Available online at www.sciencedirect.com
SCIENCE DIRECT®
Behaviour Research and Therapy 43 (2005) 337–355
BEHAVIOUR RESEARCH AND THERAPY
www.elsevier.com/locate/brt

A psychometric reanalysis of the Albany Panic and Phobia Questionnaire

Timothy A. Brown*, Kamila S. White, David H. Barlow

Center for Anxiety and Related Disorders, Boston University, 648 Beacon Street, 6th floor, Boston, MA 02215-2013, USA

Abstract

The psychometric properties of the 27-item Albany Panic and Phobia Questionnaire (APPQ) were evaluated in 1930 outpatients with *DSM-IV* anxiety and mood disorders. Although prior findings of a 3-factor latent structure were upheld in several replications (Social Phobia, Agoraphobia, Interoceptive), three items failed to load on their predicted factor (Interoceptive). Multiple-groups CFAs indicated that the measurement properties of the APPQ were invariant in male and female patients, with the exception of an intercept of one item from the Agoraphobia scale which evidenced bias against females. The three APPQ dimensions were consistently associated with high levels of scale reliability and factor determinacy. Strong evidence of concurrent validity of the Social Phobia and Agoraphobia factors was obtained in relation to interview and questionnaire measures. Although the Interoceptive factor was more strongly related to criterion measures of anxiety sensitivity and fear of panic than Social Phobia, the Agoraphobia factor had the strongest relationships with these validity indices. The results are discussed in regard to psychometric implications for the APPQ and conceptual issues pertaining to the discriminant validity of fear of agoraphobic situations and fear of sensation-producing activities.

Keywords: Panic disorder; Agoraphobia; Social anxiety disorder; (Social Phobia); Psychometrics; Psychological assessment; Confirmatory factor analysis

*Corresponding author.
E-mail address: tabrown@bu.edu (T.A. Brown).

0005-7967/$ - see front matter
doi:10.1016/j.brat.2004.03.004

1. Introduction

Fear and avoidance are hallmark features of the phobic disorders (panic disorder with agoraphobia, social phobia, specific phobia). The breadth and depth of the fear and avoidance can range from *overt* situational avoidance (e.g., public speaking, heights) to more *subtle* fears of substances or situations that may provoke autonomic symptoms similar to anxiety and panic (e.g., caffeine, aerobic activity). This distinction between the obvious and subtle forms of fear and avoidance is particularly evident in panic disorder. Phobic avoidance in panic disorder (i.e., agoraphobia) can be established from any one of the following criteria: (1) situational avoidance, (2) anxiety and distress in the phobic situation, or (3) requiring the presence of a companion in the phobic situation (American Psychiatric Association, 1994). As such, an individual may have no or few overtly avoided situations but may still have considerable fear, distress, and subtle avoidance behavior in the phobic situation. In short, the fear associated with agoraphobia is focused on the onset of distressing symptoms in situations, not the avoided situation itself. Because of this "fear of fear" (Goldstein & Chambless, 1978), a continuum of situational fears and restrictions may develop.

Whereas early conceptualizations of panic disorder with agoraphobia (PDA) centered on overt situational avoidance (e.g., Marks, 1969; Snaith, 1968), contemporary theories of PDA have emphasized the fears of autonomic sensations that are similar to those of a panic attack (Barlow, 2002; Clark, 1986). Indeed, many modern empirically supported treatments target exposure to both the phobic situations as well as the autonomic (or interoceptive) symptoms and situations (Barlow, Gorman, Shear, & Woods, 2000; Spiegel & Barlow, 2000). Because of the important role of interoceptive fear and avoidance in the development and treatment of PDA, thorough assessment of the feared sensations and activities that may produce them are essential. From a clinical standpoint, this assessment can serve as the foundation for designing interoceptive exposure exercises—a vital treatment component for decreasing a patient's anxiety about and sensitivity to physical sensations (Craske & Barlow, 2001; White & Barlow, 2002). In addition to its clinical importance, comprehensive assessment of dimensions of phobic fear and avoidance may have significant etiological and theoretical relevance for research in understanding PDA and the phobias.

Although a number of scales have been developed to assess integral aspects of PDA (cf. Antony, 2001), the assessment of interoceptive fear and avoidance has been largely neglected. An exception is the Albany Panic and Phobia Questionnaire (APPQ; Rapee, Craske, & Barlow, 1994/1995), a scale that was designed to measure the distinct dimension of fear of sensation-producing activities, in addition to fear of common agoraphobic and social phobic situations. Using principal components analysis (PCA) with equamax rotation, Rapee et al. (1994/1995) reported findings in support of the posited 3-factor solution in a sample of 405 patients with anxiety disorders and 33 persons with no mental disorder. After examination of internal consistency and factor structure (and removal of items failing to meet salient loading criteria), the final scale consisted of 27 items forming three subscales, interpreted as reflecting fear of agoraphobic situations ("Agoraphobia," 9 items), fear of activities that produce somatic sensations ("Interoceptive," 8 items), and fear of social situations ("Social Phobia," 10 items). Rapee et al. (1994/1995) subjected the APPQ to additional tests of internal consistency, temporal stability, validity, and sensitivity to change during treatment, and all evidences attested to the favorable psychometric qualities of the scale (e.g., αs ranged from .87 to .90). For instance, support for the convergent/discriminant validity of the APPQ was evidenced by a differential pattern of correlations with other clinical ratings and self-report measures of social

phobia and panic disorder and agoraphobia in the predicted direction (e.g., APPQ-Social Phobia was more strongly associated with presence/absence of social phobia and a social anxiety questionnaire than was APPQ-Agoraphobia and APPQ-Interoceptive). Moreover, a significant decrease in all three subscales was observed following cognitive-behavioral treatment of 45 patients with panic disorder with or without agoraphobia, lending support for the utility of the APPQ as a treatment outcome measure (Rapee et al., 1994/1995). More recently, Novy, Stanley, Averill, and Daza (2001) extended these results by reporting equivalent reliability and norms, as well as favorable convergent and discriminant validity, for an English and Spanish language translation of the APPQ in a bilingual sample. However, the small sample size ($N = 98$) restricted this investigation to predominantly descriptive and univariate analyses of the scale (e.g., no replication of factor structure).

Since its development, the APPQ has been used in a wide variety of contexts (cf. Antony, 2001), including large-scale treatment outcome studies (e.g., Barlow et al., 2000), latent structural analyses of the *DSM-IV* anxiety and mood disorder constructs (e.g., Brown, Chorpita, & Barlow, 1998), and analog (e.g., Veljaca & Rapee, 1998) and cross-cultural research (Novy et al., 2001). Nevertheless, this applied research has proceeded in absence of comprehensive examination of the measurement properties of the APPQ. For instance, Rapee et al. (1994/1995) represent the only published study to date that has examined the latent structure of the APPQ. Unfortunately, these analyses were limited in several ways, including: (a) use of PCA as the method of extraction, a procedure that is not based on the common factor model (cf. Fabrigar, Wegener, MacCallum, & Strahan, 1999; Thurstone, 1947; Widaman, 1993); (b) retention in the final solution of several items with salient loadings ($> .30$) on nonprimary factors; and (c) inability to cross-validate the final solution in independent samples. Moreover, Rapee et al. (1994/1995) aimed to evaluate the "factorial invariance" of the final solution between males and females. However, this objective was addressed simply by visual inspection of the structural similarities of PCAs conducted separately for males and females.

Thus, the primary aim of the current study was to provide a thorough evaluation of the latent structure of the APPQ in a large sample of patients with anxiety and mood disorders. Given the aforementioned methodological limitations of Rapee et al. (1994/1995), we began by conducting cross-validated exploratory factor analyses to better ascertain the latent form of the APPQ and to identify any poorly behaved items. Subsequent analyses on independent samples were conducted in the confirmatory factor analysis (CFA) framework which provided further replication of factor structure, and allowed for the evaluation of potential nonrandom measurement error, scale reliability, and direct statistical examination of the form and measurement invariance of the APPQ between male and female patients. Moreover, the convergent and discriminant validity of the APPQ dimensions were examined in context of the CFA solutions.

2. Method

2.1. Participants

The sample consisted of 1930 outpatients who were presented for assessment and treatment at the Center for Anxiety and Related Disorders. Women constituted the larger portion of the

sample (61%); average age was 33.26 ($SD = 11.16$, range = 18–79). The sample was predominantly Caucasian (89%; African-American = 3.5%, Asian = 4%, Latino/Hispanic = 3%). Diagnoses were established with the Anxiety Disorders Interview Schedule for *DSM-IV*: Lifetime version (ADIS-IV-L; Di Nardo, Brown, & Barlow, 1994), a semi-structured interview designed to ascertain reliable diagnosis of the *DSM-IV* anxiety, mood, somatoform, and substance use disorders, and to screen for the presence of other conditions (e.g., psychotic disorders). Moreover, the ADIS-IV-L provides dimensional assessment of the key and associated features of disorders (0–8 ratings); for many disorders (e.g., social phobia, specific phobia, generalized anxiety disorder, obsessive–compulsive disorder), such features are dimensionally rated regardless of whether a formal *DSM-IV* diagnosis is under consideration. A reliability study of a subset of the current sample ($N = 362$) who had two independent administrations of the ADIS-IV-L indicated good-to-excellent interrater agreement for current disorders (range of $\kappa s = .67-.86$) except dysthymia ($\kappa = .31$; Brown, Di Nardo, Lehman, & Campbell, 2001). For each diagnosis, interviewers assign a 0–8 clinical severity rating (CSR) that indicates the degree of distress and impairment associated with the disorder (0 = "none" to 8 = "very severely disturbing/disabling"). In patients with two or more current diagnoses, the "principal" diagnosis is the one receiving the highest CSR. For current and lifetime disorders that meet or surpass the threshold for a formal *DSM-IV* diagnosis, CSRs of 4 (definitely disturbing/disabling) or higher are assigned ("clinical" diagnoses). Current clinical diagnoses not deemed to be the principal diagnosis are referred to as "additional" diagnoses. The rates of current clinical disorders (collapsing across principal and additional diagnoses) that frequently occurred in the sample were as follows: panic disorder with or without agoraphobia ($n = 708$), generalized anxiety disorder ($n = 403$), social phobia ($n = 811$), specific phobia ($n = 396$), obsessive–compulsive disorder ($n = 222$), major depression ($n = 501$), and dysthymic disorder ($n = 156$).[1]

2.2. Measures

As noted earlier, the APPQ is a 27-item instrument designed to measure interoceptive, agoraphobic, and social situational fear (Rapee et al., 1994/1995; see Appendix). Participants respond to the items using a 0–8 scale, where 0 = "no fear" and 8 = "extreme fear." The three APPQ subscales are scored by summation, as guided by the factor solution reported in Rapee et al. (1994/1995); Social Phobia, 10 items; Agoraphobia, 9 items; Interoceptive, 8 items).

In addition to the APPQ, the following questionnaires were administered: Anxiety Sensitivity Index (ASI; Peterson & Reiss, 1992), and Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998). These measures were used in validity analyses of the latent APPQ factors. The sample estimate of the SIAS's internal consistency was .95 (Cronbach's α). Given consistent evidence of the multifactorial nature of the ASI (e.g., Zinbarg, Barlow, & Brown, 1997; Zinbarg, Mohlman, & Hong, 1999), and evidence that these ASI factors have differential relevance to the

[1]As noted in the Results, diagnostic status was used as a concurrent validity indicator by collapsing current principal and additional diagnoses (i.e., current *DSM-IV* panic disorder with/without agoraphobia, current social phobia). This approach was deemed more suitable than the common strategy of coding by principal diagnosis (i.e., use of principal diagnoses would weaken the analyses because many patients who did not have a given principal diagnosis were assigned the disorder as an additional diagnosis).

prediction of anxiety disorder outcomes (e.g., Carter, Suchday, & Gore, 2001; Zinbarg, Brown, Barlow, & Rapee, 2001), subscales rather than the total score were used in the validity analyses. Specifically, the ASI Physical Concerns ($\alpha = .87$) and ASI Social Concerns ($\alpha = .57$) subscales were analyzed under the premise that the former would be more strongly related to APPQ-Agoraphobia and APPQ-Interoceptive, whereas the latter would evidence a stronger relationship with APPQ-Social Phobia.[2]

Dimensional clinical ratings from the ADIS-IV-L were also used in the validity analyses (0–8 scales, higher scores indicate higher levels of fear/avoidance). For purposes of the present analyses, the following dimensional ratings were used: (a) fear of 13 social situations (e.g., initiating a conversation, participating at meetings/classes); (b) agoraphobic avoidance of 22 situations (e.g., malls, being at home alone); and (c) single rating of past-month fear of panic. Factor analyses reported in Brown et al. (2001) indicated that the fear ratings of 13 social situations and avoidance ratings of 22 agoraphobic situations were unidimensional. A large sample ($N = 362$) analysis of the interrater reliability for these ratings yielded test–retest correlations of .86, .86, and .53 for social situational fear, agoraphobic avoidance, and fear of panic, respectively (Brown et al., 2001).

3. Procedure

Patients completed the APPQ as part of their initial intake evaluation (consisting of the ADIS-IV-L and a questionnaire battery) at the Center for Anxiety and Related Disorders. To thoroughly examine and cross-validate the APPQ factor structure, the total sample ($N = 1930$) was randomly divided into four subsamples. The first two samples (Sample 1: $N = 300$, 110 males, 190 females; Sample 2: $N = 300$, 129 males, 171 females) were used to conduct initial exploratory factor analyses of the original 27-item APPQ and revised solutions. Because of the methodological shortcomings noted in earlier research, these initial analyses of the APPQ's latent structure were conducted in an exploratory fashion (i.e., a sufficient empirical basis was lacking to initiate analyses with CFA). Sample 3 ($N = 400$; 154 males, 246 females) was used to conduct an exploratory factor analysis (EFA) in the CFA framework as a precursor to CFA (see Results, for further justification). Sample 4 ($N = 930$; 367 males, 563 females) was used for CFA replications and generalizability analyses of the final APPQ latent structure, and to conduct concurrent validity analyses in the latent variable framework.

3.1. Data analysis

The sample variance–covariance matrices were analyzed using latent variable software programs and maximum-likelihood minimization functions (LISREL 8.52, Jöreskog & Sörbom,

[2]The first-order latent structure of the ASI is also characterized by a dimension that has been labeled, "Mental Incapacitation Concerns" (e.g., Zinbarg et al., 1997). However, this ASI factor was not included in the analyses because of the lack of hypotheses regarding its differential relationship to agoraphobic, interoceptive, and social fear (e.g., in Zinbarg et al., 2001, Physical Concerns was the only ASI factor that predicted fearful responding to CO_2 challenge in patients with anxiety disorders).

2002; Mplus 2.12, Muthén & Muthén, 2002). Goodness of fit of the CFA models was evaluated using the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990), and its 90% confidence interval (cf. MacCallum, Browne, & Sugawara, 1996), the Tucker–Lewis Index (TLI, Tucker & Lewis, 1973), and the Comparative Fit Index (CFI; Bentler, 1990). Acceptable model fit was defined by the following criteria: RMSEA ($<.08$, 90% CI $<.08$), CFI ($>.90$), and TLI ($>.90$). Multiple indices were selected because they provide different information for evaluating model fit (i.e., absolute fit, fit adjusting for model parsimony, fit relative to a null model); used together, these indices provide a more conservative and reliable evaluation of the model fit (cf. Jaccard & Wan, 1996). In instances where competing models were nested, comparative fit was evaluated with nested χ^2-tests.

4. Results

4.1. Exploratory factor analyses

Using Sample 1 ($N = 300$), the 27 APPQ items were submitted to an EFA (maximum likelihood estimation, promax rotation). Acceptability of the factor models (e.g., factor selection) was evaluated by goodness of model fit (RMSEA $<.08$, upper bound of 90% CI $<.08$), the interpretability of the solution, and the strength of the parameter estimates (e.g., primary factor loadings $>.30$, absence of salient cross-loadings). Consistent with the findings of Rapee et al. (1994/1995), a 3-factor solution fit the data well, $\chi^2(273) = 693.95$, $p<.001$, RMSEA $=.072$ (.065/.078) (variance explained $=54\%$; eigenvalues >1.0 for the unreduced correlation matrix were 8.10, 4.55, 1.83, 1.27, 1.09). A 4-factor solution converged but was associated with a factor on which only one item had a salient primary loading with no cross-loadings. As per Rapee et al. (1994/1995), these factors were labeled Social Phobia, Agoraphobia, and Interoceptive (variance accounted for $=30\%$, 17%, and 7%, respectively). As shown in Table 1, all 27 items had primary factor loadings above .30. However, three items that were reported by Rapee et al. (1994/1995) to load on the Interoceptive factor had primary loadings on the Agoraphobia factor in this solution (item 7: "getting gas at the dentist"; item 17: "drinking a strong cup of coffee"; item 26: "feeling the effects of alcohol"). Because Rapee et al. (1994/1995) had retained these three items as indicators of the interoceptive fear dimension, they were included in the Sample 2 EFA in order to ascertain if this discordant result was upheld in the second exploratory analysis.[3]

Using the Sample 2 ($N = 300$) data, the 3-factor solution was associated with sufficient goodness of fit, $\chi^2(273) = 707.26$, $p<.001$, RMSEA $=.073$ (.066/.079) (eigenvalues >1.0 for the unreduced correlation matrix were 7.69, 4.40, 1.92, 1.13, 1.12; variance explained $=52\%$). A 4-factor solution converged but the fourth factor was defined by no items that evidenced acceptable patterns of primary and cross-loadings. As in the Sample 1 solution, all APPQ items had strong primary loadings (and no salient cross-loadings) on their predicted factors, except for items 7, 17, and 26 (see Table 1). Given the consistency of this outcome, these three items were eliminated from further analyses.

[3]Although item 7 was retained by Rapee et al. (1994/1995) as an indicator of the Interoceptive factor, it had a salient cross-loading on Agoraphobia (.33) in their principal components analysis.

Table 1
Latent structure of the original (27 item) Albany Panic and Phobia Questionnaire: exploratory factor analyses in Samples 1 and 2 (*Ns* = 300)

	Factor					
	Social Phobia		Agoraphobia		Interoceptive	
	S1	S2	S1	S2	S1	S2
APPQ item						
1	**.804**	**.795**	.096	.102	.055	.053
5	**.340**	**.454**	.206	.013	.036	.158
8	**.747**	**.735**	.056	.020	.037	.014
9	**.633**	**.643**	.016	.008	.003	.018
12	**.848**	**.825**	.050	.025	.013	.074
15	**.504**	**.521**	.034	.059	.031	.003
21	**.548**	**.495**	.021	.031	.025	.015
22	**.880**	**.870**	.011	.003	.050	.055
23	**.535**	**.469**	.002	.131	.013	.038
24	**.874**	**.864**	.038	.011	.005	.069
2	.068	.053	**.677**	**.512**	.103	.035
7[a]	.033	.028	**.587**	**.487**	.055	.147
11	.100	.029	**.785**	**.820**	.045	.015
13	.006	.025	**.632**	**.578**	.030	.112
14	.017	.041	**.728**	**.718**	.116	.109
16	.075	.147	**.699**	**.746**	.053	.150
17[a]	.052	.021	**.358**	.273	.161	.279
18	.152	.062	**.595**	**.434**	.040	.041
20	.042	.033	**.741**	**.637**	.002	.078
25	.049	.048	**.631**	**.686**	.052	.003
26[a]	.016	.024	**.410**	**.310**	.229	.220
27	.109	.023	**.693**	**.670**	.028	.069
3	.011	.066	.093	.066	**.954**	**.874**
4	.041	.037	.266	.007	**.509**	**.796**
6	.016	.037	.051	.021	**.898**	**.889**
10	.070	.049	.284	.221	**.534**	**.606**
19	.023	.033	.090	.095	**.642**	**.653**

Note: APPQ = Albany Panic and Phobia Questionnaire; S1 = Sample 1; S2 = Sample 2. Exploratory factor analysis conducted by maximum likelihood estimation, promax rotation; Factor loadings ⩾ .30 are in bold.
[a]Items loaded on a factor counter to the original APPQ subscale scoring.

Table 2 presents the promax-rotated pattern matrix of the 24-item, 3-factor solution from the combined exploratory samples ($N = 600$); $\chi^2(207) = 802.00$, $p < .001$, RMSEA = .069 (.064/.074), variance explained = 56%. The factors were weakly to moderately correlated: Social Phobia–Agoraphobia = .24, Social Phobia–Interoceptive = .23, and Agoraphobia–Interoceptive = .54. As noted elsewhere (e.g., Grice, 2001), factor determinacy is important in the evaluation of factor analytic results (e.g., a highly indeterminate factor can produce radically different factor scores

that are nonetheless equally consistent with the obtained factor loadings). Accordingly, factor determinacies (i.e., validity coefficients: correlation between factor score estimates and their respective factors) were computed on the refined factor scores (least squares regression approach; Thurstone, 1947) using the SAS PROC IML routines provided by Grice (2001). As shown in Table 2, all three factors evidenced a high level of determinacy (range = .94–.97), per the guidelines forwarded by Gorsuch (1983; i.e., > .80; cf. Grice, 2001).

Table 2
Latent structure of the revised (24 item) Albany Panic and Phobia Questionnaire: exploratory factor analysis ($N = 600$, Samples 1 and 2) and exploratory factor analysis conducted within the confirmatory factor analysis framework ($N = 400$, Sample 3)

	Factor					
	Social Phobia		Agoraphobia		Interoceptive	
	EFA	E/CFA	EFA	E/CFA	EFA	E/CFA
APPQ item						
1	**.800**	**.828**	.096	.100	.045	.046
5	**.402**	**.456**	.087	.087	.106	.156
8	**.743**	**.747**	.019	.029	.014	.064
9	**.638**	**.633**	.025	.015	.002	.039
12	**.835**	**.824**	.011	.002	.044	.031
15	**.512**	**.589**	.002	.090	.017	.019
21	**.523**	**.616**	.012	.008	.013	.024
22*	**.874**	**.843**	.002	.000	.058	.000
23	**.502**	**.532**	.060	.050	.025	.050
24	**.872**	**.843**	.011	.093	.033	.080
2	.058	.029	**.579**	**.467**	.043	.029
11*	.067	.000	**.797**	**.835**	.036	.000
13	.007	.031	**.599**	**.608**	.085	.079
14	.028	.063	**.726**	**.640**	.096	.061
16	.109	.144	**.712**	**.659**	.078	.016
18	.114	.124	**.503**	**.384**	.010	.208
20	.008	.008	**.681**	**.575**	.056	.178
25	.002	.102	**.649**	**.677**	.045	.077
27	.065	.026	**.662**	**.688**	.025	.157
3*	.029	.000	.069	.000	**.910**	**.873**
4	.048	.076	.124	.018	**.648**	**.679**
6	.033	.090	.044	.131	**.894**	**.809**
10	.063	.131	.261	.068	**.576**	**.730**
19	.029	.099	.088	.080	**.648**	**.640**
Determinacy	.965	.966	.942	.942	.960	.952

Note: APPQ = Albany Panic and Phobia Questionnaire; EFA = exploratory factor analysis (maximum likelihood extraction, promax rotation); E/CFA = exploratory factor analysis within the confirmatory factor analysis framework (maximum likelihood). Items with asterisks were used as anchor indicators in the E/CFA analysis. Factor loadings ⩾ .30 are in bold.

4.2. Exploratory factor analysis within the confirmatory factor analysis framework

Samples 3 and 4 were used to cross-validate the latent structure of the revised (24-item) APPQ. As an intermediate step between EFA and CFA, the APPQ data from Sample 3 ($N = 400$) were analyzed using the EFA within the CFA framework approach (E/CFA; Jöreskog, 1969). Although infrequently used in applied factor analysis research, E/CFA is a very helpful precursor to CFA that allows measurement models to be explored more fully before moving into a confirmatory framework (cf. Brown, 2003). A typical sequence in psychological scale development is to conduct CFA immediately after the number of latent dimensions has been ascertained by EFA. However, poor-fitting CFA solutions are frequently encountered because of the multiple potential sources of misfit that are undetected by EFA (e.g., residual covariances are usually fixed to zero in initial CFA models). The researcher may then be faced with extensive post hoc model testing subject to the criticisms of specification searches in a single data set (MacCallum, 1986). The E/CFA approach represents an analytic step between EFA and CFA that provides information important in the development of realistic confirmatory solutions. In this strategy, the CFA applies the same number of identifying restrictions as EFA (m^2) by fixing factor variances to unity and by selecting one anchor item for each factor whose cross-loadings are fixed to zero (the loadings of nonanchor items are freely estimated on each factor).[4] Whereas this specification produces the same model fit as maximum likelihood EFA, the CFA estimation provides considerably more information (e.g., standard errors for determining the statistical significance of factor loadings and cross-loadings). An important aspect of the E/CFA results includes modification indices (MIs) reflecting potentially substantive indicator error covariances (e.g., method effects stemming from reverse or similarly worded items; cf. Brown, 2003; Byrne, Shavelson, & Muthén, 1989; Marsh, 1996). Thus, the researcher can develop a realistic measurement model prior to moving into the more restrictive CFA framework.

Using items 3, 11, and 22 as anchor items, the 3-factor E/CFA model provided a reasonable fit to the data, $\chi^2(207) = 622.17$, $p < .001$, RMSEA = .071 (90% CI: .064/.077), CFI = .92, TLI = .89. As shown in Table 1, the magnitudes of primary loadings were strong (range = .38–.87) and statistically significant (range of $zs = 6.53-21.23$). All factor determinacies (calculated by Mplus) were above .90 (range = .94–.97). The factor intercorrelations were: Social Phobia–Agoraphobia = .10 (*ns*), Social Phobia–Interoceptive = .18 ($p < .01$), and Agoraphobia–Interoceptive = .54 ($p < .001$). Importantly, this analysis was associated with several high MIs that pointed to the existence of appreciable error covariances among some items within each of the three major factors (cf. Gerbing & Anderson, 1984). For instance, although item 3 (playing a vigorous sport on a hot day) and item 6 (hiking on a hot day) had very strong loadings on the Interoceptive factor (.87 and .81, respectively), results indicated that a significant amount of their observed covariation was unexplained by this solution; MI = 31.45, completely standardized expected parameter change (EPC) for the correlated residual = .16. This nonrandom error appeared to be grounded substantively by the differentially high content overlap of these items (physical exertion

[4] "m" refers to the number of latent factors in the solution. Thus, in the current E/CFA involving three latent factors, there were nine identifying restrictions (3^2): factor variances were set to unity (three restrictions), and the cross-loadings of anchor items were fixed to zero (six restrictions, see Table 2).

in hot weather), relative to other items that loaded on the Interoceptive factor (e.g., "blowing up an airbed quickly"). Similarly, substantively salient error covariances appeared to exist within the Social Phobia dimension for the two public speaking items (items 9 and 12, MI = 45.37, EPC = .16) and among the three items assessing one-on-one social interaction (items 1, 21, and 22). On the Agoraphobia factor, large and substantively interpretable MIs were observed for the two items that pertain to being away from home alone (items 11 and 25, MI = 32.86, EPC = .15), the two items involving highway driving (items 14 and 27, MI = 17.88, EPC = .15), and the two items that involve confinement for protracted periods (items 18 and 20, MI = 21.43, EPC = .15). Thus, these findings suggested that while the common 3-factor model provided an adequate fit to the data, a more complex error theory (correlated uniquenesses) may be needed to represent minor factors or error covariances existing within each major dimension.

4.3. Confirmatory factor analysis

In accord with the conclusions from the E/CFA findings, a CFA fitted to the Sample 4 APPQ data ($N = 930$), where measurement error was specified as random, did not fit the data well, $\chi^2(249) = 1627.58$, $p < .001$, RMSEA = .077 (90% CI = .074/.081), CFI = .88, TLI = .87. Fit diagnostics indicated that strains in the solution were not due to the failure to estimate salient cross-loadings (highest EPC = .16). However, consistent with the results from the E/CFA, high MIs corresponding to the eight error covariances among the items noted above were obtained (range of MIs = 25.67–177.39, range of EPCs = .08–.22). Given the consistency and interpretability of this outcome, the solution was respecified allowing for these correlated residuals. In addition to improving fit over the previous solution, $\chi^2_{\text{diff}}(8) = 595.63$, $p < .001$, this model fit the data well, $\chi^2(241) = 1031.95$, $p < .001$, RMSEA = .059 (90% CI = .056/.063), CFI = .93, TLI = .92.[5] As shown in Table 3, the magnitudes of the factor loadings were strong (range = .52–.89) and the factor determinacies were quite satisfactory (range = .93–.96). Factor intercorrelations were: Social Phobia–Agoraphobia = .25, Social Phobia–Interoceptive = .16, and Agoraphobia–Interoceptive = .59 (all *ps* < .001). In addition, all eight correlated residuals, which ranged .12–.19, were significant at the .001 level.

4.3.1. Scale reliability

The scale reliabilities of the three factors were calculated within the CFA model using the approach developed by Raykov (2001). This method reconciles the problems with Cronbach's α which is a misestimator of scale reliability except in the rare instance when all elements of a multiple-item measure are tau-equivalent and free of nonrandom measurement error (Lord & Novick, 1968; McDonald, 1999; Raykov, 2001, 2004). In LISREL, the procedure entails specifying three dummy latent variables whose variances are constrained to equal the numerator (true score variance), denominator (total variance), and corresponding ratio of true score variance

[5]Data diagnostics indicated no substantial departures from normality for the APPQ items except item 19 (Running up stairs) which was positively skewed and leptokurtotic (patients rarely reported fear of this activity). Analyses were also conducted using robust estimators (e.g., robust maximum likelihood, MLM) and no substantial differences in overall model fit or the significance tests of parameter estimates were noted; e.g., MLM fit of the final CFA model with $N = 930$, $\chi^2(241) = 739.61$, $p < .001$, RMSEA = .047, CFI = .94, TFI = .93.

Table 3
Latent structure of the revised (24 item) Albany Panic and Phobia Questionnaire: confirmatory factor analysis using Sample 4 ($N = 930$) and male ($N = 367$) and female ($N = 563$) patients

	Factor								
	Social Phobia			Agoraphobia			Interoceptive		
	All	Males	Females	All	Males	Females	All	Males	Females
APPQ item									
1	.886	.785	.764						
5	.518	.492	.538						
8	.799	.777	.812						
9	.632	.647	.625						
12	.826	.808	.836						
15	.568	.583	.556						
21	.538	.559	.521						
22	.813	.781	.832						
23	.549	.568	.533						
24	.883	.872	.891						
2				.539	.480	.544			
11				.781	.779	.769			
13				.654	.653	.616			
14				.572	.529	.554			
16				.714	.659	.717			
18				.535	.503	.557			
20				.682	.694	.672			
25				.615	.559	.641			
27				.581	.565	.557			
3							.739	.677	.773
4							.690	.643	.705
6							.836	.808	.842
10							.768	.757	.776
19							.793	.763	.809
Determinacy	.959	.956	.965	.928	.920	.926	.936	.926	.931
Reliability	.891	.888	.893	.847	.821	.844	.857	.823	.870
Mean	21.49	22.83	2.62	16.86	12.23	19.87	5.00	4.32	5.44
SD	16.13	15.99	16.18	14.44	11.89	15.15	7.44	6.48	7.98

Note: APPQ = Albany Panic and Phobia Questionnaire. Means and *SD*s are based on coarse factor scores (i.e., raw score composites).

to total score variance, per the classic formula for scale reliability estimation (Lord & Novick, 1968). The resulting estimates were .89, .85, and .86 for Social Phobia, Agoraphobia, and Interoceptive, respectively, attesting to the favorable scale reliability of the three APPQ dimensions.

4.3.2. Measurement invariance between sexes

The degree of structural and measurement invariance (equal factor loadings and indicator intercepts) of the 24-item APPQ was examined in male and female patients in Sample 4 using multiple-groups CFA (*ns* = 367 and 563, respectively).[6] The two-group CFA, testing equal APPQ form between the sexes (using the final solution obtained for the full sample), fit the data well, $\chi^2(482) = 1363.03$, $p < .001$, RMSEA = .063 (90% CI = .059/.067), CFI = .97, TLI = .96. The factor loadings, factor determinacies, and scale reliabilities for males and females are provided in Table 3. Given evidence of equal form, the analysis proceeded to examine the metric equivalence of the APPQ. This analysis indicated that the males' and females' factor loadings were invariant, $\chi^2_{\text{diff}}(21) = 24.78$, *ns*; $\chi^2(503) = 1387.81$, $p < .001$, RMSEA = .062 (90% CI = .058/.066), CFI = .97, TLI = .96. However, the next model, which constrained the indicator intercepts to equality (scalar invariance), produced a significant degradation in fit, $\chi^2_{\text{diff}}(24) = 254.37$, $p < .001$; $\chi^2(527) = 1642.18$, $p < .001$, RMSEA = .068 (90% CI = .064/.071), CFI = .96, TLI = .95. Fit diagnostics revealed a stand-out MI (106.06) associated with the intercept of item 13 (walking alone in isolated areas), from the Agoraphobia factor (unstandardized EPCs = −.67 and .66 for males and females, respectively). Although noninvariance of the item location parameters may often relate to substantive group differences in the underlying construct of interest (e.g., sex differences in agoraphobic avoidance; cf. Vandenberg & Lance, 2000), this was the only intercept among the Agoraphobia items that produced significant strain in the solution (range of other MIs = .45–4.25). Given that males and females would likely respond differently to this item regardless of agoraphobia status, this aspect of the scalar invariance analysis suggests that item 13 functions differently between sexes as an indicator of agoraphobic avoidance (i.e., even when the value of the underlying dimension of Agoraphobia is zero, females' observed scores on this item are higher than males' due to sex differences in the range of activities relating to personal safety). No other stand-out MIs were noted for items loading onto the Social Phobia and Interoceptive factors.

4.4. Concurrent validity of the APPQ factors

The convergent and discriminant validity of the three APPQ dimensions were examined by regressing these latent factors onto single indicators of panic disorder/agoraphobia (current *DSM-IV* diagnosis of panic disorder with or without agoraphobia, ADIS-IV-L agoraphobia ratings, ADIS-IV-L past month fear of panic, ASI Physical Concerns subscale) and social phobia (current *DSM-IV* diagnosis of social phobia, ADIS-IV-L fear ratings of social situations, SIAS, ASI Social Concerns subscale). Measurement error of the six-dimensional single indicators was adjusted for in the analyses by imposing constraints to their error variances using known reliability information (see Method) and their sample variances (cf. Bollen, 1989).

It was predicted that the four social phobia indicators (current *DSM-IV* social phobia, ADIS-IV-L fear of social situations, SIAS, ASI Social Concerns) would be more strongly related to the APPQ-Social Phobia latent factor than to the Interoceptive and Agoraphobia latent factors. In

[6]Although the multiple-groups strategy can be used to evaluate other potential aspects of group differences (e.g., equality of factor intercorrelations and factor means), these analyses were limited to examination of measurement invariance (i.e., equal form, factor loadings, and item intercepts) in light of the psychometric objectives of the study.

addition, it was expected that while the ADIS-IV-L fear of panic and ASI Physical Concerns indicators would have significantly stronger relationships with APPQ-Agoraphobia than APPQ-Social Phobia, their largest associations would be with APPQ-Interoceptive. Similarly, although the ADIS-IV-L Agoraphobia indicator would covary more strongly with APPQ-Interoceptive than APPQ-Social Phobia, its relationship with APPQ-Agoraphobia would be of the highest magnitude. Finally, it was predicted that the indicator representing the presence or absence of current *DSM-IV* panic disorder would be more strongly related to APPQ-Agoraphobia and APPQ-Interoceptive than APPQ-Social Phobia.

The eight structural models provided a good fit to the data.[7] The structural parameter estimates of these models are presented in Table 4. Using the initial solutions as baseline models, subsequent models were evaluated placing equality constraints on the structural parameters to determine if the differential magnitude of these relationships with the APPQ factors was in accord with prediction (e.g., a significant decrease in model fit when holding the paths between *DSM-IV* social phobia and APPQ-Social Phobia and APPQ-Agoraphobia to equality would provide support for the concurrent validity of the APPQ-Social Phobia scale). The results of these analyses are also provided in Table 4.

In each analysis involving the panic disorder/agoraphobia predictors, APPQ-Social Phobia had a significantly smaller (and often nonsignificant) association, relative to APPQ-Agoraphobia and APPQ-Interoceptive. However, whereas all four predictors accounted for significant variance in APPQ-Agoraphobia and APPQ-Interoceptive, in each instance these relationships were significantly stronger for APPQ-Agoraphobia. This outcome included predictors that were expected to be more strongly related to APPQ-Interoceptive (e.g., ASI Physical Concerns).

All four social phobia predictors were strongly related to APPQ-Social Phobia. For instance, the composite of the ADIS-IV-L ratings of fear of 13 different social situations was correlated with APPQ-Social Phobia at .88. Moreover, each predictor was more strongly related to APPQ-Social Phobia than APPQ-Agoraphobia and APPQ-Interoceptive (the latter two factors did not differ in their magnitude of their relationships with the predictors in any analysis). Of particular note is the finding that the ASI Social Concerns subscale was more strongly related to APPQ-Social Phobia scale ($\gamma = .68$) than APPQ-Agoraphobia ($\gamma = .23$) and APPQ-Interoceptive ($\gamma = .22$), consistent with prior evidence of the differential relevance of the ASI subscales to the prediction of various anxiety disorder constructs (e.g., Zinbarg et al., 2001).

5. Discussion

The present findings extend the psychometric basis of the APPQ (Novy et al., 2001; Rapee et al., 1994/1995) by providing cross-validated EFAs and CFAs and tests of measurement variance and concurrent validity in large samples of outpatients with *DSM-IV* anxiety and mood disorders. Although a tripartite latent structure was upheld (Agoraphobia, Interoceptive, Social), these initial analyses revealed that three of the APPQ's original 27 items (items 7, 17, 26) failed to load

[7]The model with the highest χ^2-value involved the ADIS-IV-L agoraphobic avoidance predictor, $\chi^2(262) = 1192.67$, $p < .001$, RMSEA = .062 (90% CI = .058/.065), CFI = .92, TLI = .91. Due to space considerations, overall fit statistics and fit diagnostic information for each of these models are not provided, but are available on request from the authors.

Table 4
Differential relationships of the Albany Panic and Phobia Questionnaire Factors with panic disorder/agoraphobia and social phobia variables ($N = 930$)

	APPQ factor		
	Social Phobia	Agoraphobia	Interoceptive
Panic Disorder/Agoraphobia			
Current *DSM-IV* PD/A	$-.38^{**}$	2.35^{***}	1.18^{***}
	$(-.21)^{c}$	$(1.08)^{a}$	$(.76)^{b}$
ADIS-IV-L Agoraphobia	.00	$.08^{***}$	$.04^{***}$
	$(.03)^{c}$	$(.78)^{a}$	$(.47)^{b}$
ADIS-IV-L Fear of Panic	−.05	$.84^{***}$	$.40^{***}$
	$(-.06)^{c}$	$(.74)^{a}$	$(.50)^{b}$
ASI Physical Concerns	$.03^{**}$	$.15^{***}$	$.10^{***}$
	$(.12)^{c}$	$(.50)^{a}$	$(.47)^{b}$
Social Phobia			
Current *DSM-IV* SOC	2.33^{***}	$-.39^{**}$	$-.23^{*}$
	$(1.29)^{a}$	$(-.19)^{b}$	$(-.15)^{b}$
ADIS-IV-L Social Fear	$.09^{***}$	.01	.00
	$(.88)^{a}$	$(.07)^{b}$	$(.03)^{b}$
SIAS	$.08^{***}$	$.01^{**}$	$.01^{*}$
	$(.86)^{a}$	$(.11)^{b}$	$(.09)^{b}$
ASI Social Concerns	$.51^{***}$	$.21^{***}$	$.14^{***}$
	$(.68)^{a}$	$(.23)^{b}$	$(.22)^{b}$

Note: APPQ = Albany Panic and Phobia Questionnaire; PD/A = *DSM-IV* panic disorder with or without agoraphobia; ADIS-IV-L = Anxiety Disorders Interview Schedule for *DSM-IV*: Lifetime version; ASI = Anxiety Sensitivity Index; SOC = *DSM-IV* social phobia; SIAS = Social Interaction Anxiety Scale. Presented are unstandardized path estimates and completely standardized path estimates in parentheses (except for the PD/A and SOC dummy codes, where standardized paths—in which only the APPQ factors are standardized—are provided in place of completely standardized estimates, in which both the predictor and APPQ factors are standardized). Paths with different superscripts differ significantly in magnitude ($p < .05$).

$^{*}p < .05$.
$^{**}p < .01$.
$^{***}p < .001$.

on their conjectured factor, and one item (13) functioned differently for males and females. In the former case, items posited to be indicators of interoceptive fear of alcohol use, coffee consumption, and receiving gas at the dentist had primary loadings on the Agoraphobia factor (and did not have salient secondary loadings on Interoceptive; see Table 1). It is noteworthy that these three items had the lowest loadings on the Interoceptive component of the solution reported in Rapee et al. (1994/1995), and in fact one item (7) had a salient cross-loading on the Agoraphobia component (.33). Moreover, Rapee et al. (1994/1995) provided a solution that included several agoraphobia and interoceptive items that were ultimately excluded from the 27-item version of the scale, and thus their factor loading matrix would differ (to an unknown degree) had their analysis been confined to these 27 items. It is likely that this aspect, in tandem with several methodological refinements in the current study (e.g., use of cross-validated analyses based

on the common factor model, adherence to more stringent item retention criteria), accounted in part for the discrepancies of the current findings and those of Rapee et al. (1994/1995).

Nevertheless, it is important to consider why items that would appear to have a strong conceptual link to interoceptive fear (e.g., item 17: caffeine ingestion) failed to load on the Interoceptive factor. Although contributing reasons may be item-specific,[8] an overarching explanation for this result is that the majority of APPQ-Interoceptive items pertain to physical exertion (e.g., hiking, vigorous exercise and sports, running up stairs). In addition to possibly under-representing the construct of interoceptive fear, such subscale composition would produce differential covariation among physical exertion items which thereby form a distinct factor at the exclusion of the few "interoceptive" items unrelated to physical activity (e.g., drinking a strong cup of coffee). If interoceptive fear is indeed believed to be a construct more far-reaching than physical exertion, then further revisions of the APPQ should be considered that contain a more balanced set of items assessing a broader range of activities and physical symptoms (e.g., watching scary movies, spinning in a chair, standing in a very hot room, sexual activity). Recommendations for additional refinement of the APPQ are underscored by the current unavailability of alternative clinical measures of the fear of sensation-producing activities—features that are key targets in empirically validated treatments of panic disorder (Chambless & Ollendick, 2000; Craske & Zucker, 2001).

On the other hand, it is possible that such results indicate that the construct of interoceptive fear possesses poor discriminant validity in relation to agoraphobia. In addition to findings that some items considered to reflect interoceptive fear had primary loadings on the Agoraphobia factor, concurrent validity analyses indicated that while the APPQ Interoceptive factor was more strongly related to ADIS-IV-L Fear of Panic and ASI Physical Concerns than was the APPQ Social factor, the strongest relationships with these criterion variables were with APPQ Agoraphobia (see Table 4). Similarly, Rapee et al. (1994/1995) found that APPQ Interoceptive and APPQ Agoraphobia were related to the ASI total score at statistically equal magnitudes (.47 and .50, respectively). Although methodological factors should not be ignored (e.g., possible narrow range of the current APPQ-Interoceptive scale, appropriateness of ADIS-IV-L Fear of Panic and ASI Physical Concerns as validation measures), these results prompt the substantive question of whether situational fear and the fear of sensation-producing activities represent distinct constructs or reflect features of a broader dimension of panic apprehension (i.e., general fearfulness of situations and activities that heighten the perceived risk of panic attacks). Prior findings of significant reductions in anxiety sensitivity, fear of panic attacks, and panic attack frequency in context of treatments focused exclusively on agoraphobic avoidance (i.e., situational exposure) could be viewed as challenging the distinctiveness of interoceptive and agoraphobic fearfulness (e.g., Ghosh & Marks, 1987; Soechting et al., 1998).

Nevertheless, weaker findings regarding the concurrent validity of the APPQ-Interoceptive dimension in relation to APPQ-Agoraphobia may be due in large part to the fact that it is also a measure of feared situations and activities. Although the APPQ was developed with the intent of measuring *situational* fear of agoraphobic, social, and sensation-producing activities

[8]For instance, the poor psychometric behavior of item 7 (getting gas at a dentist) as an indicator of interoceptive fear could be due to the somewhat outdated and rare nature of this activity, and the fact that responding could be strongly influenced by agoraphobia (i.e., fear of being confined in the dentist chair).

(Rapee et al., 1994/1995), the psychometric properties of the APPQ-Interoceptive scale might be fostered by the inclusion of items that assess fear of bodily sensations. Refinement of the subscale in this manner would be consistent with conceptual models of panic disorder that emphasize misappraisal or conditioned fear of interoceptive cues (e.g., accelerated heart rate, shortness of breath) as opposed to activities that may produce such sensations (e.g., Barlow, 2002; Clark, 1986).

A series of multiple-group CFA analyses indicated that the measurement properties of the APPQ were invariant in male and female patients, with one notable exception. Although item 13 (walking alone in isolated areas) had an equivalent relationship with the underlying dimension of Agoraphobia in both sexes (factor loadings = .65 and .62 for males and females, respectively), its intercept was highly noninvariant and pointed to a substantively significant sex bias—namely, a propensity to yield higher observed scores in females than males (even at equivalent levels of the underlying dimension of Agoraphobia) because of sex differences in the range of activities relating to personal safety. Although the strong evidence of differential item functioning might suggest removal of this item (especially in studies of sex differences in panic disorder and agoraphobia), it is important to note that the current sample was drawn from an urban setting. It would be of interest to examine the generalizability of these results in samples drawn from other settings (e.g., rural and suburban) where potentially this bias would be less pronounced (e.g., this sex difference may not be evident in nonurban communities associated with higher levels of perceived personal safety). Moreover, as the current study's analyses were conducted using a predominantly Caucasian outpatient sample, additional research should examine the generalizability of the APPQ's measurement properties across culturally and racially diverse patients.

Aside from the scalar noninvariance of item 13, the study's analyses provided strong psychometric support for the original APPQ Agoraphobia and Social Phobia subscales (e.g., stable and well-defined factor structures, high levels of scale reliability and factor determinacy). The concurrent validity of these factors was also quite strong. APPQ-Social Phobia had strong differential associations with clinical and questionnaire measures of social phobia, as did APPQ-Agoraphobia with ADIS-IV-L ratings of PDA and agoraphobic avoidance. Indeed, despite the differing methods (questionnaire vs. clinical rating), validity coefficients for these APPQ factors with ADIS-IV-L measures were quite impressive (e.g., .88 for APPQ-Social Phobia and ADIS-IV-L Social Fear; .78 for APPQ-Agoraphobia and ADIS-IV-L Agoraphobia). Although alternative measures of social anxiety and agoraphobia exist (e.g., Chambless, Caputo, Jasin, Gracely, & Williams, 1985; Liebowitz, 1987; Turner, Beidel, Dancu, & Stanley, 1989), a potential practical advantage of the APPQ is its relative brevity (Agoraphobia and Social Phobia are assessed by 9- and 10-item subscales, respectively, within a single measure). Specifically, these results suggest that the APPQ offers reliable, brief dimensional assessment of these core anxiety disorder constructs at high levels of concurrent validity with more time-intensive measures.

Appendix A. Albany Panic and Phobia Questionnaire

Please rate, on the following scale, the *amount of fear* that you think you would experience in each of the situations listed below if they were to occur *in the next week*. Try to imagine yourself

actually doing each activity and how you would feel (the three items in bold were excluded in the final factor models).

Fear Scale

0	1	2	3	4	5	6	7	8
no fear		slight fear		moderate fear		marked fear		extreme fear

1. Talking to people
2. Going through a car wash
3. Playing a vigorous sport on a hot day
4. Blowing up an airbed quickly
5. Eating in front of others
6. Hiking on a hot day
7. **Getting gas at a dentist**
8. Interrupting a meeting
9. Giving a speech
10. Exercising vigorously alone
11. Going long distances from home alone
12. Introducing yourself to groups
13. Walking alone in isolated areas
14. Driving on highways
15. Wearing striking clothes
16. Possibility of getting lost
17. **Drinking a strong cup of coffee**
18. Sitting in the center of a cinema
19. Running up stairs
20. Riding on a subway
21. Speaking on the telephone
22. Meeting strangers
23. Writing in front of others
24. Entering a room full of people
25. Staying overnight away from home
26. **Feeling the effects of alcohol**
27. Going over a long, low bridge.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Antony, M. M. (2001). Measures for panic disorder and agoraphobia. In M. M. Antony, & S. M. Orsillo (Eds.). *Practitioner's guide to empirically based measures of anxiety* (pp. 95–125). Dordrecht, Netherlands: Kluwer Academic Publishers.

Barlow, D. H. (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed.). New York: Guilford Press.

Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2000). Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: A randomized control trial. *Journal of the American Medical Association*, *283*, 2529–2536.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238–246.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Brown, T. A. (2003). Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects?. *Behaviour Research and Therapy*, *41*, 1411–1426.

Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the DSM-IV anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*, *107*, 179–192.

Brown, T. A., Di Nardo, P. A., Lehman, C. L., & Campbell, L. A. (2001). Reliability of DSM-IV anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology*, *110*, 49–58.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456–466.

Carter, M. M., Suchday, S., & Gore, K. L. (2001). The utility of the ASI factors in predicting response to voluntary hyperventilation among nonclinical participants. *Journal of Anxiety Disorders*, *15*, 217–230.

Chambless, D. L., & Ollendick, T. H. (2000). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, *52*, 685–716.

Chambless, D. L., Caputo, G. C., Jasin, S. E., Gracely, E. J., & Williams, C. (1985). The mobility inventory for agoraphobia. *Behaviour Research and Therapy*, *23*, 35–44.

Clark, D. M. (1986). A cognitive approach to panic. *Behaviour Research and Therapy*, *24*, 461–470.

Craske, M. G., & Barlow, D. H. (2001). Panic disorder and agoraphobia. In D. H. Barlow (Ed.). *Clinical handbook of psychological disorders: A step-by-step treatment manual*, (3rd ed.) (pp. 1–59). New York: Guilford Press.

Craske, M. G., & Zucker, B. G. (2001). Consideration of the APA practice guideline for the treatment of patients with panic disorder: Strengths and limitations for behavior therapy. *Behavior Therapy*, *32*, 259–281.

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety disorders interview schedule for DSM-IV: Lifetime version (ADIS-IV-L)*. San Antonio, TX: Psychological Corporation.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, *11*, 572–580.

Ghosh, A., & Marks, I. M. (1987). Self-treatment of agoraphobia by exposure. *Behavior Therapy*, *18*, 3–16.

Goldstein, A. J., & Chambless, D. L. (1978). A reanalysis of agoraphobia. *Behavior Therapy*, *9*, 47–59.

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

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods*, *6*, 430–450.

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

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*, 183–202.

Jöreskog, K., & Sörbom, D. (2002). *LISREL 8.52 [Computer software]*. Chicago: Scientific Software International.

Liebowitz, M. R. (1987). Social Phobia. *Modern Problems in Pharmacopsychiatry*, *22*, 141–173.

Lord, F. M., & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*, 107–120.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149.

Marks, I. M. (1969). *Fears and phobias*. London: Heinemann.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors?. *Journal of Personality and Social Psychology*, *70*, 810–819.

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety fear. *Behaviour Research and Therapy*, *36*, 455–470.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Muthén, L. K., & Muthén, B. O. (2002). *Mplus 2.12 (Computer software)*. Los Angeles: Author.

Novy, D. M., Stanley, M. A., Averill, P., & Daza, P. (2001). Psychometric comparability of English- and Spanish-language measures of anxiety and related affective symptoms. *Psychological Assessment*, *13*, 347–355.

Peterson, R. A., & Reiss, S. (1992). *Anxiety Sensitivity Index manual* (2nd ed.). Worthington, OH: IDS Publishing.

Rapee, R. M., Craske, M. G., & Barlow, D. H. (1994/1995). Assessment instrument for panic disorder that includes fear of sensation-producing activities: The Albany Panic and Phobia Questionnaire. *Anxiety*, *1*, 114–122.

Raykov, T. (2004). Behavioral scale reliability and measurement invariance evaluation using latent variable modeling. *Behavior Therapy*, *35*, 299–331.

Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, *54*, 315–323.

Snaith, R. P. (1968). A clinical investigation of phobias. *British Journal of Psychiatry*, *114*, 673–697.

Soechting, I., Taylor, S., Freeman, W., DeKoning, E., Segerstrom, S., & Thordarson, D. (1998). In vivo exposure for panic disorder and agoraphobia: Does a cognitive rationale enhance treatment efficacy?. In E. Sanavio (Ed.). *Behavior and cognitive therapy today: Essays in honor of Hans J. Eysenck* (pp. 279–288). Oxford, England: Elsevier Science.

Spiegel, D. A., & Barlow, D. H. (2000). *Sensation-focused intensive treatment for panic disorder and specific phobias*. Unpublished treatment manual, Boston University.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25*, 173–180.

Thurstone, L. L. (1947). *Multiple-factor analysis*. Chicago: University of Chicago Press.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10.

Turner, S. M., Beidel, D. C., Dancu, C. V., & Stanley, M. A. (1989). An empirically derived inventory to measure social fears and anxiety: The social phobia and anxiety inventory. *Psychological Assessment*, *1*, 35–40.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–69.

Veljaca, K. A., & Rapee, R. M. (1998). Detection of negative and positive audience behaviours by socially anxious subjects. *Behaviour Research and Therapy*, *36*, 311–321.

White, K. S., & Barlow, D. H. (2002). Panic disorder and agoraphobia. In D. H. Barlow (Ed.). *Anxiety and its disorders: The nature and treatment of anxiety and panic*, (2nd ed.) (pp. 328–379). New York: Guilford Press.

Widaman, K. F. (1993). Common factor analysis versus principal components analysis: Differential bias in representing model parameters?. *Multivariate Behavioral Research*, *28*, 263–311.

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., Mohlman, J., & Hong, N. (1999). Dimensions of anxiety sensitivity. In S. Taylor (Ed.). *Anxiety sensitivity: Theory, research, and treatment of the fear of anxiety* (pp. 83–114). Mahwah, NJ: Erlbaum.

Zinbarg, R. E., Brown, T. A., Barlow, D. H., & Rapee, R. M. (2001). Anxiety sensitivity, panic, and depressed mood: A reanalysis teasing apart the contributions of the two levels in the hierarchical structure of the Anxiety Sensitivity Index. *Journal of Abnormal Psychology*, *110*, 372–377.