Psychological Assessment
2004, Vol. 16, No. 3, 244–254

1040-3590/04/$12.00 DOI: 10.1037/1040-3590.16.3.244

Psychometric Evaluation of the Behavioral Inhibition/Behavioral Activation Scales in a Large Sample of Outpatients With Anxiety and Mood Disorders

Laura Campbell-Sills, Gabrielle I. Liverant, and Timothy A. Brown
Boston University

The latent structure, reliability, and validity of the Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS; C. L. Carver & T. L. White, 1994) were examined in a large sample of outpatients (N = 1,825) with anxiety and mood disorders. Four subsamples were used for exploratory and confirmatory factor analyses. In addition to generally upholding a latent structure found previously in nonclinical samples, results indicated measurement invariance of the BIS/BAS between genders and a higher order structure of the BAS scales. Convergent and discriminant validity of the BIS/BAS were supported by findings that the subscales correlated most strongly with measures of neighboring personality constructs (e.g., BIS with neuroticism, BAS with positive affect) than with measures of current anxiety and depression symptoms. Overall, the results support the psychometric properties of the BIS/BAS in this clinical sample.

Carver and White's (1994) Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS) have become widely used measures of personality in psychological research. The BIS/BAS scales are based on Gray's (1987) biopsychological theory of emotion, which posits that three systems—the behavioral inhibition, behavioral activation, and fight–flight systems—underlie learning, affective, and motivational processes. The behavioral inhibition system is hypothesized to direct behavior in response to threats and novel stimuli, whereas the behavioral activation system modulates behavior in response to incentives. The fight–flight system coordinates rapid response to immediate threats to survival. Varying degrees of strength and sensitivity of these physiological systems lead to characteristic patterns of responding that may comprise a significant component of human personality.

Carver and White (1994) reviewed Gray's (1987) theory of brain function and behavior in their original description of the BIS/BAS. The BIS is hypothesized to control the experience of anxiety in response to novel or threatening cues and to inhibit behavior that might produce negative or painful consequences. The physiologically independent BAS is thought to produce positive feelings in response to signals of reward and to provoke movement toward desired goals. Although Gray specified physiological, affective, and behavioral consequences that resulted from varying strengths of these systems, Carver and White elected to focus primarily on the affective consequences of the BIS and BAS in their self-report measure.

More specifically, the authors of the BIS/BAS focused on vulnerability to experiencing emotions in certain contexts rather than the everyday experience of emotions. Carver and White (1994) reasoned that everyday experiences of emotions like anxiety and happiness were imperfect indicators of BIS and BAS strength because other factors could influence the frequency and intensity of emotions. A person with a very sensitive BIS might be prone to experiencing anxiety in novel situations but could learn to avoid such situations and thereby reduce the everyday experience of anxiety to a minimum. Carver and White constructed items that asked respondents to consider their probable affective responses to a variety of situations (e.g., making a mistake, being criticized, engaging in fun activities) to increase the likelihood of isolating dispositional differences.

Following this general approach to operationalizing Gray's (1987) theory, Carver and White (1994) measured BIS strength as the degree of negative affect experienced in response to threatening cues (e.g., "I feel worried when I think I have done poorly at something"). In contrast, BAS strength was primarily defined as degree of positive affect experienced in response to rewards. However, in addition to measuring affective responses to incentives, Carver and White's BAS scale also incorporated items that assessed for behavioral responses to rewards. Rapid, determined movement toward goals and action motivated by a desire to experience pleasure are measured in the BAS scale, whereas behavioral avoidance is not measured in the BIS scale. The focus on behavioral responses to impending reward may be one basis for another important difference between the scales: The BIS scale has been consistently described as unidimensional, whereas the BAS scale has been described as having three distinct factors (more discussion of the factor structure of the BIS/BAS is provided below).

Carver and White (1994) labeled the three factors of the BAS scale: Reward Responsiveness, Drive, and Fun Seeking. Examination of the items that comprise these scales reveals that the Reward Responsiveness scale exclusively assesses affective responding (e.g., "When I get something I want, I feel excited and energized"), the Drive scale exclusively assesses behavioral re-

Laura Campbell-Sills, Gabrielle I. Liverant, and Timothy A. Brown, Department of Psychology, Boston University.

Correspondence concerning this article should be addressed to Timothy A. Brown, Center for Anxiety and Related Disorders, Department of Psychology, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215. E-mail: tabrown@bu.edu

sponding (e.g., “If I see a chance to get something I want, I move on it right away”), and the Fun Seeking scale assesses for both affective and behavioral responding (e.g., “I crave excitement and new sensations” and “I will often do things for no other reason than that they might be fun”). Carver and White indicated that they avoided behavioral items in constructing the BIS scale because it had been demonstrated that “how people react behaviorally when they confront circumstances that induce anxiety depends on a wide range of factors other than the existence of the anxiety state per se” (p. 322). However, they noted that they adopted a more “divergent” strategy in generating BAS items due to the “absence of complete consensus about exactly how BAS sensitivity is likely to be manifest” (p. 322). The greater depth of knowledge pertaining to the BIS may have allowed for a greater number and variety of items that captured the affective component of the BIS, whereas items pertaining to behavioral approach may have been needed to augment the BAS scale. Regardless of the basis for this discrepancy in the character of the items comprising the scales, it has obvious implications for the structure of the BIS/BAS items (i.e., behavioral and affective items may not tend to cluster together as much as solely affective items do).

Using a college student sample, Carver and White (1994) evaluated their instrument with principal-components analysis (PCA) and found the aforementioned components of Behavioral Inhibition, Reward Responsiveness, Drive, and Fun Seeking. Subsequent investigations have also reported four-factor solutions based on exploratory analyses conducted in college and community samples (Heubeck, Wilkinson, & Cologon, 1998; Jorm et al., 1999; Ross, Millis, Bonebright, & Bailley, 2002). However, two of these studies also reported results suggesting possible two-, three-, or five-factor solutions (Heubeck et al., 1998; Jorm et al., 1999). These discrepancies were attributed in part to differences in estimation procedures (PCA vs. principal-axis factoring; Heubeck et al., 1998) and factor selection guidelines (Kaiser–Guttman rule vs. scree test; Jorm et al., 1999), and both studies reported four-factor solutions in an effort to “maximize the comparability” with the findings of Carver and White (1994).

Recent studies have also examined the latent structure of the BIS/BAS using confirmatory factor analysis (CFA). Heubeck et al. (1998) used CFA to test two-, three-, and four-factor models of the BIS/BAS scales. Although a correlated four-factor model provided the best fit to the data, goodness-of-fit statistics did not meet conventional criteria for adequate model fit (e.g., goodness-of-fit index = .83, comparative fit index [CFI] = .80). In contrast, Ross et al. (2002) tested one-, two-, and four-factor models in a college student sample, and found that the four-factor model met most criteria for acceptable fit, and was clearly superior to competing models.

Despite discordant findings on the BIS/BAS’s latent structure, results of studies investigating the validity of these scales have been quite favorable. The BIS scale correlates highly with negative affect, neuroticism, and processing of unpleasant information, whereas the BAS scales are strongly related to positive affect, extraversion, and processing of pleasant information (Gomez & Gomez, 2002; Heubeck et al., 1998; Jorm et al., 1999). Psychophysiological studies have demonstrated that higher BAS scores correlate with relatively greater left- than right-frontal cortical activity (Harmon-Jones & Allen, 1997), and higher BIS scores correlate with greater relative right-frontal activation (Sutton & Davidson, 1997). These results converge with previous findings linking left-frontal activity with behavioral approach tendencies and right-frontal activity with behavioral avoidance (Davidson, 1992). Overall, the BIS/BAS has demonstrated conceptually meaningful relationships to measures of personality, information processing, and cerebral activity.

The BIS/BAS is potentially important for psychopathological research, because it affords a method for measuring components of personality that may relate to specific psychological disorders and implicate particular neuroanatomical pathways. The behavioral inhibition and fight–flight systems have featured prominently in recent conceptualizations of anxiety and mood disorders (e.g., Barlow, 2002). Similarly, researchers have hypothesized that BAS hyperactivity is associated with disorders characterized by elevated mood and impulsivity (Meyer, Johnson, & Carver, 1999; Meyers, Johnson, & Winters, 2001), whereas BAS hypoactivity is related to disorders that involve low mood, withdrawal, and decreased pursuit of goals. For example, a recent study linked lower BAS levels to increased severity and diminished treatment response in a sample of depressed patients (Kasch, Rottenberg, Arnow, & Gotlib, 2002).

Despite the increasing interest and use of the BIS/BAS in psychopathological research (e.g., Campbell, Brown, & Grisham, 2003; Kasch et al., 2002), no studies have examined the psychometric properties of this measure in clinical samples. Although frequently overlooked in applied clinical research, classical test theory underscores the importance of evaluating the generalizability of a scale’s measurement properties for use in novel populations (Lord & Novick, 1968; McDonald, 1999). Indeed, clinical and nonclinical samples differ in many ways that could result in disparate covariances and distributions among BIS/BAS items, and hence discordant latent structures. For instance, as noted by Widiger, Verheul, and van den Brink (1999), “personality and psychopathology influence the presentation or appearance of one another” (p. 347) in a variety of manners that may be more prominent in clinical samples (e.g., differential susceptibility to mood–state distortion or acquiescence response set). Moreover, the affective circumstances under which clinical and nonclinical groups complete the questionnaire may differ markedly, in that clinical respondents (especially at the time of diagnostic assessment or treatment seeking) are typically experiencing persistent and abnormal emotional distress. As the majority of the items of the BIS/BAS inquire about affective experience, clinical participants may focus on and respond to different aspects of the items than do nonclinical respondents. In addition, individuals with anxiety and mood disorders may reflect on a qualitatively different set of recent life experiences and emotional experiences in forming their responses to the BIS/BAS items. Thus, although the generalizability or invariance of a scale’s measurement properties should never be simply assumed, these issues are particularly salient when the possibility exists that distinct examinee populations perceive items or attributes sufficiently differently to raise the issue of whether the same constructs are being measured in these groups (McDonald, 1999).

Similarly, although prior studies have reported significant differences between men and women on the BIS/BAS scales (Carver & White, 1994; Jorm et al., 1999), these comparisons were undertaken without first establishing measurement invariance of the BIS/BAS between genders. To make valid comparisons of groups of persons with regard to their level on a given trait or attribute, one must determine that the measurement properties of the testing instrument are equivalent in these groups with regard to the construct being measured and the ability to provide unbiased estimates of individuals' relative standing on the latent dimension (cf. Byrne, Shavelson, & Muthén, 1989; Reise, Widaman, & Pugh, 1993). In applied research, this fundamental aspect of psychometrics is often assumed or ignored (Reise et al., 1993) despite the possibility that mean differences between groups reflect substantively misleading artifacts of noninvariant measurement (e.g., item content or wording that is biased against a given group) rather than true differences in the latent construct (for a recent example of salient noninvariance of a clinical measure between male and female patients, see Brown, White, & Barlow, in press).

The potential sources of noninvariance between genders often vary across assessment contexts; in the case of personality measures such as the BIS/BAS, differences in the affective and behavioral patterns of men and women may be particularly germane. For instance, research suggests that different emotional "display rules" exist for men and women (Brody, 1999). Certain emotions such as pride and aggression appear to be more permissible for men to express in social contexts, whereas others such as anxiety and sadness seem more acceptable for women to display. If men and women indeed learn different rules that govern the expression of emotion, responses to items on the BIS/BAS may differ in men and women even if the actual levels of responsiveness of these systems do not. Specifically, it may be less socially desirable for men to endorse the experience of anxiety contained in the BIS items, whereas women may be less apt to endorse the aggressive style implicit in the Drive items of the BAS scale.

The current study aimed to provide an extensive psychometric analysis of the BIS/BAS in a large sample of patients with anxiety and mood disorders, and to potentially clarify discordant results regarding its factor structure. In addition, tests of measurement invariance were conducted in male and female patients. Unlike prior studies, the analysis of BIS/BAS latent structure also included evaluation of a possible higher order behavioral activation (BA) factor. The higher order analysis was relevant to the conceptual position that a broad BA dimension (representing the neurological system specified by Gray, 1987) was underlying the three BAS factors that had been reported in previous studies (e.g., Carver & White, 1994). It was hypothesized that a higher order solution would indicate that considerable covariance among the BAS items could be explained by a broader BA dimension, as well as the lower order factors of Reward Responsiveness, Drive, and Fun Seeking. Finally, in an effort to augment the already promising literature on the validity of the BIS/BAS, we examined its concurrent validity in relation to measures of personality and clinical psychopathology. Specifically, we predicted that the BIS/BAS subscales would be more strongly related to measures of conceptually similar personality traits (e.g., neuroticism, positive affect) than to measures of current anxiety and depression symptoms.

Method

Participants

The sample consisted of 1,825 patients who presented for assessment and treatment at a specialty anxiety and mood disorders clinic. Women constituted the larger portion of the sample (61.0%); average age was 33.27 years ($SD = 11.11$, range = 18 to 75). Most participants were Caucasian (88.9%), with smaller numbers of individuals identifying as Asian (4.1%), African American (3.5%), Hispanic (2.9%), and other (0.7%). Diagnoses were established with the Anxiety Disorders Interview Schedule for *DSM-IV*: Lifetime version (ADIS-IV-L; Di Nardo, Brown, & Barlow, 1994), a semistructured interview designed to reliably diagnose the *Diagnostic and Statistical Manual of Mental Disorders—IV* (American Psychiatric Association, 1994) anxiety, mood, somatoform, and substance use disorders, and to screen for the presence of other conditions (e.g., psychotic disorders; cf. Brown, Di Nardo, Lehman, & Campbell, 2001). The breakdown of current disorders in the sample was as follows: social phobia ($n = 764$), panic disorder with or without agoraphobia ($n = 688$), generalized anxiety disorder ($n = 381$), specific phobia ($n = 381$), obsessive–compulsive disorder ($n = 213$), posttraumatic stress disorder ($n = 49$), agoraphobia without a history of panic disorder ($n = 35$), major depression ($n = 474$), dysthymic disorder ($n = 154$), and other anxiety/mood disorder (e.g., anxiety or depressive disorder not otherwise specified; $n = 258$).

Measures

As noted earlier, the BIS/BAS (Carver & White, 1994) is a 20-item instrument designed to measure behavioral inhibition (i.e., concern over and reactivity to aversive events) and behavioral activation (i.e., responsiveness to incentives, drive, and fun seeking). Participants respond to the items using a 1 to 4 scale, ranging from 1 (*quite untrue of you*) to 4 (*quite true of you*). After reverse-scoring two items on the BIS, subscale scores are computed by summing the scores corresponding to items on the BIS scale (7 items), the Reward Responsiveness scale (5 items), the Drive scale (4 items), and the Fun Seeking scale (4 items). The Reward Responsiveness, Drive, and Fun Seeking subscales comprise the BAS.

In addition to the BIS/BAS, the following questionnaires were administered: Beck Depression Inventory (BDI; Beck & Steer, 1987), Beck Anxiety Inventory (BAI; Beck & Steer, 1990), Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995), Eysenck Personality Inventory–Neuroticism Scale (EPI-N; Eysenck & Eysenck, 1968), and the trait ("in general") version of the Positive and Negative Affect Scales (PANAS; Watson, Clark, & Tellegen, 1988). These measures were used in concurrent validity analyses as indicators of Depression (BDI, DASS-Depression), Anxious Arousal (BAI, DASS-Anxiety), Negative Affect (PANAS-Negative Affect [PANAS-N]), Neuroticism (EPI-N), and Positive Affect (PANAS-Positive Affect [PANAS-P]).

Procedure

Participants completed the BIS/BAS scales and other questionnaires as part of their initial evaluation (consisting of the ADIS–IV–L and a questionnaire battery). To thoroughly replicate the BIS/BAS factor structure, we randomly divided the full sample ($N = 1,825$) into four subsamples. The demographic composition of these subsamples was equivalent; for example, percentages of women in the four samples ranged from 57% to 64%, $\chi^2(3, N = 1,825) = 3.68$, *ns*; mean age ranged from 33.12 to 33.35 years, $F(3, 1821) < 1$.

Given that a strong empirical basis for CFA was lacking (i.e., consistent evidence with regard to the appropriate number of factors, patterning of item-factor relationships, and appropriate error theory; see e.g., Heubeck et al., 1998; Jorm et al., 1999; Ross et al., 2002), analysis of the BIS/BAS's latent structure was conducted in a sequential fashion entailing increasingly restrictive solutions (i.e., exploratory factor analysis [EFA], exploratory factor analysis with the CFA framework [E/CFA], CFA, multiple-group

CFA models).[1] Sample 1 ($n = 300$) was used to conduct an initial EFA of the original 20-item BIS/BAS. Sample 2 ($n = 300$) was used to conduct an independent EFA to confirm the factor structure found in the first analysis. Samples 1 and 2 were combined ($n = 600$) to test revised solutions (e.g., solutions that removed problematic items) and obtain estimates for the final EFA solution. Sample 3 ($n = 300$) and Sample 4 ($n = 925$) were used to replicate the final BIS/BAS structure obtained in combined Samples 1 and 2. As a precursor to CFA, Sample 3 was used for an E/CFA. A common 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 not detected by EFA (e.g., salient residual covariances). The researcher is then faced with extensive post hoc model testing subject to the criticisms of specification searches in a single data set (MacCallum, 1986). Thus, the results of the analyses conducted in Samples 1–3 greatly informed the specification of an appropriate CFA solution in Sample 4. Sample 4 was also used for analyses of measurement invariance between male and female patients, a higher order BA solution, and concurrent validity of the BIS/BAS with measures of personality and anxiety/mood disorder symptoms.

Data Analysis

The sample variance–covariance matrices were analyzed using latent variable software programs and maximum-likelihood minimization functions (LISREL 8.52, Jöreskog & Sörbom, 2001; Mplus 2.12, Muthén & Muthén, 1998). Goodness of fit was evaluated using the root-mean-square error of approximation (RMSEA; Steiger, 1990) and its 90% confidence interval (90% CI), p value for test of close fit (CFit; RMSEA < 0.05), standardized root-mean-square residual (SRMR; Jöreskog & Sörbom, 1986), and CFI (Bentler, 1990). 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 chi-square difference tests.

Results

EFA

Using Sample 1 ($n = 300$), we submitted the 20 BIS/BAS items to an EFA with maximum-likelihood estimation and promax rotation. Acceptability of the models (e.g., factor selection) was evaluated by goodness-of-model fit (RMSEA $< .08$, CFit, 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). A four-factor solution fit the data best, $\chi^2(116) = 255.07$, $p < .001$; RMSEA $= .063$, 90% CI $=$ .053–.074, CFit $= .02$ (eigenvalues > 1.0 for the unreduced correlation matrix were: 4.02, 3.14, 1.82, 1.33, 1.21, 1.01; RMSEAs for two- and three-factor solutions were .11 and .08, respectively). However, as shown in Table 1, three items did not have salient loadings on any factor (Item 5 “When I’m doing well at something, I love to keep at it,” Item 13 “Even if something bad is about to happen to me, I rarely experience fear or nervousness,” and Item 16 “I have very few fears compared to my friends”).

A second EFA was conducted with Sample 2. As before, a four-factor solution provided a good fit to the data, $\chi^2(116) = 211.19$, $p < .001$; RMSEA $= .052$, 90% CI $=$ .041–.063; CFit $=$ .36 (eigenvalues $> 1.0 =$ 4.96, 3.01, 1.47, 1.21, 1.15). Items 5, 13, and 16 again had the lowest loadings on their corresponding

Table 1
Latent Structure of the 20-Item Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS): Exploratory Factor Analyses in Samples 1 and 2 (Ns = 300)

BIS/BAS item	Behavioral Inhibition S1	Behavioral Inhibition S2	Reward S1	Reward S2	Drive S1	Drive S2	Fun Seeking S1	Fun Seeking S2
1	**.333**	**.444**	.164	.200	.018	.044	.157	.068
4	**.547**	**.570**	.075	.132	.047	.116	.006	.085
7	**.758**	**.683**	.191	.005	.066	.095	.071	.069
10	**.817**	**.765**	.096	.042	.053	.039	.071	.151
13	.225	.247	.186	.086	.063	.065	.216	.078
15	**.595**	**.664**	.047	.080	.103	.036	.058	.018
16	.231	**.324**	.079	.070	.084	.085	.205	.219
2	.146	.077	**.757**	**.693**	.018	.065	.046	.082
5	.061	.025	.277	.216	.145	.219	.021	.077
8	.045	.066	**.555**	**.739**	.013	.093	.016	.032
11	.010	.055	**.450**	**.425**	.097	.029	.111	.115
14	.029	.003	**.478**	**.503**	.171	.165	.157	.054
3	.013	.034	.025	.121	**.791**	**.754**	.028	.077
6	.032	.062	.006	.099	**.866**	**.842**	.091	.110
9	.035	.100	.070	.171	**.675**	**.570**	.014	.050
12	.080	.030	.002	.013	**.535**	**.628**	.149	.093
17	.051	.044	.092	.204	.141	.094	**.720**	**.575**
18	.107	.094	.020	.056	.003	.088	**.811**	**.773**
19	.053	.019	.001	.072	.008	.081	**.805**	**.759**
20	.017	.059	.074	.015	.054	.011	**.533**	**.656**

Note. Exploratory factor analysis was conducted with maximum-likelihood estimation, promax rotation. Factor loadings $\geq .30$ are in bold. S1 = Sample 1; S2 = Sample 2.

factors, with two of them below .30 and the third (Item 16) only slightly above .30 (as compared with loadings of .44 to .84 for other items; see Table 1).

Items 5, 13, and 16 had consistently nonsalient loadings, suggesting that they were weakly related to their conjectured latent construct. Thus, an EFA was rerun in the combined Samples 1 and 2 ($n = 600$) after removal of these three items.[2] A four-factor model provided an acceptable fit to the data, $\chi^2(74) = 110.85$, $p < .01$; RMSEA $= .041$, 90% CI $=$ .040–.058; CFit $= .83$ (eigenvalues $> 1.0 =$ 4.77, 2.71, 1.43, and 1.16). The promax-rotated factor pattern matrix of this solution is presented in Table 2. The first factor, labeled Behavioral Inhibition (BIS) consisted of all the original BIS items, with the exception of the two reverse-worded items (Items 13 and 16). The second factor, labeled Reward

[1] Even if consistent evidence of the BIS/BAS’s latent form had been obtained in nonclinical samples (e.g., number of factors), conducting initial analyses in less restrictive exploratory frameworks was warranted given the unknown generalizability of these results to clinical samples (see introduction) and the failure of previous studies to consider the potential role of nonrandom measurement error (e.g., method effects arising from items with highly overlapping content).

[2] It is noteworthy that two of the three poorly loading items were the BIS/BAS’s only reverse-worded items. Often, reversed items are associated with psychometric complications such as method effects and relatively lower loadings on the substantive latent dimension (cf. Brown, 2003; Marsh, 1996).

Table 2
Latent Structure of the 17-Item Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS): Exploratory Factor Analysis (N = 600, Samples 1 and 2) and Exploratory Factor Analysis Conducted Within the Confirmatory Factor Analysis Framework (N = 300, Sample 3)

	Factor							
	Behavioral Inhibition		Reward		Drive		Fun Seeking	
BIS/BAS item	EFA	E/CFA	EFA	E/CFA	EFA	E/CFA	EFA	E/CFA
1	**.386**	**.520**	.134	.122	.018	.071	.108	.075
4	**.573**	**.840**	.021	.184	.018	.033	.053	.164
7	**.692**	**.744**	.048	.106	.006	.089	.020	.076
10	**.747**	**.653**	.004	.064	.026	.053	.064	.092
15[a]	**.655**	**.673**	.008	.000	.040	.000	.021	.000
2	.071	.248	**.737**	**.748**	.020	.006	.081	.023
8[a]	.041	.000	**.638**	**.513**	.054	.000	.027	.000
11	.039	.285	**.445**	**.679**	.043	.105	.101	.001
14	.030	.279	**.510**	**.701**	.157	.057	.079	.051
3	.016	.150	.077	.229	**.767**	**.730**	.052	.052
6[a]	.052	.000	.061	.000	**.867**	**.911**	.005	.000
9	.081	.154	.136	.226	**.631**	**.556**	.001	.007
12	.010	.101	.001	.185	**.585**	**.544**	.128	.006
17	.073	.137	.127	.189	.121	.085	**.656**	**.496**
18[a]	.069	.000	.054	.000	.035	.000	**.803**	**.743**
19	.017	.096	.062	.054	.035	.162	**.784**	**.825**
20	.007	.024	.029	.006	.039	.017	**.590**	**.539**
Determinacy	.885	.902	.856	.873	.921	.940	.910	.906

Note. Factor loadings ≥ .30 are in bold. EFA = exploratory factor analysis (maximum-likelihood extraction, promax rotation); E/CFA = exploratory factor analysis within the confirmatory factor analysis framework (maximum likelihood).
[a] Items were used as anchor indicators in the E/CFA analysis.

Responsiveness (Reward), consisted of all but one of the original items intended to load on this subscale (Item 5). The third and fourth factors, Drive and Fun Seeking (Fun), were identical to the original scale design. The BAS factors (Reward, Drive, and Fun) were significantly intercorrelated (*r*s = .46–.54), whereas none of the BAS factors demonstrated strong correlations with the BIS factor (*r*s = −.04–−.18).

Although reported infrequently in applied psychometric research, factor determinacy data are important in the evaluation of factor analytic findings (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 (cf. Grice, 2001). As shown in Table 2, all four factors evidenced a favorable level of determinacy (range = .86–.92), per the recommendations of Gorsuch (1983; i.e., >.80).

E/CFA

Sample 3 (n = 300) was used to cross-validate the four-factor EFA solution involving the 17 retained BIS/BAS items. As an intermediate step between EFA and CFA, the BIS/BAS data from Sample 3 were analyzed using an E/CFA approach (Jöreskog, 1969). Although underused in applied factor analysis research, E/CFA is a very helpful precursor to CFA. In this strategy, the CFA applies only to the identifying restrictions used in EFA (e.g., fix factor variances to 1.0 for m restrictions; for each factor, select an anchor item to freely load onto it but fix its cross-loadings to 0; allow nonanchor items to be freely estimated). Whereas this specification uses the same number of restrictions (m^2) and produces the same overall fit as EFA, the CFA maximum-likelihood estimation provides a great deal of additional information that can be important in developing realistic CFA models in subsequent analyses (e.g., standard errors to determine the significance of factor loadings and factor correlations). For instance, E/CFA provides modification indices for indicator error covariances that might point to the existence of minor factors or a more complex error structure that would be undetected by EFA. We hypothesized that the residuals of two sets of items would be correlated due to similar content and wording (Item 4 “I worry about making mistakes” with Item 15 “I feel worried when I think I have done poorly at something” and Item 3 “When I want something, I usually go all-out to get it” with Item 6 “I go out of my way to get things I want”).

Using Items 6, 8, 15, and 18 as anchor items, we found that the four-factor E/CFA model provided a good fit to the data, $\chi^2(74)$ = 118.06, $p < .001$; RMSEA = .045, 90% CI = .029–.059; CFit = .71; SRMR = .027; CFI = .97. As shown in Table 2, the magnitudes of primary loadings were strong (range = .50–.91) and statistically significant (range of zs = 6.23–18.00), and factor determinacies were high (range = .87–.94). The factor intercorrelations were as follows: BIS with Reward, Drive, and Fun (.36,

−.14, .18, respectively); Reward with Drive and Fun (.22, .40, respectively); Drive with Fun (.33). The correlations between BIS and Drive, BIS and Fun, and Reward and Drive were not statistically significant. Counter to our error theory predictions involving Items 4 and 15 and Items 3 and 6, modification indices pertaining to correlated residuals revealed no points of strain in the E/CFA model.

CFA

On the basis of the solutions obtained in Samples 1, 2, and 3, a four-factor CFA model was fit to the Sample 4 data (n = 925). This model provided a good fit to the data, $\chi^2(113) = 406.95$, $p <$ .001; RMSEA = .053, 90% CI = .048–.059; CFit = .18; SRMR = .045; CFI = .94. However, sequential fit diagnostic evaluation indicated that the points of ill-fit pertained to the error covariances of Items 3 and 6 and Items 4 and 15. Because these were the two highest modification indices pertaining to error covariances, and because these findings were consistent with the a priori error theory, the CFA solution was respecified correlating the residuals of these item pairs. In addition, fit diagnostics from this CFA suggested that Item 14 loaded on Drive in addition to Reward (modification index [MI] = 34.71, standardized expected parameter change = .29). Examination of Item 14's content ("When I see an opportunity for something I like, I get excited right away") revealed that the item contained both the focus on affective experience that the Reward items shared and the sense of urgency that most of the Drive items contained (cf. Item 9 "If I see a chance at something I want, I move on it right away"). Because this MI appeared to reflect a substantive relationship between Item 14 and the Drive factor, this path was freed in the revised solution.

The revised model provided a better fit to the data, $\chi^2_{\text{diff}}(3)$ = 149.66, $p < .001$; $\chi^2(110) = 257.29$, $p < .001$; RMSEA = .038, 90% CI = .032–.044; CFit = 1.00; SRMR = .037; CFI = .97. Fit diagnostics indicated no salient points of strain in the solution. The error variances between Items 3 and 6 and Items 4 and 15 were significantly correlated (rs = .12 and .16, respectively, ps < .001), as hypothesized on the basis of item wording. Although Item 14 demonstrated statistically significant loadings on both the Reward and Drive factors, the secondary loading on the Drive factor was relatively weak (.29). As shown in Table 3, the magnitudes of primary loadings were strong (range = .51–.79), and the factor determinacies were quite satisfactory (range = .87–.92). Factor intercorrelations were as follows: BIS with Reward, Drive, Fun (.27, −.03, −.09, respectively; Reward with Drive, Fun (.55, .45, respectively); Drive with Fun (.42). The correlation between BIS and Drive was not statistically significant.

We calculated the scale reliabilities of the four factors within the CFA model using the approach developed by Raykov (2001). This method reconciles the problems with Cronbach's alpha, which is a misestimator of scale reliability except in the rare and restrictive

Table 3
Latent Structure of the 17-Item Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS): Confirmatory Factor Analysis Using Sample 4 (N = 925) and Male (N = 365) and Female (N = 560) Patients

	Factor											
	Behavioral Inhibition			Reward			Drive			Fun Seeking		
Item	All	M	F	All	M	F	All	M	F	All	M	F
BIS/BAS item												
1	.507	.517	.488									
4	.609	.609	.650									
7	.696	.665	.707									
10	.708	.730	.677									
15	.577	.580	.579									
2				.700	.714	.663						
8				.643	.676	.587						
11				.506	.539	.492						
14				.487	.512	.477						
3							.756	.788	.732			
6							.748	.795	.722			
9							.771	.754	.785			
12							.740	.746	.735			
14							.287	.292	.267			
17										.690	.667	.703
18										.704	.753	.676
19										.788	.770	.799
20										.622	.629	.620
Determinacy	.872	.871	.873	.871	.883	.856	.917	.928	.912	.904	.906	.904
Reliability	.728	.728	.733	.745	.774	.771	.816	.840	.799	.795	.798	.795
M	16.98	16.57	17.24	13.40	12.99	13.67	13.80	13.65	13.90	10.05	10.03	10.06
SD	2.94	3.03	2.85	2.22	2.42	2.04	3.61	3.66	3.59	3.16	3.05	3.23

Note. Means and standard deviations are based on coarse factor scores (i.e., raw score composites). All = entire Sample 4 (N = 925); M = male patients; F = female patients.

instance when all elements of a multiple-item measure are tau-equivalent (Lord & Novick, 1968; McDonald, 1999; Raykov, 2001). 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 to total score variance, per the classic formula for scale reliability estimation (Lord & Novick, 1968). As shown in Table 3, the scale reliabilities were quite acceptable, ranging from .73 to .82.

Measurement Invariance of the BIS/BAS in Male and Female Patients

The measurement invariance (equal latent form, factor loadings, indicator intercepts) of the 17-item BIS/BAS was examined in male and female patients using multiple-groups CFA in Sample 4 ($n = 925$).[3] Prior to the multiple-groups solutions, CFAs were conducted separately to verify adequate fit in the male and female samples (*n*s = 365 and 560, respectively). The factor loadings, factor determinacies, and scale reliabilities obtained for men and women are provided in Table 3. Next, a two-group CFA was conducted to test equal BIS/BAS form between sexes. The simultaneous equal form solution fit the data well, $\chi^2(220) = 366.99$, $p < .001$; RMSEA = .038, 90% CI = .031–.045; CFit = 1.00; SRMR = .043; CFI = .97. Given evidence of equal form in the male and female samples, the next analysis addressed metric invariance by imposing equality constraints on the factor loadings of the male and female solutions. These constraints did not significantly degrade the fit of the solution, indicating that the factor loadings were equivalent in the male and female samples, $\chi^2_{\text{diff}}(14) = 16.33$, *ns*. The next analysis examined the scalar invariance of the BIS/BAS by placing equality constraints on the intercepts of the 17 items. Although these constraints resulted in a degradation in model fit, $\chi^2_{\text{diff}}(13) = 64.91$, $p < .001$, fit diagnostics revealed no salient localized points of strain in the solution (i.e., no standout MIs, no potential noninvariance with apparent substantive implications; cf. Brown et al., in press). Although suggesting no important noninvariance with respect to the item intercepts, these results possibly reflect the oversensitivity of chi-square in measurement invariance evaluation (cf. Cheung & Rensvold, 2002; Vandenberg & Lance, 2000).

Higher Order Factor Structure of the BAS Scales

Next, the viability of a higher order BA dimension was considered—specifically, that the BAS factors represent more differentiated components of the broader dimension of BA. Because of the just-identified nature of this model (i.e., a single higher order BA factor, three lower order BAS factors), the appropriateness of hierarchical structure was assessed in terms of the magnitude of factor loadings of the BAS factors and individual BAS items. Data from Sample 4 (men and women combined; $n = 975$) were used for these analyses.

Using factor correlations obtained from the first-order solution, results indicated that the Reward, Drive, and Fun factors loaded strongly on the higher order BA factor (loadings = .77, .72, and .59, respectively, *p*s < .001). A Schmid–Leiman (Schmid & Leiman, 1957) transformation was then conducted to obtain loadings between the BAS items and the higher order BA factor, and residualized loadings between the BAS items and the three lower order BAS factors. As shown in Table 4, every BAS item had a salient loading on the higher order BA factor (range = .37–.58). Moreover, with the exception of the Item 14 cross-loading, all of the residualized first-order loadings were above .30 suggesting that the component dimensions of Reward, Drive, and Fun Seeking accounted for unique salient variance in their constituent indicators, above the variance explained by the higher order BA factor (as shown in Table 4, residualized primary loadings ranged from .31 to .45 for Reward, from .52 to .54 for Drive, and from .50 to .64 for Fun).

The scale reliability of the higher order BA factor was estimated using the LISREL parameterization developed by Raykov and Shrout (2002) for composites with general structure. Indeed, the scale reliability estimate of the general BA factor was quite favorable ($\rho = .88$).

Relationships of the BIS/BAS Factors With Negative Affect, Neuroticism, Positive Affect, Anxious Arousal, and Depression

To evaluate the concurrent validity of the BIS/BAS scales, we developed additional models using the Sample 4 data to determine the associations of the four BIS/BAS factors with the dimensions of negative affect, neuroticism, positive affect, anxious arousal, and depression (cf. Brown, Chorpita, & Barlow, 1998; L. A. Clark, Watson, & Mineka, 1994). Several patients in Sample 4 had missing data for at least one of these validity indicators, resulting in a sample size of 749 for these analyses. Along the lines of Brown et al. (1998), the BAI and DASS-Anxiety scales were specified as indicators of a latent Anxious Arousal factor, and the BDI and DASS-Depression scales were used as indicators of a latent Depression factor. The PANAS-N, EPI-N, and PANAS-P were used as single indicators of negative affect, neuroticism, and positive affect, respectively. For these single indicators, error variances were computed from the sample reliability (α) and variance estimates, and these values were used to constrain the corresponding error matrices in the analysis. Correlated error was specified between DASS-Anxiety and DASS-Depression (method covariance of subscales from the same questionnaire).

On the basis of the notion that the BIS/BAS measures stable personality traits, it was expected that the BIS/BAS factors would evidence stronger relationships to measures of neighboring personality dimensions than to measures of current (past week) anxiety and depression symptoms. Accordingly, we predicted that the BIS factor would manifest stronger correlations with negative affect and neuroticism, than with the Anxious Arousal and Depression factors. We also hypothesized that the BAS scales would have a stronger relationship with positive affect than with the Anxious Arousal and Depression factors. Given evidence of the differential relationship between positive affect/behavioral activation and mood disorder symptoms (e.g., Brown et al., 1998), it was

[3] 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.

Table 4
Higher Order Factor Loadings and Residualized Primary Loadings for the 12-Item Behavioral Activation Scales (BAS; N = 925)

BAS item	Higher order factor loading	Residualized primary loading		
	General BAS	Reward	Drive	Fun Seeking
2	.538	.448		
8	.494	.412		
11	.389	.324		
14	.580	.312		
3	.542		.527	
6	.537		.521	
9	.553		.537	
12	.531		.516	
14	—		.200	
17	.407			.557
18	.416			.568
19	.466			.636
20	.368			.502

Note. Loadings were transformed with the Schmid–Leiman procedure. The dash indicates that the higher order factor loading of Item 14 presented with items loading on BAS-Reward.

predicted that the BAS factors would be more strongly related to Depression than Anxious Arousal.

Prior to interpretation of structural coefficients, the nine-dimension measurement model was examined. This model provided a good fit to the data, $\chi^2(216) = 524.29$, $p < .001$; RMSEA = .044, 90% CI = .044–.049; CFit = .98; SRMR = .044; CFI = .96. Fit diagnostics revealed no localized points of strain in the model, and factor loadings for the Anxious Arousal and Depression latent variables indicated that the anxiety and depression measures were reasonable indicators of these constructs (Anxious Arousal: factor loadings = .87 and .92 for BAI and DASS-Anxiety, respectively; Depression: factor loadings = .93 and .84 for BDI and DASS-Depression, respectively; Anxious Arousal–Depression factor intercorrelation = .57).

The correlations (ϕs) of the BIS/BAS factors with negative affect, neuroticism, positive affect, Anxious Arousal, and Depression are provided in Table 5. The differential magnitude of the correlations of each BIS/BAS factor with negative affect, neuroticism, positive affect, Anxious Arousal, and Depression was evaluated using the *z*-test procedure of Meng, Rosenthal, and Rubin (1992). As shown in Table 5, the BIS factor was most strongly correlated with neuroticism and negative affect, and the BAS scales were most strongly correlated with positive affect. The BIS factor manifested a stronger correlation with neuroticism than with negative affect, which may be explained in part by the fact that scores on the PANAS-N, an indicator of negative affect whose constituent items are single adjectives, are more apt to be influenced by current mood state than scores on the EPI-N, the indicator of neuroticism (cf. Widiger et al., 1999). Consistent with prediction, each BAS factor was more strongly related to Depression than Anxious Arousal.

Discussion

The current study extends the psychometric basis of the BIS/BAS by evaluating its factor structure, reliability, and validity in a large sample of outpatients with anxiety and mood disorders. The BIS/BAS factor structure was consistent across clinical replication samples, and was largely invariant in male and female patients. In accordance with results from nonclinical samples (e.g., Heubeck et al., 1998; Jorm et al., 1999), the BIS/BAS possessed good reliability in these clinical samples and evidenced strong convergent validity in relation to conceptually similar traits such as neuroticism and positive affect. Moreover, the use of cross-validated exploratory and confirmatory factor analytic procedures fostered the clarification of the BIS/BAS latent structure. Overall, a four-factor structure with factors corresponding to Carver and White's (1994) original Behavioral Inhibition, Reward Responsiveness, Drive, and Fun Seeking factors was supported. However, these analyses indicated that three BIS/BAS items had nonsalient loadings on their primary factors, suggesting an abridged, 17-item version of the BIS/BAS for research with clinical samples. Moreover, CFA analyses suggested that an item (Item 14) from the BAS-Reward subscale had a secondary loading on BAS-Drive. Whereas this cross-loading produced a localized area of strain in the CFA solutions, it is important to note that the magnitude of this loading was relatively weak (range = .27–.29), a relationship that was further diminished in the higher order factor transformations (loading = .20). Thus, this particular result does not appear to have

Table 5
Differential Relationships of the Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS) Factors With Negative Affect, Positive Affect, and Neuroticism, Anxious Arousal, and Depression (N = 749)

BIS/BAS factor	Validity dimension				
	Neuroticism	Negative affect	Positive affect	Depression	Anxious arousal
Behavioral Inhibition	$.75_a$	$.61_b$	$-.34_d$	$.48_c$	$.26_d$
Reward	$.06_c$	$-.04_c$	$.36_a$	$-.16_b$	$-.05_c$
Drive	$-.06_{c,d}$	$-.07_c$	$.43_a$	$-.19_b$	$-.01_d$
Fun Seeking	$-.08_c$	$-.16_b$	$.33_a$	$-.15_b$	$-.06_c$

Note. Coefficients are completely standardized parameter estimates (ϕ) from the measurement model. Parameters in the same row but with different subscripts differ significantly in magnitude ($p < .05$); ϕs $\geq |.15|$ are significant at $p < .001$.

strong implications for revisions to the composite scoring of the individual BAS subscales.

However, the analysis of the higher order structure of the BAS scales supports the conceptual position that the BAS factors represent more differentiated components of the broader dimension of BA. Although some investigators have suggested that the BAS subscales be treated separately and not combined into a general BAS composite (e.g., Ross et al., 2002), the present findings indicated that each BAS item had a salient loading on the higher order BAS dimension, in addition to its primary factor. Indeed, for the items comprising the Reward Responsiveness and Drive factors, loadings pertaining to the higher order factor were stronger than those pertaining to the lower order factors. Moreover, the scale reliability estimate of the higher order BAS dimension was quite favorable ($\rho = .88$) and in fact was higher than estimates obtained for the lower order BAS dimensions (range = .73–.82). These results suggest that although it is appropriate to use the BAS subscale scores in clinical research, a total BAS score may also be used as an indicator of the broader BA domain (cf. Gray, 1987).

Concurrent validity analyses were consistent with the conceptualization that the BIS/BAS factors are more reflective of temperamental traits than transient mood states. The BIS dimension was related most strongly to neuroticism and negative affect, and correlated more highly with these conceptually similar personality traits than to positive affect, anxiety, and depression. As has been found in previous studies (e.g., Carver & White, 1994; Jorm et al., 1999), BIS scores were significantly correlated with anxious and depressive symptoms; however, the current study established that these relationships are weaker than the relationship of BIS to more stable personality traits. The significant negative relationship of BIS to positive affect diverges from the lack of relationship generally found in nonclinical samples (Carver & White, 1994; Heubeck et al., 1998) and may be due to response tendencies that are particular to a clinical or treatment-seeking sample. For example, a respondent experiencing current emotional distress may endorse high levels of BIS sensitivity because the BIS items inquire about emotions such as anxiety and fear. In addition, the same respondent might be less likely to endorse positive affect—even on a trait measure—because positive emotions are discrepant with his or her dominant emotional experience at the moment. Indeed, questionnaires such as the BIS/BAS and dispositional form of the PANAS require respondents to reflect upon their history of emotional experience, and the effects of mood state on memory are well documented (e.g., D. M. Clark & Teasdale, 1982; Teasdale, Taylor, & Fogarty, 1980). Clinical participants' responses may be influenced by mood-congruent memory effects that lead to overendorsement of items that refer to negative affect and underendorsement of items that pertain to positive affect.

BAS scores correlated highly with positive affect and demonstrated substantially weaker (and generally nonsignificant) relationships to neuroticism, negative affect, anxiety, and depression. Each BAS subscale also demonstrated a small but statistically significant correlation with depressive symptoms, but not with anxiety symptoms. This finding may reflect that an underactive BAS constitutes a specific temperamental vulnerability to depression. Indeed, the symptoms that distinguish depression from anxiety (e.g., anhedonia, diminished motivation) are intricately related to theoretical conceptualizations of the BAS and specifically addressed in the Reward Responsiveness, Drive, and Fun Seeking scale items.

Overall, the differential relationships of the BIS/BAS factors to relevant personality traits and measures of anxious and depressive states is impressive, given that personality measures given at the time of an intake clinical assessment may be influenced by patients' current psychopathology (Widiger et al., 1999). In a particularly compelling demonstration of the distinctiveness of the BIS/BAS dimensions from clinical emotional disorder symptoms, Kasch et al. (2002) found each of the BIS/BAS scales to be highly stable over an 8-month period in 41 depressed participants and 21 nondepressed controls. In fact, although over a third of depressed participants were classified as no longer depressed at the 8-month follow-up, BIS/BAS scores displayed the same high level of temporal stability in this group (e.g., Time 1 and Time 2 BIS *M*s = 24.2 and 23.3, respectively) as in the group of participants who were depressed at both assessment points (e.g., Time 1 and Time 2 BIS *M*s = 24.0 and 23.9, respectively). As noted earlier, however, Time 1 BAS scores (Reward Responsiveness, Drive) were found to be significantly associated with the persistence of depression over the 8-month interval. This finding also contributes to evidence that BAS hypoactivity may constitute a specific vulnerability feature for depressive disorders.

Results of the current study also generally supported Gray's (1987) and Carver and White's (1994) conceptualization of BIS and BAS as relatively orthogonal dimensions. Whereas the correlations among the BAS subscales were moderate to strong (.42 to .55 for the final CFA model), the correlations of the BAS subscales with the BIS scale were fairly weak ($r = -.03$ to .27 for the final CFA model). The most substantial relationship was a positive correlation between BIS and Reward, a finding that converges with the results obtained in nonclinical samples (*r*s = .27–.28; Carver & White, 1994; Heubeck et al., 1998; Jorm et al., 1999). The positive relationship between BIS and Reward may be due to the considerations noted above, in which items from the BIS and Reward scales assess affective responding, whereas items from the Drive and Fun Seeking scales focus on behavior. Individuals with higher trait emotionality (i.e., greater affective intensity across the emotional spectrum) may score higher on both the BIS and Reward scales. In addition, individuals with greater insight into their emotional reactions may endorse greater BIS and Reward sensitivity, because they are more apt to be aware of their feelings of anxiety or excitement.

Despite the relatively large sample sizes used in this study, one limitation is the predominance of Caucasian participants. Future studies should attempt to test the psychometric properties of the BIS/BAS scales in more diverse samples, including tests of measurement invariance across racial/ethnic groups. Additionally, the clinical samples used for these analyses were comprised of individuals who presented to an anxiety and mood disorders clinic for assessment and treatment. Although many participants had lifetime histories of other psychological disorders (e.g., substance use disorders, somatization disorders), the present results may not generalize to other types of clinical samples (e.g., inpatient samples, samples comprised of patients with personality disorders, psychotic disorders, etc.). Notwithstanding these limitations to generalizability, the collective results provide a strong psychometric basis for the use of the BIS/BAS in clinical samples of patients with anxiety and mood disorders (clinical disorders that are very germane to the underlying theory of the BIS/BAS; cf. Gray, 1987). Given the theoretical importance of the behavioral inhibition and activation systems to models of psychopathology, the BIS/BAS holds considerable promise for future studies attempting to elucidate the roles of these systems in the development, manifestation, and course of the emotional disorders.

References

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

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

Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck Depression Inventory.* San Antonio, TX: Psychological Corporation.

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

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

Brody, L. R. (1999). *Gender, emotion, and the family.* Cambridge, MA: Harvard University Press.

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.

Brown, T. A., White, K. S., & Barlow, D. H. (in press). A psychometric reanalysis of the Albany Panic and Phobia Questionnaire. *Behaviour Research and Therapy.*

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.

Campbell, L. A., Brown, T. A., & Grisham, J. R. (2003). The relevance of age of onset to the psychopathology of generalized anxiety disorder. *Behavior Therapy, 34,* 31–48.

Carver, C. L., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67,* 319–333.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness of fit indices for testing measurement invariance. *Structural Equation Modeling, 2,* 233–255.

Clark, D. M., & Teasdale, J. D. (1982). Diurnal variation in clinical depression and accessibility of memories of positive and negative experiences. *Journal of Abnormal Psychology, 91,* 87–95.

Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology, 103,* 103–116.

Davidson, R. J. (1992). Emotion and affective style: Hemispheric substrates. *Psychological Science, 3,* 39–43.

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.

Eysenck, H. J., & Eysenck, S. B. G. (1968). *Manual for the Eysenck Personality Inventory.* San Diego, CA: Educational and Industrial Testing Service.

Gomez, A., & Gomez, R. (2002). Personality traits of the behavioral approach and inhibition systems: Associations with processing of emotional stimuli. *Personality and Individual Differences, 32,* 1299–1316.

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

Gray, J. A. (1987). *The psychology of fear and stress.* Cambridge, England: Cambridge University Press.

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

Harmon-Jones, E., & Allen, J. J. B. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *Journal of Abnormal Psychology, 106,* 159–163.

Heubeck, B. G., Wilkinson, R. B., & Cologon, J. (1998). A second look at Carver and White's (1994) BIS/BAS scales. *Personality and Individual Differences, 25,* 785–800.

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. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood and least square methods.* Mooresville, IN: Scientific Software.

Jöreskog, K., & Sörbom, D. (2001). LISREL 8.52 [Computer software]. Chicago: Scientific Software International.

Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Rodgers, B. (1999). Using the BIS/BAS scales to measure behavioural inhibition and behavioural activation: Factor structure, validity, and norms in a large community sample. *Personality and Individual Differences, 26,* 49–58.

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *Journal of Abnormal Psychology, 111,* 589–597.

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

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33,* 335–342.

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

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

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

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

Meyer, B., Johnson, S. L., & Carver, C. S. (1999). Exploring behavioral activation and inhibition sensitivities among college students at risk for bipolar spectrum symptomatology. *Journal of Psychopathology and Behavioral Assessment, 21,* 275–292.

Meyer, B., Johnson, S. L., & Winters, R. (2001). Responsiveness to threat and incentive in bipolar disorder: Relations of the BIS/BAS scales with symptoms. *Journal of Psychopathology and Behavioral Assessment, 23,* 133–143.

Muthén, L. K., & Muthén, B. O. (1998). Mplus 2.12 [Computer software]. Los Angeles: Author.

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.

Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling, 9,* 195–212.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114,* 552–566.

Ross, S. R., Millis, S. R., Bonebright, T. L., & Bailley, S. E. (2002). Confirmatory factor analysis of the Behavioral Inhibition and Activation Scales. *Personality and Individual Differences, 33,* 861–865.

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53–61.

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

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science, 8,* 204–210.

Teasdale, J. D., Taylor, R., & Fogarty, S. J. (1980). Effects of induced elation-depression on the accessibility of memories of happy and unhappy experiences. *Behaviour Research and Therapy, 18,* 339–346.

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.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Widiger, T. A., Verheul, R., & van den Brink, W. (1999). Personality and psychopathology. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 347–366). New York: Guilford Press.

Received March 5, 2003
Revision received February 9, 2004
Accepted March 4, 2004 ■