079, .084), $\chi^2(305, N = 2,361) = 2,690.52, p < .001$. Thus, the three-factor model was associated with better values than the one- and two-factor models on all four fit indices. Whereas the three-factor model yielded a good fit to the data on all four indices, the one- and two-factor models yielded a good fit on only two of four indices. Inspection of the CIs for the RMSEA indicated that the three-factor model had a significantly better fit than both the one- and two-factor models, as indicated by nonoverlapping CIs. The one- and two-factor models did not differ from one another in their degree of fit to the data.

The three-factor model also had a superior fit to the one- and two-factor models in terms of Satorra–Bentler corrected differences in chi-square values: Factor 1 versus 2, $\chi^2_{diff}(1, N = 2,361) = 40.99, p < .001$; Factor 2 versus 3, $\chi^2_{diff}(2, N = 2,361) = 245.34, p < .001$; and Factor 1 versus 3, $\chi^2_{diff}(3, N = 2,361) = 257.31, p < .001$.

In summary, in the present study an 18-item three-factor model provided the best fit to the data, for both women and men. The items from this scale form the ASI–3, which was derived from the ASI–R and contains five items from the original ASI (one to two items per subscale, as shown in Table 1). The ASI–3 has an overall Flesch–Kincaid reading level of Grade 6.5, indicating that it would be readily comprehended by the majority of adults.

Study 2: Factorial Validity

Factorial validity is a form of construct validity established through factor analysis. In the case of the ASI–3, this form of validity was investigated by determining whether the three-factor model could be replicated across different samples. Two types of tests of factorial validity were conducted, as part of a stepwise multigroup confirmatory factor analysis. The more liberal one involved determining whether the three factors provided a good fit to the data in separate confirmatory factor analyses of each of six replication samples, including a sample in which a two-group (women vs. men) analysis was conducted (i.e., in Subsample 2; the remaining samples were not large enough for such a two-group analysis). A more rigorous, multigroup confirmatory factor analysis was also conducted, in which a common factor structure was

Table 1
Study 1: U.S.–Canadian Subsample 1 (n = 2,361)—Loadings (and Standard Errors) for Final, Multigroup Three-Factor Solution of the ASI–3

| Item no. | Item | Factor 1: Physical Concerns | Factor 2: Cognitive Concerns | Factor 3: Social Concerns |
|----------|--|-----------------------------------|------------------------------------|---------------------------------|
| 4 | When my stomach is upset, I worry that I might be seriously ill. ^a | .79 (.02) | | |
| 12 | When I notice my heart skipping a beat, I worry that there is something seriously wrong with me. | .76 (.02) | | |
| 8 | When I feel pain in my chest, I worry that I'm going to have a heart attack. | .69 (.02) | | |
| 7 | When my chest feels tight, I get scared that I won't be able to breathe properly. | .68 (.02) | | |
| 15 | When my throat feels tight, I worry that I could choke to death. | .67 (.02) | | |
| 3 | It scares me when my heart beats rapidly. ^a | .66 (.02) | | |
| 14 | When my thoughts seem to speed up, I worry that I might be going crazy. | | .87 (.01) | |
| 18 | When my mind goes blank, I worry there is something terribly wrong with me. | | .84 (.01) | |
| 10 | When I feel "spacey" or spaced out I worry that I may be mentally ill. | | .83 (.02) | |
| 16 | When I have trouble thinking clearly, I worry that there is something wrong with me. | | .83 (.01) | |
| 2 | When I cannot keep my mind on a task, I worry that I might be going crazy. ^a | | .77 (.02) | |
| 5 | It scares me when I am unable to keep my mind on a task. ^a | | .62 (.02) | |
| 9 | I worry that other people will notice my anxiety. | | | .85 (.01) |
| 6 | When I tremble in the presence of others, I fear what people might think of me. | | | .79 (.01) |
| 11 | It scares me when I blush in front of people. | | | .75 (.02) |
| 13 | When I begin to sweat in a social situation, I fear people will think negatively of me. | | | .70 (.02) |
| 17 | I think it would be horrible for me to faint in public. | | | .59 (.02) |
| 1 | It is important for me not to appear nervous. ^a | | | .54 (.02) |

Note. Factor model was simultaneously fitted to the samples of women and men, matching loadings, item errors, and factor correlations. ASI–3 = Anxiety Sensitivity Index–3.

^a Items from the original Anxiety Sensitivity Index.

simultaneously fitted to all the samples in this article (i.e., Subsample 1 and the six replication samples, without splitting gender into distinct groups). In this multigroup analysis, the degree of fit was evaluated for a model that fitted exactly the same factor structure to each sample (i.e., groups were matched on the size of factor loadings, factor correlations, and error variances). It was predicted that the three-factor model would have a good fit to the data, regardless of the language in which the scale was adminis-

tered and regardless of the patient status of the samples (clinical vs. nonclinical participants). In addition, it was predicted that the degree of fit for the three-factor model would be superior to the fit of the one- and two-factor models. It was also predicted that the three-factor model of the ASI–3 would show incremental factorial validity over the three-factor model of the original ASI. That is, the three-factor model for the ASI–3 was predicted to have better fit indices than the corresponding model from the ASI, because the

Table 2
Study 1: Polychoric Correlations Among ASI–3 Items for Subsample 1 (Pooled Across the Samples of Women and Men; n = 2,361)

| ASI–3 item no. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 2 | .27 | — | | | | | | | | | | | | | | | |
| 3 | .43 | .41 | — | | | | | | | | | | | | | | |
| 4 | .26 | .57 | .51 | — | | | | | | | | | | | | | |
| 5 | .20 | .52 | .26 | .43 | — | | | | | | | | | | | | |
| 6 | .29 | .47 | .31 | .45 | .39 | — | | | | | | | | | | | |
| 7 | .20 | .41 | .44 | .46 | .33 | .42 | — | | | | | | | | | | |
| 8 | .21 | .45 | .46 | .50 | .35 | .36 | .51 | — | | | | | | | | | |
| 9 | .54 | .41 | .40 | .41 | .38 | .65 | .40 | .34 | — | | | | | | | | |
| 10 | .30 | .65 | .41 | .55 | .51 | .46 | .39 | .52 | .52 | — | | | | | | | |
| 11 | .38 | .45 | .43 | .43 | .33 | .52 | .31 | .30 | .62 | .54 | — | | | | | | |
| 12 | .29 | .42 | .55 | .55 | .33 | .38 | .50 | .59 | .44 | .54 | .43 | — | | | | | |
| 13 | .31 | .36 | .25 | .36 | .38 | .62 | .38 | .32 | .56 | .43 | .52 | .36 | — | | | | |
| 14 | .30 | .65 | .40 | .61 | .48 | .49 | .41 | .51 | .52 | .75 | .51 | .55 | .47 | — | | | |
| 15 | .08 | .39 | .29 | .43 | .39 | .45 | .61 | .42 | .38 | .49 | .36 | .45 | .39 | .49 | — | | |
| 16 | .39 | .62 | .48 | .57 | .52 | .48 | .35 | .41 | .56 | .65 | .49 | .50 | .38 | .74 | .37 | — | |
| 17 | .45 | .23 | .35 | .31 | .16 | .40 | .27 | .22 | .59 | .30 | .39 | .33 | .37 | .35 | .22 | .42 | — |
| 18 | .29 | .68 | .38 | .58 | .56 | .53 | .42 | .50 | .48 | .67 | .49 | .50 | .42 | .69 | .56 | .70 | .35 |

Note. See Table 1 for item wording. Given the large sample size, all correlations are significant at $p < .001$ (two-tailed). ASI–3 = Anxiety Sensitivity Index–3.

former but not the latter was designed to be a multidimensional scale.

Method

Participants. As mentioned in Study 1, the U.S.–Canadian participants (Subsample 2; $n = 2,359$) were mostly undergraduate students. The mean age was 19.5 years ($SD = 3.7$), and 67% were women. The other nonclinical participants were undergraduates recruited from the Université de Lyon II (France; $n = 701$, M age = 20.4 years, $SD = 4.2$, 89% women), University of Mexico (Mexico; $n = 418$, M age = 22.2 years, $SD = 1.8$, 41% women), University of Groningen (the Netherlands; $n = 536$, M age = 20.8 years, $SD = 2.4$, 58% women), and the Universidad Nacional de Educación a Distancia (Spain; $n = 480$, M age = 26.4 years, $SD = 7.9$, 67% women). For the clinical sample (from the United States and Canada), the sample size was 390, mean age was 36.2 years ($SD = 12.0$), and 65% were women.

Ethnicity data were formally collected only for the U.S.–Canadian samples. This was due to the archival nature of the study, in which data were pooled across sites that differed in the number of demographic variables they assessed. For U.S.–Canadian Subsample 2, the majority of participants were White (62%), with the remainder being African American (19%), Asian (7%), Hispanic (3%), or other (9%). For the clinical sample, the majority were White (95%), with the remainder being African American (1%), Asian (2%), Hispanic (1%), or other (2%). The ethnicity of the other samples largely reflected the country in which the data were collected; participants from France, the Netherlands, and Spain were mostly White, and those from Mexico were primarily Hispanic.

For the clinical sample, primary diagnoses (in terms of severity) were panic disorder ($n = 143$), obsessive–compulsive disorder ($n = 104$), social anxiety disorder ($n = 38$), generalized anxiety disorder ($n = 30$), specific phobia ($n = 29$), hypochondriasis ($n = 15$), posttraumatic stress disorder ($n = 9$), trichotillomania ($n = 5$), major depressive disorder ($n = 3$), behavioral medicine conditions (e.g., irritable bowel syndrome, pain disorder, Raynaud's phenomenon; $n = 9$), and miscellaneous other conditions (e.g., Tourette's syndrome, kleptomania, adjustment disorder; $n = 5$). The proportion of patients taking psychotropic medication, assessed only at the University of British Columbia (UBC) site, was 48%. Medications were mostly benzodiazepines (e.g., lorazepam), selective serotonin reuptake inhibitors (e.g., fluoxetine), or tricyclic antidepressants (e.g., imipramine).

Participants who served in Study 1 (Subsample 1) were also included in those analyses in which the factor solutions for the replication samples were compared with those of Subsample 1.

Measures and procedure. Participants completed a short questionnaire assessing demographic features. U.S. participants (clinical and nonclinical) completed the 36-item ASI–R. All other participants completed the 42-item ASI–R. Participants completed the measures in classroom or individual settings.

All measures were administered in English, except for those administered to the samples in France, the Netherlands, Mexico, and Spain, in which cases the measures were translated into either French, Dutch, or Spanish. Translations were consistent with contemporary guidelines and practices (Butcher & Pancheri, 1976; Geisinger, 1994). Translators were doctoral-level psychologists

who were native speakers of the dominant language of a given country and also fluent in English. Translators were knowledgeable about the culture in which the scales were to be administered and were also familiar with the nature and assessment of AS. The scales were independently back-translated to ensure accuracy. All translators had previous experience in translating scales.

The clinical sample was recruited from cognitive–behavioral outpatient programs specializing primarily in anxiety disorders at UBC Hospital ($n = 155$) and the Mayo Clinic Anxiety/Obsessive–Compulsive Disorder (OCD) program ($n = 235$). All patients were referred by physician. Patients were diagnosed with an unstructured interview by the referring physician, who forwarded the diagnostic information in a referral letter. When patients attended the cognitive–behavioral outpatient programs, they were re-diagnosed with structured or unstructured clinical interviews according to *Diagnostic and Statistical Manual of Mental Disorders (DSM–IV)* criteria (American Psychiatric Association, 2000), with all diagnoses reviewed for accuracy in weekly staff meetings. Sixty-two percent of UBC patients were diagnosed with the Anxiety Disorders Interview Schedule for *DSM–IV* (DiNardo, Brown, & Barlow, 1994), and 60% of patients at the Mayo Clinic were assessed with either the Structured Clinical Interview for *DSM–IV* (First, Spitzer, Gibbon, & Williams, 1996) or the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998). The remaining patients from both sites were diagnosed with an unstructured clinical interview. Diagnosticians were either doctoral-level psychologists or predoctoral interns working under supervision. The reliability of diagnostic procedures has been previously established for the UBC and Mayo clinics (e.g., Taylor et al., 1996). Of the 390 clinical participants, 358 completed the measures before the commencement of cognitive–behavioral treatment. The remaining 32 completed the measures after they had received a full or partial course of therapy. Inclusion of both types of patients increased the range of scores for the factor analyses.

Statistical methods. The confirmatory factor analytic procedures in this study were essentially the same as those for the first study; polychoric correlations with asymptotic weights and robust weighted least-squares were used, and error terms were not permitted to be correlated. Given that this was a replication study, model modification indices were not used. The fit indices and criteria for interpreting these indices were the same as those used in Study 1. For the ASI–3, the 18-item three-factor model was tested in a multigroup (women vs. men) analysis for Subsample 2. This model was tested in single-group analyses for each of the other replication samples. The model was also tested in a multigroup analysis, which was conducted across all seven samples (not split by gender). The latter analysis was a highly stringent test of the replicability of the ASI–3 factor model; it required that factor correlations, factor loadings, and error variances were matched across samples. These analyses were then repeated for the one- and two-factor models of the ASI–3.

For the analyses comparing the ASI–3 with the ASI, the Canadian participants from Subsamples 1 and 2 were included ($n = 478$) because they—like the French, Dutch, Mexican, and Spanish samples—completed the 42-item ASI–R, which contains items that form the ASI–3 and ASI. U.S. participants from Subsamples 1 and 2 completed the 36-item ASI–R, which does not contain all the ASI items. Similarly, the comparison of ASI–3 and ASI was limited to Canadian patients ($n = 155$) because they were the only

patients who completed the 42-item version of the ASI-R. For the comparison of ASI-3 and ASI, the latter was scored to yield subscales measuring physical, cognitive, and social concerns. The selection of items for the three-factor model for the ASI was based on previous research (Zinbarg et al., 1997). The items composing each ASI subscale were as follows: Physical Concerns (3, 4, 6, 8, 9, 10, 11, 14), Cognitive Concerns (2, 12, 15, 16), and Social Concerns (1, 5, 7, 13).

Results and Discussion

For the 18 ASI-3 items, the percentage of missing data for a given item ranged from 0% to 0.4% for Subsample 2. The corresponding ranges for the other samples were as follows: clinical (0% to 0.8%), France (0%), Mexico (0% to 0.5%), the Netherlands (0% to 0.6%), and Spain (0% to 0.4%). The ranges of item-level missing data for the ASI were as follows: Canadian participants from Subsamples 1 and 2 (0% to 0.4%), Canadian participants from the clinical sample (0% to 1.9%), France (0%), Mexico (0% to 1.0%), the Netherlands (0% to 0.4%), and Spain (0% to 0.6%).

Table 3 shows the main results of the confirmatory factor analyses for the replication samples. With the exception of one sample—the clinical sample—the three-factor model had a good fit to the data for all fit indices in all analyses. For the clinical sample, the model had a good fit on three out of four indices; the exception was the RMSEA.

A series of analyses were conducted to determine why the RMSEA did not produce an acceptable goodness of fit for the clinical sample. RMSEA was not improved when eight multivariate outliers were removed from the data set. RMSEA also did not vary across the site of recruitment of the clinical samples (Mayo Clinic vs. UBC). We also examined the model modification indices of the three-factor model for the clinical sample. These indices indicated that no paths from factors to items should be added or deleted. This is consistent with other findings reported in this article, in which it was found that, for all the samples including the clinical sample, the three-factor solution had a better degree of fit than the one- and two-factor solutions. Model modification indices for the three-factor model suggested that goodness of fit would be improved if item uniqueness terms were allowed to be correlated. This could indicate that more factors should be specified for the clinical sample. To investigate this possibility, we conducted an exploratory factor analysis, using principal axis factor analysis and oblique rotation. Multiple rules were used to determine the number of factors to extract: the eigenvalue greater than one rule, visual inspection of the scree plot, and parallel analysis. Each indicated a three-factor solution. The pattern of loadings indicated good simple structure, with every item of a given ASI-3 subscale having strong (>.50) loadings on only one factor and small (<.20) loadings on the other factors. In other words, the results clearly supported a factor solution defined by physical, cognitive, and social concerns. This suggests that the lack of fit on the RMSEA was not because the type and number of factors for the clinical sample were different from the factor structure in the other samples.

It is possible that the goodness of fit on the RMSEA for the clinical sample may have been improved if we had been able to control for extraneous variables contributing correlations among the item uniqueness terms. Extraneous influences may include psychotropic medications and other treatments, which have been shown to influence AS scores (Taylor, 1999). A small proportion

(8%) of the clinical sample was assessed after partial or full cognitive-behavioral treatment, whereas the remainder was assessed prior to such treatment. RMSEA results did not change when we controlled for the effects of this form of treatment. Unfortunately, however, we were unable to control for the effects of psychotropic medication. Medication data were not available for the larger (Mayo Clinic) site. For the UBC site, 48% of patients were on medication, although data on the dose and duration of medication were not available. Further research is needed to investigate the factor structure of the ASI-3 in clinical samples, particularly research that controls for extraneous variables influencing AS scores such as psychotropic medication.

For the seven-group solution, in which the three-factor model was simultaneously fitted to Subsample 1 and the six replication samples, Table 3 shows that the model had a good fit for three of the four fit indices. Again, the exception was the RMSEA. It is noteworthy, however, that exactly the same factor model—in terms of loadings, factor correlations, and error variances—provided a good fit to all seven samples for the majority of the fit indices.

The results in Table 3 show that, across the various samples, the three factors tended to be highly correlated with one another. When the ASI-3 subscales were computed as the unit-weighted sum of their items, the correlations among subscales tended to be much lower than the correlations among their corresponding factors. For each replication sample, the ranges of correlations among unit-weighted subscales were as follows: Subsample 2 (.50 to .59), clinical (.41 to .53), France (.26 to .43), Mexico (.59 to .63), the Netherlands (.44 to .54), and Spain (.37 to .54).

Fit indices for the one- and two-factor models of the ASI-3 for the replication samples and for the seven-group solution were also examined and compared with the three-factor model. Recall from Study 1 that the two-factor model combines the items assessing cognitive and social concerns into a single factor, with the items assessing physical concerns forming a separate factor. In the one-factor model, all items load on a single factor. The results (appearing in tables in the appendix, available online) indicated that, for most analyses, the one- and two-factor models poorly fitted the data and that the goodness of fit was superior for the three-factor model.

A goal in developing the ASI-3 was to develop a scale with a more robust (replicable) factor structure than the original ASI. Accordingly, we expected that the degree of fit of the three-factor model of the ASI-3 should be better than that of the ASI. Table 4 shows the main results relevant to this prediction for participants who had completed both scales. Three findings in the table are noteworthy. First, the fit indices for the ASI-3 tend to be better than those of the ASI; 26 of 28 indices in the table indicate a good fit for the ASI-3, whereas a good fit is indicated for 20 of 28 indices for the ASI. Second, the magnitude of the fit indices is generally better (i.e., higher for the CFI and TLI, and lower for the SRMR and RMSEA) for the ASI-3 compared with the ASI. This result holds for 27 of the 28 indices. Third, the ASI-3 and ASI had nonoverlapping RMSEA CIs in five of seven instances. The latter indicates that, for most analyses, the three-factor model of the ASI-3 had a significantly better fit than the corresponding model of the ASI.

A further series of confirmatory factor analyses examined the one- and two-factor solutions of the ASI, with these models defined in the same way that they were for the ASI-3. The results (appearing in the

Table 3
Study 2: Fit Indices and Correlations Among Factors for the Three-Factor Model of the ASI-3 for Each Replication Sample and for All Seven Samples (Scale Construction and Replication Samples)

| Sample | Range of correlations among factors | SRMR | RMSEA (and 90% CI) | CFI | TLI | $\chi^2(df)$ |
|--|-------------------------------------|-------------|--------------------------|-------------|-------------|------------------|
| U.S.–Canadian Subsample 2: two-group analysis (matching women and men to the same factor model; $n = 2,359$) | .69–.78 | .061 | .055 (.052, .058) | .986 | .986 | 1,397.13 (303) |
| U.S.–Canadian clinical sample ($n = 390$) | .45–.61 | .061 | .090 (.083, .098) | .965 | .959 | 550.43 (132) |
| France ($n = 701$) | .32–.59 | .054 | .043 (.036, .049) | .985 | .982 | 298.65 (132) |
| Mexico ($n = 418$) | .74–.86 | .053 | .040 (.031, .049) | .992 | .991 | 220.52 (132) |
| The Netherlands ($n = 536$) | .59–.70 | .051 | .032 (.023, .040) | .994 | .994 | 202.90 (132) |
| Spain ($n = 480$) | .43–.68 | .061 | .056 (.049, .064) | .985 | .983 | 330.95 (132) |
| 7-group solution: fitting the factor model (factor loadings, item errors, and factor correlations) to Subsample 1 and to all 6 replication samples (not split by gender; $N = 7,245$) | .70–.79 | .101 | .078 (.076, .079) | .962 | .965 | 8,355.19 (1,158) |

Note. ASI-3 = Anxiety Sensitivity Index-3; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CI = confidence interval; CFI = comparative fit index; TLI = Tucker-Lewis index; Boldface type = fit index indicating that the three-factor model provided a good fit to the data (SRMR \leq .08, RMSEA \leq .06, CFI \geq .95, or TLI \geq .95).

appendix, available online) indicated that the three-factor model of the ASI had a better fit than the one- and two-factor models. In other words, of the factor models examined in the present article, the best-fitting (three-factor) model for the ASI tended to have a poorer fit to the data than did the corresponding model for the ASI-3. Overall, the findings support the factorial validity of the ASI-3 and indicate that this scale has stronger factorial validity than the ASI.

Study 3: Reliability as Internal Consistency

Cronbach’s coefficient alpha was computed for each of the ASI-3 subscales. Coefficients greater than or equal to .70 were defined as acceptable, and those greater than or equal to .80 were defined as good (Nunnally & Bernstein, 1994). It was predicted that these values would be as good as or better than the corre-

Table 4
Study 2: ASI-3 Versus ASI: Fit Indices and Correlations Among Factors for the Three-Factor Model

| Sample | ASI-3 | | | | | | ASI | | | | | |
|--|-------------------------------------|-------------|-------------------------|-------------|-------------|----------------|-------------------------------------|-------------|--------------------|-------------|-------------|----------------|
| | Range of correlations among factors | SRMR | RMSEA (and 90% CI) | CFI | TLI | $\chi^2(df)$ | Range of correlations among factors | SRMR | RMSEA (and 90% CI) | CFI | TLI | $\chi^2(df)$ |
| Canadian nonclinical ($n = 478$) ^a | .63–.75 | .055 | .049 (.041,.049) | .989 | .987 | 282.83 (132) | .65–.73 | .076 | .081 (.074,.090) | .960 | .952 | 420.47 (101) |
| Canadian clinical ($n = 155$) ^a | .47–.73 | .064 | .074 (.060,.089) | .979 | .976 | 244.38 (132) | .62–.81 | .079 | .093 (.077,.108) | .966 | .959 | 234.02 (101) |
| France ($n = 701$) | .32–.59 | .054 | .043 (.036,.049) | .985 | .982 | 298.65 (132) | .57–.72 | .066 | .063 (.056,.070) | .954 | .946 | 380.95 (101) |
| Mexico ($n = 418$) | .74–.86 | .053 | .040 (.031,.049) | .992 | .991 | 220.57 (132) | .58–.83 | .068 | .063 (.054,.072) | .973 | .968 | 268.56 (101) |
| The Netherlands ($n = 536$) | .59–.70 | .051 | .032 (.023,.040) | .994 | .994 | 202.90 (132) | .68–.82 | .071 | .065 (.057,.073) | .978 | .974 | 330.04 (101) |
| Spain ($n = 480$) | .43–.68 | .061 | .056 (.049,.064) | .985 | .983 | 330.95 (132) | .56–.63 | .066 | .065 (.055,.074) | .976 | .972 | 264.56 (101) |
| 6-group solution: fitting all the above samples simultaneously to 3-factor model ($n = 2,768$) | .51–.69 | .090 | .047 (.044,.050) | .987 | .988 | 1,973.64 (987) | .64–.73 | .103 | .068 (.065,.071) | .964 | .967 | 2,456.91 (781) |

Note. Boldface type indicates fit index showing that the three-factor model provided a good fit to the data (SRMR \leq .08, RMSEA \leq .06, CFI \geq .95, or TLI \geq .95). ASI-3 = Anxiety Sensitivity Index-3; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis index; CI = confidence interval.

^a Of the Canadian and U.S. samples, only the Canadian samples completed all the ASI items, so the U.S. samples were not used in the comparison of the ASI-3 and ASI factor models.

sponding values of the ASI subscales. This is because the ASI was not designed to be a multidimensional measure, and the Cognitive and Social Concerns subscales of the ASI each contain only four items. Although such short scales could, in principle, yield acceptable alpha values, this would be most likely to occur if the items had high content validity. Given the concerns raised earlier in this article about the contents of some of the ASI items, it was expected that four-item subscales would not be sufficiently long to yield acceptable alphas.

Method

Participants. Participants were those used in Studies 1 and 2. Samples were those in which both ASI-3 and ASI data were available, that is, the Canadian clinical and nonclinical samples and the samples from France, Mexico, the Netherlands, and Spain.

Measures and procedure. These were the same as in Studies 1 and 2.

Statistical methods. Coefficient alphas for the ASI-3 and ASI subscales were compared by computing a 95th percentile CI around each alpha, according to the methods described by Duhaček and Iacobucci (2004).

Recall that the ASI-3 and ASI can be conceptualized as having hierarchic structures in which the three lower order factors load on a common, general AS factor. Total scores on the ASI-3 and ASI are useful for assessing the general factor, providing that the general factor accounts for a substantial proportion of variance. To investigate this for each of the six samples in which data on the ASI-3 and ASI were available, we calculated the proportion of variance due to the general factor (known as ω_H ; Zinbarg, Yovel, Revelle, & McDonald, 2006), and we also calculated the proportion of variance due to the lower order factors, separate from the variance due to the general factor. This was done by performing a Schmid-Leiman analysis of the loadings from the confirmatory factor analyses reported earlier in this article, using the procedures described by Zinbarg et al. (2006).

Results and Discussion

The proportions of missing data were the same as for Study 2. Table 5 shows the coefficient alphas for the ASI-3 and ASI subscales, along with the corresponding 95th percentile CIs. The values for the ASI-3 were all in the range considered to be acceptable or good (Nunnally & Bernstein, 1994), which supports the intended use of these subscales as research instruments. The ASI-3 and ASI did not differ in the magnitude of coefficients for the Physical Concerns subscales. However, the ASI-3 generally had significantly larger coefficients than did the ASI for the Cognitive Concerns and Social Concerns subscales.

For each of the six samples, the proportion of variance due to the general factor (ω_H) of the ASI-3 was consistently slightly higher than that of the ASI (respectively, $M = 0.36$, $SD = 0.06$, and $M = 0.33$, $SD = 0.06$). For each sample and scale, the total proportion of variance from all three lower order factors was calculated, after controlling for variance due to the general factor. Across all six samples, the proportions of variance for ASI-3 were consistently higher than those for the ASI (respectively, $M = 0.40$, $SD = 0.06$, and $M = 0.33$, $SD = 0.06$). In summary, results indicate that for the ASI-3, about 36% of the variance in scale scores is due to a

general AS factor, and 40% of variance is due to the lower order factors (i.e., 76% total explained variance and 24% error variance). The corresponding figures for the ASI were 33% and 33% (66% explained variance, 34% error variance). These results support the use of total and subscale scores of the ASI-3 and also show that scores on the ASI-3 contain less error variance than those on the ASI.

Study 4: Convergent, Discriminant, and Criterion-Related Validities

Convergent validity of the ASI-3 was examined by intercorrelating the subscales of the ASI-3 and ASI. It was predicted that similar subscales (e.g., ASI-3 and ASI Physical Concerns subscales) would be highly correlated ($r \geq .50$; i.e., large correlations according to Cohen, 1988). Discriminant validity for a measure of a given construct would be supported when the measure is more highly correlated with similar or theoretically related constructs than when it is correlated with dissimilar or theoretically unrelated constructs. Accordingly, it was predicted that similar subscales (e.g., ASI-3 and ASI Physical Concerns) would be more highly correlated than dissimilar subscales (e.g., ASI-3 Physical Concerns correlated with ASI Cognitive Concerns or with ASI Social Concerns). The ASI-3 and ASI contain five overlapping items, which would inflate correlations between ASI-3 and ASI subscales. Therefore, in conducting tests of convergent and discriminant validity, we removed the overlapping items from the ASI. We did this because the ASI-3 subscales were the focus of investigation, so it would be inappropriate to eliminate the overlapping items from the ASI-3.

Regarding criterion-related (known groups) validity, we tested predictions about how scores on the ASI-3 subscales would differ across different groups. U.S.–Canadian data from students (i.e., nonclinical controls [NC]) and pretreatment data from the four largest clinical groups—panic disorder (PD), OCD, social anxiety disorder (SANX), and generalized anxiety disorder (GAD)—enabled us to test a prediction for each of the ASI-3 subscales. The first prediction had two parts: (a) that the PD group would be associated with higher scores on the Physical Concerns subscales than would all other groups and (b) that the remaining clinical groups would have higher scores than would the NC group. These predictions were based on prior theory and research implicating the importance of heightened physical concerns in PD and other anxiety disorders (Taylor, 1999). Research indicates that physical concerns are elevated in these disorders compared with NCs, with PD tending to be associated with the highest scores (Taylor, 1999).

The second prediction also had two parts: (a) SANX, compared with all other groups, would be associated with higher scores on social concerns, and (b) the other clinical groups would score higher on social concerns than would the NCs. The rationale was that SANX, compared with other anxiety disorders and NCs, is characterized by strong fear of negative evaluation (American Psychiatric Association, 2000). Therefore, SANX should be associated with higher scores than the other groups. In turn, GAD, OCD, and PD should be associated with higher scores than NCs, because social fears tend to be greater in these groups compared with NCs (e.g., Rapee, Sanderson, & Barlow, 1988).

The third prediction was that each of the clinical groups would score higher than NCs on the Cognitive Concerns subscale. The

Table 5
Study 3: Subscale Reliability for the ASI-3 Versus ASI: Coefficient Alphas and Their 95th Percentile Confidence Intervals

| Sample | Physical Concerns | | Cognitive Concerns | | Social Concerns | |
|--|-------------------|----------------|-----------------------|----------------|-----------------------|----------------|
| | ASI-3 | ASI | ASI-3 | ASI | ASI-3 | ASI |
| Canadian nonclinical (<i>n</i> = 478) | .79 (.76, .82) | .83 (.81, .86) | .83 (.81, .86) | .77 (.74, .80) | .78 (.75, .81) | .53 (.46, .60) |
| Canadian clinical (<i>n</i> = 155) | .86 (.83, .90) | .89 (.86, .92) | .91 (.89, .93) | .84 (.80, .88) | .86 (.83, .89) | .60 (.50, .70) |
| France (<i>n</i> = 701) | .76 (.73, .79) | .76 (.73, .79) | .79 (.77, .82) | .68 (.64, .72) | .75 (.72, .78) | .41 (.34, .49) |
| Mexico (<i>n</i> = 418) | .83 (.81, .86) | .81 (.78, .84) | .82 (.79, .85) | .77 (.73, .81) | .73 (.69, .77) | .53 (.46, .60) |
| The Netherlands (<i>n</i> = 536) | .80 (.77, .82) | .83 (.81, .85) | .81 (.79, .84) | .80 (.77, .83) | .76 (.72, .79) | .44 (.36, .52) |
| Spain (<i>n</i> = 480) | .84 (.81, .86) | .82 (.80, .85) | .87 (.85, .89) | .81 (.78, .84) | .84 (.81, .86) | .66 (.61, .71) |

Note. Confidence intervals are in parentheses. For pairwise comparisons between ASI-3 and ASI subscales, for a given sample and content domain (e.g., Physical Concerns), the significantly larger alpha is in boldface type, as indicated by nonoverlapping confidence intervals. ASI-3 = Anxiety Sensitivity Index-3.

rationale was that all four of the anxiety disorders have been found to be associated with particular forms of cognitive concern. This prediction was based on research that used measures of AS in PD studies and on research on other disorders with the use of instruments similar to the Cognitive Concerns subscale (e.g., measures of beliefs about the importance of controlling one’s thoughts in OCD research). To illustrate the various forms of cognitive concerns, researchers have found PD to be associated with strong fears of cognitive phenomena (e.g., derealization), associated with beliefs that these phenomena have catastrophic consequences, such as permanent mental incapacitation (Taylor, 1999). Similarly, SANX is associated with fears that one may not be able to perform adequately in social situations, such as fears that one’s mind will go blank (Clark & Wells, 1995). This is a situationally specific form of cognitive concern, which should contribute to high scores on the Cognitive Concerns subscale. Research has shown that people with OCD, compared with control participants, are more likely to overestimate the dangerousness of cognitive dyscontrol, which is consistent with contemporary cognitive models of OCD (Frost & Steketee, 2002), which propose that OCD arises, in part, from such distorted beliefs about the harmful consequences of cognitive dyscontrol. Similarly, theory and research indicate that GAD is associated with “meta-worry,” that is, worry about the deleterious effects of uncontrollable worry (Wells, 2005).

Empirical research has shown that one particular facet of AS—physical concerns—is especially elevated in PD compared with most other anxiety disorders (although it is not clear that people with PD score any higher than those with posttraumatic stress disorder; Taylor, 1999). It is unclear whether PD is associated with higher scores, compared with other disorders, on the Cognitive and Social Concerns subscales. It is for these reasons that we chose not to include an a priori prediction about group differences in ASI-3 total score. Previous research on this issue has been based largely on the ASI. Total scores on that scale are weighted toward the assessment of physical concerns, which represent half of the items of the scale. Therefore, previous findings that PD is associated with the highest ASI total score (Taylor, 1999) might simply reflect the ASI’s overemphasis on physical concerns. We also did not compare the known groups validity of the ASI-3 with that of the ASI. This was because there were too few clinical participants who completed both measures (*n* = 155), which meant that the corresponding clinical groups would be too small for analysis.

Method

Participants. Participants were those used in Studies 1 and 2. For the analyses involving group comparisons, the samples used are shown in Table 6. For the other analyses, samples were those

Table 6
Study 4: Means (and Standard Deviations) on the ASI-3 Subscales

| Sample | Physical Concerns | Cognitive Concerns | Social Concerns | Total score |
|--|-------------------|--------------------|-----------------|-------------|
| Nonclinical groups | | | | |
| U.S.–Canada: women (<i>n</i> = 3,153) | 4.3 (4.2) | 2.6 (3.8) | 5.9 (4.7) | 12.8 (10.5) |
| U.S.–Canada: men (<i>n</i> = 1,567) | 3.9 (4.2) | 2.8 (3.8) | 6.0 (4.8) | 12.8 (10.8) |
| U.S.–Canada: total (<i>n</i> = 4,720) | 4.2 (4.2) | 2.7 (3.8) | 5.9 (4.7) | 12.8 (10.6) |
| France (<i>n</i> = 701) | 5.0 (4.0) | 2.8 (3.4) | 8.5 (4.8) | 16.4 (9.1) |
| Mexico (<i>n</i> = 418) | 5.5 (4.8) | 3.5 (4.1) | 6.1 (4.3) | 15.2 (11.3) |
| The Netherlands (<i>n</i> = 536) | 2.7 (3.2) | 1.7 (2.8) | 5.7 (4.0) | 10.7 (8.1) |
| Spain (<i>n</i> = 480) | 4.5 (3.9) | 2.8 (3.7) | 6.9 (4.7) | 14.2 (9.8) |
| Selected clinical groups (pretreatment) ^a | | | | |
| Panic disorder (<i>n</i> = 120) | 11.3 (6.7) | 9.0 (6.4) | 12.3 (5.8) | 32.6 (14.3) |
| Obsessive-compulsive disorder (<i>n</i> = 102) | 8.3 (6.2) | 7.7 (6.0) | 10.3 (6.7) | 26.3 (16.8) |
| Social anxiety disorder (<i>n</i> = 38) | 6.2 (4.5) | 7.9 (6.1) | 17.3 (4.8) | 31.4 (11.9) |
| Generalized anxiety disorder (<i>n</i> = 30) | 8.1 (5.3) | 8.9 (7.4) | 10.5 (7.0) | 27.5 (16.5) |

^a Selected for inclusion in tests of criterion-related (known groups) validity. ASI-3 = Anxiety Sensitivity Index-3.

in which both ASI-3 and ASI data were available, that is, the Canadian clinical and nonclinical samples, as well as the samples from France, Mexico, the Netherlands, and Spain.

Measures and procedure. These were the same as in Studies 1 and 2.

Statistical methods. For the correlational and group comparisons analyses, items were summed (unit-weighted) to form ASI-3 and ASI subscales. All correlations were Pearson product-moment correlations unless stated otherwise. As mentioned, items that overlapped with ASI-3 were omitted from the ASI subscales. To disentangle the evaluation of convergent and discriminant validity from effects due to subscale reliability, we disattenuated the correlations in the validity tests for unreliability (Nunnally & Bernstein, 1994). That is, each correlation between a pair of measures was divided by the square root of the product of the reliability (α) of each measure. Tests of differences between correlations were computed according to the methods described by Meng, Rosenthal, and Rubin (1992).

Regarding the tests of group differences in ASI-3 subscale scores (i.e., known groups validity), conventional multiple comparisons (e.g., Tukey's comparisons) were inappropriate because our data violated the assumptions of equal sample sizes and homogeneous variances. A further problem was that conventional multiple comparisons yield intransitive results. That is, they often yield ambiguous or logically impossible results. To illustrate, for a three-group comparison of three means {1,2,3}, ordered from highest to lowest, multiple comparisons such as the Tukey's test often yield results such as the following: $1 = 2$, $2 = 3$, and yet $1 \neq 3$. This intransitivity adds confusion to the interpretation of the findings. To circumvent all of these problems, we used the paired-comparisons information criterion (PCIC; Dayton, 2003), which computes the Akaike Information Criterion (AIC) for all logically possible subsets of groups. For example, for three groups, ordered from highest to lowest scores, there would be the following "models," in which the commas indicate that a group differs from other groups: {123} (null model in which there are no group differences), {1,2,3}, {12,3}, and {1,2,3}. The model with the smallest AIC represents the best-fitting model. The PCIC does not involve computing statistical significance tests, so the control of Type I and II error is not an issue. The method does not produce intransitive results and does not require that sample sizes or variances be equal across groups. The AIC has a slight bias for selecting more complicated models than the true model, and so an omnibus test (e.g., multivariate analysis of variance) is used to test the null hypothesis of no group differences. Monte Carlo studies indicate that the protected PCIC is superior to other multiple comparison procedures in correctly identifying patterns of group differences (Dayton, 2003).⁵

Results and Discussion

Preliminary analyses. The proportions of missing data were the same as for Study 2. For descriptive purposes, ASI-3 and ASI correlations with demographic variables were computed. Across six samples (Canadian students, Canadian patients, and samples from France, Mexico, the Netherlands, and Spain), the ASI-3 (and ASI) total score correlations with age ranged from $-.09$ to $.05$ ($-.11$ to $.08$). Polyserial correlations with gender (scored *women* = 1, *men* = 2) ranged from $-.16$ to $.07$ ($-.22$ to $.08$).

Correlations with education level,⁶ available only for the Canadian samples, ranged from $-.06$ to $.15$ ($-.06$ to $.12$). Correlations with ethnicity (*Non-White* = 1, *White* = 2), available only for the Canadian samples, ranged from $-.22$ to $-.12$ ($-.19$ to $-.12$). In summary, the patterns of correlations for the ASI-3 and ASI were similar to one another. Although some of the correlations were statistically significant because of the large sample sizes, all correlations were in the range that would be classified as small or trivial, according to Cohen's (1988) classification scheme. In other words, the ASI-3 performed similarly to the ASI; both were largely unrelated to demographic variables.

Convergent and discriminant validity. Table 7 shows that when the subscales were corrected for less-than-perfect reliability, each ASI-3 subscale measured essentially the same content domain (e.g., physical concerns) as its ASI counterpart, as indicated by correlations that approach unity. These large correlations support the convergent validity of the ASI-3 subscales. Correlations between similar subscales from the ASI-3 and ASI (e.g., Physical Concerns) were, for almost all analyses, significantly larger than the correlations between dissimilar subscales (e.g., ASI-3 Physical Concerns correlated with either ASI Cognitive or Social Concerns). This is shown by the results for the planned contrasts (Z scores) in Table 7 and provides supporting evidence of the discriminant validity of the ASI-3 subscales.

Criterion-related (known groups) validity. A multivariate analysis of variance was conducted as an omnibus test, in which the dependent variables were the three subscale scores on the ASI-3, and the independent variable (group) was defined by the five groups (PD, OCD, GAD, SANX, and NC). The group factor for the omnibus test was significant, Pillai $F(12, 15,012) = 67.74$, $p < .001$, $\eta^2 = .05$. The corresponding means (and standard deviations) appear in Table 6. The best-fitting results of the PCIC comparisons were as follows: for the ASI-3 Physical Concerns subscale, $PD > (OCD, GAD) > SANX > NC$ ($\eta^2 = .08$). For Cognitive Concerns, $(PD, GAD) > (OCD, SANX) > NC$ ($\eta^2 = .10$). For Social Concerns, $SANX > PD > (OCD, GAD) > NC$ ($\eta^2 = .09$). Thus, each of the three predictions was supported regarding the criterion-related validity of the ASI-3 subscales.

General Discussion

As researchers increasingly focus their attention on dimensions of AS, rather than simply looking at AS as a global construct, it is important to develop robust, psychometrically sound measures of

⁵ PCIC was used instead of planned contrasts because of the violation of the assumptions of equal sample size and homogeneity of variance. A further reason for not using planned contrasts was because the latter are performed in the context of null hypothesis significance testing, in which the purpose of the contrasts is to reduce the odds of Type II error by conducting comparisons only among particular pairs of means instead of among all possible pairs. Type II and Type I error rates are not relevant to PCIC (Dayton, 2003).

⁶ Education level was rated by respondents on the following scale: 1 = *Grade 6 or less*; 2 = *Grades 7–12, without graduating from high school*; 3 = *high school or equivalent*; 4 = *partial college*; 5 = *graduated from a 2-year college program*; 6 = *graduated from a 4-year college or university program*; 7 = *partial graduate or professional school*; and 8 = *completed graduate or professional school*.

Table 7

Study 4: Pearson Product–Moment Correlations Among ASI–3 and ASI Subscales, With (and Without) Correction for Attenuation

| Sample and subscale | ASI–3 and ASI | | | Planned contrasts among correlations that were corrected for attenuation | |
|--------------------------------|---------------|-----------|-----------|--|--------|
| | Physical | Cognitive | Social | Contrast | Z* |
| Canadian nonclinical (n = 478) | | | | | |
| Physical | .93 (.73) | .69 (.49) | .72 (.37) | Phys vs. Cog, Soc | 20.01 |
| Cognitive | .62 (.50) | .99 (.73) | .67 (.36) | Cog vs. Phys, Soc | 77.18 |
| Social | .62 (.49) | .76 (.53) | .99 (.54) | Soc vs. Phys, Cog | 72.15 |
| Canadian clinical (n = 155) | | | | | |
| Physical | .99 (.85) | .66 (.53) | .71 (.42) | Phys vs. Cog, Soc | 31.66 |
| Cognitive | .72 (.64) | .94 (.78) | .92 (.56) | Cog vs. Phys, Soc | 6.36 |
| Social | .59 (.50) | .74 (.60) | .99 (.67) | Soc vs. Phys, Cog | 46.41 |
| France (n = 701) | | | | | |
| Physical | .92 (.67) | .55 (.35) | .64 (.27) | Phys vs. Cog, Soc | 28.36 |
| Cognitive | .61 (.45) | .98 (.64) | .59 (.25) | Cog vs. Phys, Soc | 49.16 |
| Social | .47 (.34) | .54 (.34) | .99 (.52) | Soc vs. Phys, Cog | 106.70 |
| Mexico (n = 418) | | | | | |
| Physical | .92 (.72) | .74 (.57) | .74 (.40) | Phys vs. Cog, Soc | 17.25 |
| Cognitive | .77 (.60) | .93 (.71) | .65 (.35) | Cog vs. Phys, Soc | 20.39 |
| Social | .85 (.63) | .74 (.53) | .99 (.54) | Soc vs. Phys, Cog | 63.26 |
| The Netherlands (n = 536) | | | | | |
| Physical | .96 (.75) | .58 (.44) | .83 (.36) | Phys vs. Cog, Soc | 62.59 |
| Cognitive | .70 (.55) | .83 (.64) | .91 (.40) | Cog vs. Phys, Soc | 0.03 |
| Social | .59 (.45) | .63 (.47) | .99 (.48) | Soc vs. Phys, Cog | 206.48 |
| Spain (n = 480) | | | | | |
| Physical | .98 (.79) | .61 (.45) | .53 (.38) | Phys vs. Cog, Soc | 39.66 |
| Cognitive | .73 (.59) | .86 (.64) | .54 (.39) | Cog vs. Phys, Soc | 11.84 |
| Social | .59 (.47) | .55 (.41) | .92 (.65) | Soc vs. Phys, Cog | 21.44 |

Note. Item overlap between the ASI–3 and ASI was eliminated by omitting overlapping items from the ASI. ASI = Anxiety Sensitivity Index; Phys = Physical; Cog = Cognitive; Soc = Social.

* All $ps < .001$, apart from $Z = 0.03$, for which $p > .10$.

these dimensions. The present study developed and evaluated such a scale, which measures the three most widely replicated dimensions: physical, cognitive, and social concerns. This scale, called the ASI–3, had a stable three-factor structure across gender and across seven different samples, according to most fit indices. The findings were especially encouraging given that they were obtained from different populations (clinical vs. nonclinical) and from different countries and different language versions of the measures. The ASI–3 shares 5 of its 18 items with the ASI, with 1 to 2 overlapping items on each of the ASI–3's 6-item subscales. The overlap is a result of the scale construction process for the ASI–3, in which we selected, on the basis of content validity and psychometric properties, the best of the ASI items for inclusion in the ASI–3.

Although the ASI–3 is only two items longer than the original ASI, it tended to be a better measure of the AS dimensions, as assessed by reliability as internal consistency and by factorial validity. Evidence also supported the convergent, discriminant, and criterion-related (known groups) validities of the ASI–3. The findings suggest that the ASI–3 would be preferable to the ASI in studies of the dimensions of AS. The ASI–3 subscales may also be preferable to the ASI subscales in taxometric studies because of the greater reliability of the former. Subscales are commonly used as indicators in taxometric studies, and indicator reliability influences the taxometric validity of the indicators, that is, their ability to distinguish between taxonic groups (Meehl, 1992).

The multisample data set used in this article has both strengths and limitations. The sample was clinically and nationally diverse,

and to our knowledge it is the largest data set to have ever been used in AS research. However, the items for the ASI–3 were embedded in the larger pool of items forming the ASI–R, and so it remains to be seen whether the results of the present study generalize to situations in which the ASI–3 is administered as a stand-alone instrument. Nevertheless, it is noteworthy that the original ASI was also embedded among the ASI–R items, and the performance of the ASI in the present study resembled that found in many other studies (e.g., the identification of a three-factor solution; Taylor, 1999).

Further research is needed to more fully evaluate the validity of the ASI–3. Studies are required to test whether, across different types of samples, the ASI–3 subscales are more strongly correlated with theoretically related variables (e.g., anxiety-related variables) than with theoretically unrelated variables (e.g., extraversion or other traits that are not correlated with the tendency to experience negative emotions). Laboratory studies, such as symptom-provocation studies using carbon dioxide inhalation (which induces intense dyspnea), could also be conducted to test the convergent validity of the ASI–3. Studies of the ASI have shown that scores on this scale predict anxiety evoked by carbon dioxide inhalation (Taylor, 1999). The same is predicted for the ASI–3.

The ASI–3's test–retest reliability remains to be studied. Prospective studies are required to determine whether the ASI–3, like the ASI, predicts the risk of psychopathology such as panic attacks, and whether the prediction of psychopathology differs among the ASI–3 subscales. If the ASI–3 proves to be sensitive to treatment effects, then this short scale could be repeatedly admin-

istered during treatment studies to investigate the relative effects of treatments on the three dimensions of AS and to study the mediators and moderators of treatment-related changes in AS. Finally, further studies are needed to fully investigate the cross-cultural similarities and differences in AS.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Bernstein, A., Zvolensky, M. J., Kotov, R., Arrindell, W. A., Taylor, S., Sandin, B., et al. (2006). Taxonicity of anxiety sensitivity: A multinational analysis. *Journal of Anxiety Disorders, 20*, 1–22.
- Blais, M. A., Otto, M. W., Zuckerman, B. G., McNally, R. J., Schmidt, N. B., Fava, M., et al. (2001). The Anxiety Sensitivity Index: Item analysis and suggestions for refinement. *Journal of Personality Assessment, 77*, 272–294.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Butcher, J. N., & Pancheri, P. (1976). *Handbook of cross-national MMPI research*. Minneapolis, MN: University of Minnesota Press.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. Liebowitz, D. Hope, & F. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). New York: Guilford Press.
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dayton, C. M. (2003). Information criteria for pairwise comparisons. *Psychological Methods, 8*, 61–71.
- Deacon, B. J., Abramowitz, J. A., Woods, C. M., & Tolin, D. F. (2003). The Anxiety Sensitivity Index—Revised: Psychometric properties and factor structure in two nonclinical samples. *Behaviour Research and Therapy, 41*, 1427–1449.
- Dehon, C., Weems, C. F., Stickle, T. R., Costa, N. M., & Berman, S. L. (2005). A cross-sectional evaluation of the factorial invariance of anxiety sensitivity in adolescents and young adults. *Behaviour Research and Therapy, 43*, 799–810.
- DiNardo, P., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV*. New York: Graywind.
- Duhachek, A., & Iacobucci, D. (2004). Alpha's standard error (ASE): An accurate and precise confidence interval estimate. *Journal of Applied Psychology, 89*, 792–808.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured Clinical Interview for Axis I DSM-IV disorders—Patient edition*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Frost, R. O., & Steketee, G. (2002). *Cognitive approaches to obsessions and compulsions: Theory, assessment and treatment*. Oxford, England: Elsevier.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6*, 304–312.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453.
- 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.
- Jöreskog, K., & Sörbom, D. (2005). *LISREL 8.72* [manual and computer program]. Lincolnwood, IL: Scientific Software International.
- Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60*, 117–174.
- Meng, X.-L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111*, 172–175.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Peterson, R. A., & Reiss, S. (1992). *Anxiety Sensitivity Index revised manual*. Worthington, OH: IDS Publishing.
- Rapee, R. M., Sanderson, W. C., & Barlow, D. H. (1988). Social phobia features across the DSM-III-R anxiety disorders. *Journal of Psychopathology and Behavioral Assessment, 10*, 287–299.
- Reiss, S., & McNally, R. J. (1985). Expectancy model of fear. In S. Reiss & R. R. Bootzin (Eds.), *Theoretical issues in behavior therapy* (pp. 107–121). San Diego, CA: Academic Press.
- Roth, D. A., Coles, M. E., & Heimberg, R. G. (2002). The relationship between memories of childhood teasing and anxiety and depression in adulthood. *Journal of Anxiety Disorders, 16*, 149–164.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33.
- Stewart, S. H., Taylor, S., & Baker, J. M. (1997). Gender differences in dimensions of anxiety sensitivity. *Journal of Anxiety Disorders, 11*, 179–200.
- Taylor, S. (1999). *Anxiety sensitivity: Theory, research, and treatment of the fear of anxiety*. Mahwah, NJ: Erlbaum.
- Taylor, S., & Cox, B. J. (1998a). An expanded Anxiety Sensitivity Index: Evidence for a hierarchic structure in a clinical sample. *Journal of Anxiety Disorders, 12*, 463–484.
- Taylor, S., & Cox, B. J. (1998b). Anxiety sensitivity: Multiple dimensions and hierarchic structure. *Behaviour Research and Therapy, 36*, 37–51.
- Taylor, S., Koch, W. J., Woody, S., & McLean, P. (1996). Anxiety sensitivity and depression: How are they related? *Journal of Abnormal Psychology, 105*, 474–479.
- Wells, A. (2005). The metacognitive model of GAD: Assessment of meta-worry and relationship with DSM-IV generalized anxiety disorder. *Cognitive Therapy and Research, 29*, 107–121.
- Zinbarg, R. E., Barlow, D. H., & Brown, T. A. (1997). The hierarchical structure and general factor saturation of the Anxiety Sensitivity Index: Evidence and implications. *Psychological Assessment, 9*, 277–284.
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ω_H . *Applied Psychological Measurement, 30*, 121–144.
- Zvolensky, M. J., Arrindell, W. A., Taylor, S., Bouvard, M., Cox, B. J., Stewart, S. H., et al. (2003). Anxiety sensitivity in six countries. *Behaviour Research and Therapy, 41*, 841–859.

Received November 28, 2005

Revision received December 15, 2006

Accepted December 22, 2006 ■