Journal of Abnormal Psychology
2007, Vol. 116, No. 2, 313–328

0021-843X/07/$12.00 DOI: 10.1037/0021-843X.116.2.313

Temporal Course and Structural Relationships Among Dimensions of Temperament and *DSM–IV* Anxiety and Mood Disorder Constructs

Timothy A. Brown
Boston University

The temporal stability and directional relations among dimensions of temperament (e.g., neuroticism) and selected *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV;* American Psychiatric Association, 1994) disorder constructs (depression, generalized anxiety disorder, social phobia) were examined in 606 outpatients with anxiety and mood disorders, assessed on 3 occasions over a 2-year period. Neuroticism/behavioral inhibition (N/BI) and behavioral activation/positive affect (BA/P) accounted for the cross-sectional covariance of the *DSM–IV* constructs. Although N/BI evidenced the most change of the constructs examined, initial levels of N/BI predicted less improvement in 2 of the 3 disorder constructs. Unlike the *DSM–IV* disorder constructs, the temporal stability of N/BI increased as a function of initial severity. Moreover, N/BI explained all the temporal covariation of the *DSM–IV* disorder constructs. The results are discussed in regard to conceptual models of temperament that define N/BI and BA/P as higher order dimensions accounting for the course and covariation of emotional disorder psychopathology.

Keywords: temperament and psychopathology of emotional disorders, predictors of longitudinal course of anxiety and mood disorders, classification of emotional disorders, vulnerability for anxiety and depression, structural equation modeling

A wealth of evidence at the genetic, diagnostic, and symptom level attests to the high degree of covariance among the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV;* American Psychiatric Association, 1994) anxiety and mood disorders (e.g., T. A. Brown & Barlow, 2002; Hettema, Prescott, Myers, Neale, & Kendler, 2005; Mineka, Watson, & Clark, 1998). Although these findings might be interpreted as suggesting that the *DSM–IV* anxiety and mood disorder constructs possess inadequate discriminant validity (cf. Andrews, 1996), recent evidence and theory indicate much of this overlap stems from the fact that these disorders emerge from the same biologic/genetic and psychosocial diatheses (e.g., Barlow, 2002; Clark, Watson, & Mineka, 1994). That is, although emotional disorders co-occur in part because they arise from core diatheses, this common vulnerability is manifested heterogeneously (i.e., as different *DSM* disorders) as a result of exposure to differing environmental influences, other genetic/biologic agents, and so forth (e.g., Barlow, 2002; Clark, 2005; Hettema et al., 2005). For example, whereas a genetically based trait such as neuroticism may underlie social phobia and depression, whether one or both conditions become manifest may depend on environmental determinants such as direct experiences with social humiliation, rejection, or vicarious exposure (e.g., parental modeling of ineffective coping). Many researchers would concur that, although the overlap among anxiety and mood disorders is considerable, differentiation is useful because it has important implications for understanding clinical course, complications, and treatment.

Two genetically based core dimensions of temperament have been posited to be instrumental in the etiology and course of the emotional disorders: neuroticism/negative affectivity and extraversion/positive affectivity. Whereas neuroticism/negative affectivity is considered to be etiologically relevant to the full range of emotional disorders, the influence of extraversion/positive affectivity is more specific to depression and social phobia (e.g., T. A. Brown, Chorpita, & Barlow, 1998; Mineka et al., 1998; Watson, Clark, & Carey, 1988). Although the theoretical frameworks were developed independently (cf. Eysenck, 1981; Tellegen, 1985), neuroticism/negative affectivity and extraversion/positive affectivity are closely related to Gray's (1987) constructs of behavioral inhibition and behavioral activation, respectively, at both the conceptual and empirical levels (e.g., Campbell-Sills, Liverant, & Brown, 2004; Carver & White, 1994; Clark et al., 1994; Kasch, Rottenberg, Arnow, & Gotlib, 2002).

A growing literature attests to the heritability of these constructs (e.g., Fanous, Gardner, Prescott, Cancro, & Kendler, 2002; Hettema, Prescott, & Kendler, 2004; Viken, Rose, Kaprio, & Koskenvuo, 1994) and their roles in accounting for the onset, overlap, and maintenance of anxiety and depression (e.g., T. A. Brown et al., 1998; Gershuny & Sher, 1998; Kasch et al., 2002; Watson, Clark, & Tellegen, 1988). For instance, in a sample of outpatients, T. A. Brown et al. (1998) found that virtually all the considerable covariance among latent variables corresponding to the *DSM–IV* constructs of unipolar depression, social phobia, generalized anxiety disorder, obsessive–compulsive disorder, and panic disorder/agoraphobia was explained by the higher order dimensions of

I thank David Barlow for his comments on this article and the following individuals for their assistance with this study: Gabrielle Liverant, Dylan Kollman, Amy Lawrence, Jessica Grisham, and Laura Campbell-Sills.

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 02215-2013. E-mail: tabrown@bu.edu

negative affect and positive affect. Although the results were consistent with the notion of neuroticism/negative affect as a broadly relevant dimension of vulnerability, the *DSM–IV* disorder constructs were differentially related to negative affect, with depression and generalized anxiety disorder evidencing the strongest associations (e.g., standardized γs = .67 and .74, respectively). In accord with a reformulated hierarchical model of anxiety and depression (Mineka et al., 1998), positive affect was predictive of depression and social phobia only (see also Watson, Clark, & Carey, 1988). Recent comorbidity and structural genetic findings also reveal a differential relationship between mood disorders and social phobia, which could be seen as consistent with the existence of a temperamental vulnerability dimension specific to these two disorders (T. A. Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Eley & Brown, 2006).

Nonetheless, empirical tests of the specific nature of the relationship between temperament and psychopathology have been limited by many factors. An overarching complication is that there are several manners in which these constructs may be interrelated: (a) *predispositional*—dimensions of temperament act as vulnerability to emotional disorders, (b) *pathoplastic*—temperament primarily influences the course and expression of disorders, (c) *complication/scar*—changes in temperament are due to the experience of emotional disorders, and (d) *continuity spectrum*—dimensions of temperament and disorders reflect the same underlying process at varying degrees of severity (e.g., Clark et al., 1994; Widiger, Verheul, & van den Brink, 1999). Further differentiation of these explanations is possible (cf. Clark, 2005). Although leading conceptual models of emotional disorders espouse predispositional and pathoplastic explanations (e.g., Barlow, 2002; Clark, 2005), study design limitations have precluded a compelling evaluation of the directionality of the relationships between temperament and psychopathology. Common methodological limitations have involved sampling (e.g., use of college student samples), measurement (e.g., use of adjective checklists and instructional sets that foster mood-state distortions on personality measures; see next paragraph), scope (e.g., consideration of only one or two personality and *DSM* disorder constructs at a time, limiting the evaluation of the specificity of relationships and theoretically viable competing hypotheses), and perhaps most saliently, the predominant use of cross-sectional designs (Widiger & Seidlitz, 2002; Widiger et al., 1999). More prospective studies are needed to address the viability of these varying, but not mutually exclusive, accounts.

However, longitudinal studies of clinical and nonclinical samples each pose unique limitations and challenges. For instance, in clinical samples, predispositional and scar hypotheses cannot be addressed (i.e., because disorders are already present). In addition, clinical studies must recognize and deal with the effects of *mood-state distortion*; that is, patients' self-reports of temperament are affected (augmented) by their current clinical state. Accordingly, even if temperament constructs truly operate as temporally stable dimensions that characterize everyday functioning (i.e., traits), their measurement in clinical samples will contain a considerable amount of "state" variance that is due to, and thus covaries with, general clinical distress (e.g., Clark, Vittengl, Kraft, & Jarrett, 2003; Widiger et al., 1999). These issues may be less problematic in nonclinical samples. Yet, nonclinical studies have restrictive methodological requirements (e.g., very large samples and protracted lengths of follow-up) to ensure adequate statistical power and a sufficient incidence of target events (e.g., disorder onsets). Indeed, given the requisite methodological demands (e.g., following individuals from infancy through adulthood), it is likely that no single study will unambiguously extricate the viability of these varying accounts (Widiger & Seidlitz, 2002). Despite these issues, studies of the directional relationships among temperament and psychopathology constructs can inform etiological models and the classification and treatment of mental disorders in many ways. These include (a) identification of dimensions that transcend the features of disorders themselves that account for the overlap among disorders and may have a stronger genetic basis; (b) verification of the nature and extent to which putative risk dimensions predict the course, complications, and treatment response of disorders; and (c) validation and refinement of higher order phenotypes of temperament that represent more robust constructs for biological analysis (e.g., genetic studies), may merit their own treatment focus (e.g., may not respond to disorder-specific treatment and leave patients at risk for return of comorbidity; cf. T. A. Brown, Antony, & Barlow, 1995; Kasch et al., 2002), and may warrant inclusion in a future *DSM* classification system that incorporates dimensional elements (cf. T. A. Brown & Barlow, 2005; Clark, 2005; Widiger & Samuel, 2005).

Although the literature on the temporal relationships between temperament and emotional disorder psychopathology is sparse, several studies of this nature have been conducted in the context of treatment outcome research for major depression. These studies were based on the premise that because neuroticism/negative affect and extraversion/positive affect are construed as risk factors for depression, these dimensions should predict depression treatment outcome (i.e., higher neuroticism/negative affect and lower extraversion/positive affect are associated with poorer treatment response). However, results have been mixed, with some studies showing that pretreatment levels of temperament predict poorer treatment outcome (e.g., Geerts & Bouhuys, 1998; Joyce, Mulder, & Cloninger, 1994; Kasch et al., 2002), and other studies reporting no such effects (e.g., Boyce & Parker, 1985; Clark et al., 2003; Sato et al., 1999).

Given their conceptualization as trait vulnerability dimensions, it might be expected that neuroticism/negative affect and extraversion/positive affect are temporally stable in clinical samples. However, most of the aforementioned depression treatment studies have found that scores on measures of temperament change significantly from pre- to posttreatment (i.e., improvement in depression is accompanied by a reduction in neuroticism and an increase in extraversion). Although several interpretations are possible (e.g., trait vulnerability is responsive to treatment), it seems likely that the pre- to posttreatment changes in temperament observed in these studies were due in large part to a mood-state effect—in addition to a stable (trait) component, the measurement of temperament also taps "state" variance (clinical distress), which is less stable and covaries with disorder severity. Clark et al. (2003) addressed this explanation by using multivariate statistical procedures to separate the variance of their measure of temperament into trait and state components. Indeed, the authors found that changes in depression severity correlated with state changes in temperament measurement, but not with stable trait scores.

The findings of Kasch et al. (2002) are a notable exception in the depression treatment outcome literature. These authors examined the temporal stability (8 months) and predictive utility of self-

reported levels of behavioral inhibition and behavioral activation in 41 individuals with major depression (most of whom received treatment during the follow-up interval). Lower Time 1 behavioral activation predicted poorer clinical outcome of depression at the 8-month reassessment. Moreover, the self-report measures of behavioral inhibition and activation (i.e., the Behavioral Inhibition/Behavioral Activation Scales [BIS/BAS]; Carver & White, 1994) were remarkably stable over time and clinical state. In fact, although over a third of 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 a subgroup 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). Thus, unlike other studies using depression treatment samples, these results suggest that dimensions of temperament are stable, do not covary with clinical state, and are relatively unaffected by treatment. These findings also highlight that study differences in the measurement of temperament may be partly responsible for the inconsistent findings on the stability and predictive validity of these dimensions (e.g., a measure such as the BIS/BAS may be more temporally stable than an adjective checklist).

With these issues in mind, the present study examined the temporal course and structural relationships among dimensions of temperament and selected *DSM–IV* anxiety and mood disorder constructs (unipolar depression, social phobia, generalized anxiety disorder) in a large sample of treatment-seeking outpatients followed over a period of 2 years. Selection of the three *DSM–IV* disorder constructs was guided by findings indicating potential diagnostic boundary problems with one another (e.g., generalized anxiety disorder and mood disorders) or with dimensions of temperament (e.g., generalized anxiety disorder and neuroticism), as well as the aforementioned evidence indicating a differential relationship of positive affect with mood disorders and social phobia (cf. T. A. Brown, Barlow, & Liebowitz, 1994; T. A. Brown et al., 1998; T. A. Brown, Di Nardo, Lehman, & Campbell, 2001; Mineka et al., 1998). Although results have been mixed in the extant literature (cf. Clark et al., 2003; Kasch et al., 2002), the key hypotheses were forwarded in accord with leading theoretical models of trait vulnerability in the emotional disorders (which espouse a predispositional explanation of the nature of the relationship between temperament and psychopathology; e.g., Barlow, 2002): (a) Although a decrease in severity will be seen in all constructs, dimensions of temperament will evidence greater temporal stability and less response to treatment, relative to the *DSM–IV* disorder constructs; and (b) directional effects will be obtained such that initial levels of temperament will predict the extent of change in the *DSM–IV* disorder constructs in the expected manner (i.e., high negative affectivity predicts lesser change in all disorder constructs, low positive affectivity predicts less change in depression and social phobia only), but not vice versa (i.e., if the initial severity of the *DSM–IV* disorder constructs predicted the temporal course of dimensions of temperament, this would be at odds with conceptual models of trait vulnerability that posit unidirectional temperament → psychopathology effects). Finally, along the lines of T. A. Brown et al. (1998), it was predicted that dimensions of temperament will act as higher order variables accounting for the cross-sectional and temporal covariance of the *DSM–IV* disorder constructs. The time-series aspect of this prediction was based on the aforementioned evidence of a "state" variance component of temperament that is expected to covary with clinical distress (cf. Clark et al., 2003).

The hypotheses were tested with single- and parallel-process latent growth models (LGMs) using latent variables of the temperament and disorder constructs as outcomes. Although it increases analytic complexity, the use of outcomes defined by multiple indicators has many advantages. For instance, it responds to the call for multivariate research on temperament and psychopathology that emphasizes constructs over individual measures (Clark et al., 1994; Kasch et al., 2002). Moreover, growth is analyzed within the framework of a latent variable measurement model such that the outcomes are theoretically free of measurement error and estimated in the context of a formal evaluation of longitudinal measurement invariance (e.g., time-specific variance and measurement error variance are not confounded, statistical power is fostered by smaller standard errors of growth factor parameters; cf. Curran & Bollen, 2001; Hancock, Kuo, & Lawrence, 2001). Incorporation of latent variable measurement models in parallel-process LGMs (i.e., simultaneous estimation of more than one growth process) addresses the discriminant validity of the purportedly distinct constructs of temperament and emotional disorder.

Method

Participants

The sample consisted of 606 outpatients who presented for assessment or treatment at the Center for Anxiety and Related Disorders (CARD). Women constituted the larger portion of the sample (63%); average age was 34.72 years (*SD* = 11.89, range = 18 to 74). The sample was predominantly Caucasian (89%; African American = 4%, Asian = 3%, Latino/Hispanic = 3%). Intake diagnoses (Time 1 [T1]) 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 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). Patients were reevaluated at 12 months (Time 2 [T2]) and 24 months (Time 3 [T3]) with the follow-up version of the ADIS–IV (ADIS–IV–FU), which is identical to the ADIS–IV–L except (a) sections for past diagnoses are omitted, and (b) a section is included to assess treatment follow-up (e.g., nature and extent of treatments received since intake). Both ADIS–IV versions provide dimensional assessment of the key and associated features of disorders (0–8 ratings); 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), which had two independent administrations of the ADIS–IV–L, indicated good to excellent interrater agreement for current disorders (range of κs = .67 to .86), except for dysthymia (κ = .31; T. A. Brown, Di Nardo, et al., 2001). As assessed by the ADIS–IV–L, the rates of current clinical disorders that frequently occurred in the sample at intake were as follows: social phobia (42%), panic disorder with or without agoraphobia (38%), mood disorders (i.e., major depression, dysthymic disorder, depressive disorder not otherwise specified; 36%), generalized anxiety disor-

der (22%), specific phobia (20%), and obsessive–compulsive disorder (12%). The rate of generalized anxiety disorder at intake increases to 34% when the *DSM–IV* hierarchy rule with mood disorders is ignored (cf. T. A. Brown, Campbell, et al., 2001).

Latent Variables and Indicators in the Structural Models

In addition to the temperament dimensions of neuroticism/behavioral inhibition (N/BI) and behavioral activation/positive affect (BA/P),[1] three *DSM–IV* disorder constructs were examined in the structural models: depression (DEP), social phobia (SOC), and generalized anxiety disorder (GAD).

N/BI. The following measures were used as indicators of the latent construct of N/BI: (a) Eysenck Personality Inventory—Neuroticism (EPI–N; Eysenck & Eysenck, 1968); (b) the BIS of the BIS/BAS (Carver & White, 1994);[2] and (c) Negative Affect Scale of the Positive and Negative Affect Scales (PANAS–N, "in general" time frame instructions; Watson, Clark, & Tellegen, 1988).

BA/P. The BA/P latent variable was defined by the BAS of the BIS/BAS and the Positive Affect Scale of the PANAS (PANAS–P). Psychometric studies of the BIS/BAS have found that behavioral activation is composed of three subdimensions (reward, drive, fun seeking; e.g., Campbell-Sills et al., 2004; Carver & White, 1994). However, the broader dimension of behavioral activation was used in the present analyses on the basis of recent findings of the existence of this higher order factor, which evidences more favorable psychometric properties (e.g., $\rho = .88$) than its constituent subdimensions (Campbell-Sills et al., 2004).

SOC. Two questionnaires and a clinical rating measure were used as indicators of the SOC latent variable. The questionnaires were (a) the Social Interaction Anxiety Scale (Mattick & Clarke, 1998; cf. E. J. Brown et al., 1997) and (b) the Social Phobia Scale of the Albany Panic and Phobia Questionnaire (Rapee, Craske, & Barlow, 1994/1995; cf. T. A. Brown, White, & Barlow, 2005). In addition, patients' fear of 13 social situations (e.g., initiating a conversation, participating at meetings and/or classes) was rated by the clinician during administration of the ADIS–IV–L and ADIS–IV–FU (0 = *no fear* to 8 = *very severe fear*). Analyses reported in T. A. Brown, Di Nardo, et al. (2001), using a subset of the current sample, indicate these ratings are unifactorial and have favorable interrater reliability (test–retest $r = .86$ on the basis of independent administrations of the ADIS–IV–L). A sum of these ratings was used in the analyses (α in present sample = .91).

DEP. A latent variable of unipolar depression was formed with the use of two questionnaire indicators and an ADIS–IV clinical rating composite: (a) Beck Depression Inventory (BDI; Beck & Steer, 1987), (b) Depression Scale of the 42-item version of the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995; cf. Antony, Bieling, Cox, Enns, & Swinson, 1998; T. A. Brown, Chorpita, Korotitsch, & Barlow, 1997), and (c) the ADIS–IV dimensional ratings of the symptom criteria of *DSM–IV* major depression (0 = *none* to 8 = *very severe*; interrater $r = .74$; T. A. Brown, Di Nardo, et al., 2001). In accord with the procedures of T. A. Brown et al. (1998), the BDI was scored using the 10 items that load on a Cognitive/Affective factor (Items 1–9, 13) that are more specific to the unipolar mood disorders ($\alpha = .87$). Similarly, only the symptoms of depressed mood, anhedonia/loss of interest, worthlessness/guilt, and thoughts of death or suicide were used to form the sum composite of the ADIS–IV dimensional ratings of depression ($\alpha = .83$).

GAD. A *DSM–IV* factor of GAD was created using three ADIS–IV dimensional rating measures (all 0–8 scales): (a) excessiveness of worry in eight areas (e.g., finances, minor matters; $\alpha = .82$), (b) a single rating of difficulty controlling worry, and (c) frequency and severity of the six symptoms comprising the *DSM–IV* associated symptom criterion of GAD ($\alpha = .83$).[3] Each of these indicators evidences favorable interrater reliability (e.g., test–retest *r*s = .73, .73, and .83, respectively; cf. T. A. Brown, Di Nardo, et al., 2001). Internal consistencies (α) of these composites in the present sample were .82 or higher.

Data analysis. The raw data were analyzed using a latent variable software program and maximum-likelihood minimization functions (Mplus 3.13; Muthén & Muthén, 1998–2005). Although attrition was negligible at T2 (7%), it increased to 50% at T3. Thus, missing data were accommodated in all models using direct maximum likelihood under the assumption of missingness at random (cf. Allison, 2003; Raykov, 2005). Goodness of fit of the models was evaluated using the root-mean-square error of approximation (RMSEA) and its 90% confidence interval (CI) and test of close fit (CFit), the Tucker–Lewis index (TLI), the comparative fit index (CFI), and the standardized root-mean-square residual (SRMR). Acceptable model fit was defined in part by the criteria forwarded by Hu and Bentler (1999): RMSEA values close to .06 or below (90% CI upper limit close to $\leq .06$, nonsignificant CFit), CFI and TLI values close to .95 or above, and SRMR values close to .08 or below. 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 model fit. In instances in which competing models were nested, comparative fit was evaluated with nested chi-square tests (χ^2_{diff}). The acceptability of the models was further evaluated by the presence or absence of salient localized areas of strains in the solutions (e.g., modification indices) and the strength and interpretability of the parameter estimates.

---

[1] These latent variables were so named to make the marker indicator more apparent (i.e., N/BI = EPI–N is marker, BA/P = BAS is marker) to foster interpretability of unstandardized estimates (marker indicator = indicator whose unstandardized factor loading is fixed to 1.0 in order to define the metric of the latent variable).

[2] The BIS was reversed scored for inclusion in study analyses to be consistent with the scaling of the EPI–N and PANAS–N (i.e., higher scores reflect higher levels of neuroticism, behavioral inhibition, and negative affect).

[3] The original study design used the Penn State Worry Questionnaire (PSWQ; J. T. Meyer, Miller, Metzger, & Borkovec, 1990) as an indicator of GAD. However, initial analyses revealed that the PSWQ evidenced poor discriminant validity in the models that included the N/BI latent variable (e.g., the PSWQ was more strongly correlated with indicators of N/BI than with indicators of GAD). To allow the temporal analyses to proceed, the PSWQ was from subsequent models.

Results

Diagnostic Outcome Over Follow-Up

Three quarters (76%) of the sample received treatment at CARD following the intake assessment. As expected, the overall rate of anxiety and mood disorders in the sample declined markedly at the 12-month follow-up (i.e., from 100% to 64%; McNemar test $p < .001$), with negligible additional change at the 24-month assessment (58%, $p = .80$). A similar pattern was found for the specific *DSM–IV* disorders examined in this study. The respective rates of each disorder at intake, 12 months, and 24 months were social phobia, 42%, 28%, and 26%; generalized anxiety disorder (excluding hierarchy), 34%, 22%, and 20%; and mood disorders (major depression, dysthymia, depression not otherwise specified), 36%, 20%, and 24%.

Longitudinal Measurement Models

To establish the suitability of using latent variables as outcomes in the LGMs (i.e., to ensure that temporal change in a latent construct is not confounded by change in its measurement across time), longitudinal measurement models were evaluated for each temperament and *DSM–IV* disorder construct. This is exemplified by the measurement model for DEP in Figure 1. As shown in Figure 1, indicators of the same variable assessed at different times (e.g., BDI_1, BDI_2, BDI_3) were specified to have correlated uniquenesses (cf. T. A. Brown, 2006; Kenny & Zautra, 2001). A baseline model, which contained no equality constraints on the factor loadings and indicator intercepts, fit the data well, $\chi^2(15) = 34.73$, $p = .003$, SRMR = .02, RMSEA = 0.05 (90% CI = 0.03–0.07, CFit =.58), TLI = 0.98, CFI = .99. The next model tested the temporal invariance of the factor loadings. These constraints produced a statistically nonsignificant increase in model chi-square, indicating that the loadings were time invariant, $\chi^2_{diff}(4) = 1.00$, *ns*. A final model confirmed that the indicator intercepts were equal over time, $\chi^2_{diff}(9) = 4.68$, *ns*. The completely standardized parameter estimates of this solution are presented in Figure 1.

The same modeling process was undertaken for the dimensions of SOC, GAD, N/BI, and BA/P. The descriptive goodness of fit of the final measurement model (equal loadings and intercepts) is presented in Table 1. In three instances, the longitudinal measurement models were found to possess partial measurement invariance (cf. Byrne, Shavelson, & Muthén, 1989); that is, most but not all measurement parameters were invariant over time. Growth modeling with latent factors requires at least one time-invariant indicator in addition to the marker indicator (Bollen & Curran, 2005; T. A. Brown, 2006; Byrne et al., 1989). As seen in Table 1, this condition was met, as there were two or more time-invariant

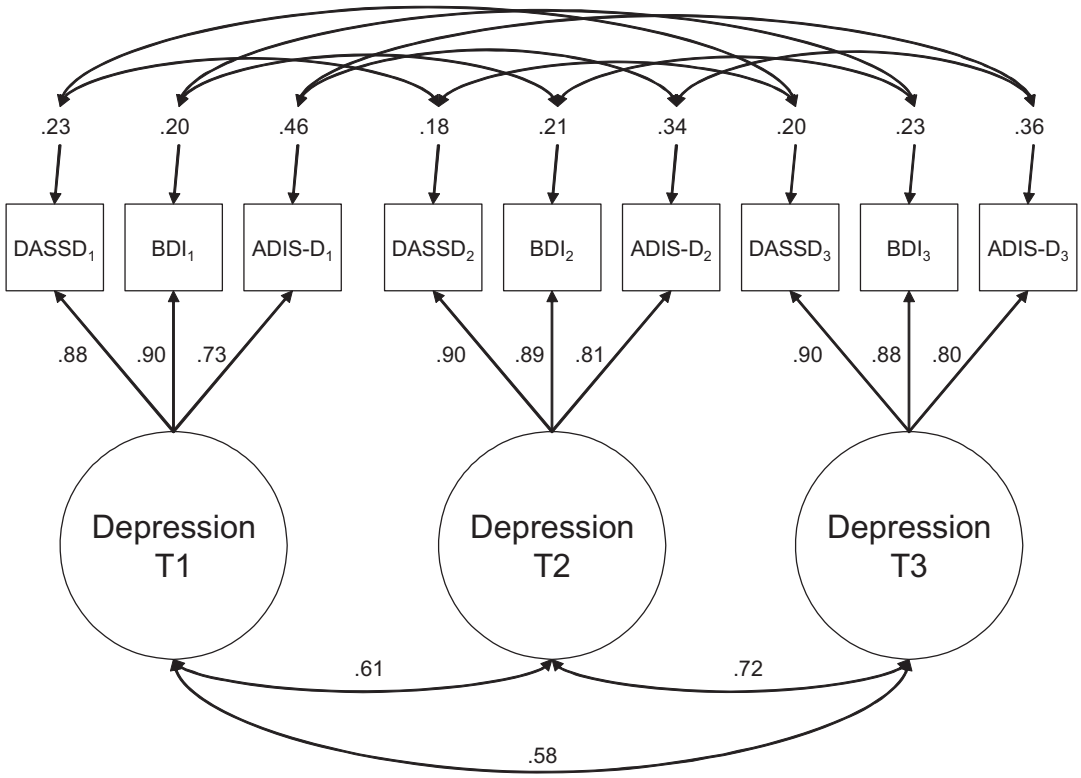


Figure 1. Longitudinal measurement model of depression. Completely standardized parameter estimates are provided. All corresponding unstandardized parameter estimates are statistically significant (*p*s < .001). For presentational clarity, correlated error estimates are not presented (range = .03 to .10). DASSD = Depression Scale of the Depression Anxiety Stress Scales; BDI = Beck Depression Inventory; ADIS–D = Anxiety Disorders Interview Schedule for *DSM–IV* dimensional ratings of major depression; T1 (Time 1) = intake; T2 (Time 2) = 12-month follow-up; T3 (Time 3) = 24-month follow-up.

Table 1
Longitudinal Measurement Models of Neuroticism/Behavioral Inhibition, Behavioral Activation/Positive Affect, Depression, Social Phobia, and Generalized Anxiety Disorder

Factor loading	T1	T2	T3	Model constraint
Neuroticism/behavioral inhibition[a]				
EPI–N	.85	.94	.94	All held equal except λ and τ of PANAS–N at T1
BIS	.66	.71	.70	
PANAS–N	.79	.79	.80	
Behavioral activation/positive affect[b]				
BAS	.85	.88	.83	All λs and τs held equal
PANAS–P	.51	.56	.51	
Depression[c]				
DASS–D	.88	.90	.90	All λs and τs held equal
BDI	.90	.89	.88	
ADIS–D	.73	.81	.80	
Social phobia[d]				
SIAS	.91	.89	.90	All held equal except τs of APPQ–S at T1–T3; λ of APPQ–S at T2
ADIS–S	.85	.86	.86	
APPQ–S	.90	.86	.88	
Generalized anxiety disorder[e]				
ADIS–Ex	.84	.86	.88	All held equal except τ of ADIS–As at T1
ADIS–Co	.89	.92	.92	
ADIS–As	.76	.74	.74	

Note. First measure listed for each latent variable used as marker indicator (e.g., EPI–N for Neuroticism/Behavioral Inhibition); λ = factor loading; τ = indicator intercept; T1 (Time 1) = intake; T2 (Time 2) = 12-month follow-up; T3 (Time 3) = 24-month follow-up; EPI–N = Neuroticism Scale of Eysenck Personality Inventory; BIS = Behavioral Inhibition Scale; PANAS–N = Negative Affect Scale of the Positive and Negative Affect Scales; BAS = Behavioral Activation Scale; PANAS–P = Positive Affect Scale of the Positive and Negative Affect Scales; DASS–D = Depression Scale of the Depression Anxiety Stress Scales; BDI = Beck Depression Inventory; ADIS–D = Anxiety Disorders Interview Schedule for *DSM–IV* (ADIS–IV) ratings of major depression; SIAS = Social Interaction Anxiety Scale; ADIS–S = ADIS–IV ratings of situational social fear; APPQ–S = Social Phobia Scale of Albany Panic and Phobia Questionnaire; ADIS–Ex = ADIS–IV ratings of excessiveness of worry; ADIS–Co = ADIS–IV ratings of uncontrollability of worry; ADIS–As = ADIS–IV ratings of associated symptoms of generalized anxiety disorder; SRMR = standardized root-mean-square residual; RMSEA = root-mean-square error of approximation; CI = confidence interval; CFit = test of close fit; TLI = Tucker–Lewis index; CFI = comparative fit index.
[a] Model fit: $\chi^2(21) = 21.50$, $p = .43$, SRMR = .04, RMSEA = 0.01 (90% CI = 0.00–0.03, CFit = 1.0), TLI = 1.00, CFI = 1.00.
[b] Model fit: $\chi^2(7) = 5.11$, $p = .65$, SRMR = .01, RMSEA = 0.00 (90% CI = 0.00–0.04, CFit = .98), TLI = 1.00, CFI = 1.00.
[c] Model fit: $\chi^2(23) = 39.40$, $p = .02$, SRMR = .03, RMSEA = 0.03 (90% CI = 0.01–0.05, CFit = .92), TLI = 0.99, CFI = .99.
[d] Model fit: $\chi^2(20) = 52.56$, $p < .001$, SRMR = .02, RMSEA = 0.05 (90% CI = 0.04–0.07, CFit = .40), TLI = 0.99, CFI = .99.
[e] Model fit: $\chi^2(22) = 23.44$, $p = .38$, SRMR = .02, RMSEA = 0.01 (90% CI = 0.00–0.04, CFit = 1.0), TLI = 1.00, CFI = 1.00.

indicators (for both factor loadings and intercepts) in each longitudinal measurement model.

Temporal Variability of N/BI, BA/P, and DSM–IV Disorder Latent Constructs

The prediction that the constructs of N/BI and BA/P would evidence greater temporal stability than *DSM–IV* disorder constructs was first evaluated by considering the test–retest correlations of these dimensions (i.e., parameter estimates from the longitudinal models described in Table 1). These estimates are provided in Table 2. As expected, the T1–T2 correlations were of lesser magnitude than the T2–T3 correlations because the majority of the sample received treatment before T2, and there was not an appreciable change in disorder status between T2 and T3. Although it is likely that many patients received additional treatment during the T2–T3 interval (not formally assessed), the T2–T3

Table 2
Temporal Variation in Neuroticism/Behavioral Inhibition, Behavioral Activation/Positive Affect, Depression, Social Phobia, and Generalized Anxiety Disorder

Parameter estimate	N/BI	BA/P	DEP	SOC	GAD
Test–retest correlations					
T1–T2	.72	.85	.61	.82	.66
T2–T3	.81	.94	.72	.94	.69
T1–T3	.70	.84	.58	.79	.58
Growth intercept[a]					
Variance	13.18***	11.57***	62.58***	294.71***	58.88***
Growth slope					
Mean	−3.34***	0.66***	−4.56***	−6.30***	−3.78***
Variance	4.53*	2.16, *ns*	22.62***	90.51***	15.63**
Effect size (*d*)	.70	.19	.52	.39	.43
T2 slope loading	.85	1.00	.94	.90	.92
Intercept–Slope					
Covariance	3.64**	0.12, *ns*	−15.80***	−78.49***	−12.80**
Correlation	.47	.02	−.42	−.48	−.42
Treatment covariate → Slope					
Unstandardized	−2.04***	n/a	−3.00***	−4.52***	−1.51*
Standardized	−0.98	n/a	−0.63	−0.47	−0.38

Note. Fit of latent growth models are as follows: neuroticism/behavioral inhibition (N/BI), $\chi^2(23) = 22.58$, $p = .49$, standardized root-mean-square residual (SRMR) = .03, root-mean-square error of approximation (RMSEA) = 0.00 (90% confidence interval [CI] = 0.00–0.03, test of close fit [CFit] = .99), Tucker–Lewis index (TLI) = 1.00, comparative fit index (CFI) = 1.00; behavioral activation/positive affect (BA/P), $\chi^2(10) = 7.91$, $p = .64$, SRMR = .04, RMSEA = 0.00 (90% CI = 0.00–0.04, CFit = .99), TLI = 1.00, CFI = 1.00; depression (DEP), $\chi^2(25) = 40.25$, $p = .03$, SRMR = .02, RMSEA = 0.03 (90% CI = 0.01–0.05, CFit = .96), TLI = 0.99, CFI = .99; social phobia (SOC), $\chi^2(22) = 53.02$, $p < .001$, SRMR = .02, RMSEA = 0.05 (90% CI = 0.03–0.06, CFit = .54), TLI = 0.99, CFI = .99; generalized anxiety disorder (GAD), $\chi^2(24) = 26.96$, $p = .31$, SRMR = .03, RMSEA = 0.01 (90% CI = 0.00–0.04, CFit = 1.00), TLI = 1.00, CFI = 1.00. Time 1 (T1) = intake; Time 2 (T2) = 12-month follow-up; Time 3 (T3) = 24-month follow-up; n/a = no significant variance in slope to be explained by treatment covariate.
[a] Mean of intercept fixed at zero.
* $p < .05$. ** $p < .01$. *** $p < .001$.

test–retest correlations might be roughly interpreted as estimates of the untreated course (temporal stability) of these constructs given the relative lack of diagnostic change over the 2nd year of follow-up.

The differential magnitude of the test–retest coefficients offered partial support for the prediction of greater temporal stability of N/BI and BA/P. As seen in Table 2, the temporal correlations of N/BI and BA/P were generally higher than those of DEP and GAD. BA/P evidenced the highest degree of temporal stability, with a 2-year test–retest correlation of .84 (T1–T3) and a 1-year correlation as high as .94 (T2–T3). However, the temporal stability of SOC was comparable to BA/P (e.g., T1–T3 $r = .79$). In fact, these test–retest estimates suggested SOC was more temporally stable than N/BI.

These correlational models are limited because no inferences can be made about the extent of change or about predictors of individual differences in change over time. Thus, the longitudinal measurement models were respecified as LGMs. Because the majority of symptom change occurred during the 1st year of follow-up, linear growth was untenable. Accordingly, with the exception of the growth model for BA/P,[4] the Slope factor loadings were specified as follows for T1, T2, and T3, respectively: 0, *, and 1 (* = freely estimated). This factor loading specification centers the Intercept factor on the initial assessment point (i.e., mean and variance of the outcome when time = 0), and the mean of the Slope factor provides the estimates of the extent of change in the latent construct (in the metric of the marker indicator) over the 2-year period. Identification and parsimony of the models was further fostered by fixing the intercepts of the latent factors to zero and holding their residual variances to equality.

As shown in Table 2, each of these LGMs fit the data well, for example, N/BI: $\chi^2(23) = 22.58$, $p = .49$, SRMR = .03, RMSEA = 0.00 (90% CI = 0.00–0.03, CFit = .99), TLI = 1.00, CFI = 1.00. The freed T2 Slope factor loadings indicate the proportion of change that occurred between T1 and T2 relative to the total change over 2-year follow-up. As expected, the majority of change in each construct occurred in the 1st year of follow-up (e.g., 85% for N/BI). The direction and statistical significance ($ps < .001$) of the Slope means indicated that, on average, patients evidenced significant change in a less symptomatic direction on each temperament and disorder construct over follow-up. With the exception of the BA/P Slope, the variances of the Intercept and Slope of each construct were statistically significant, indicating there were nontrivial individual differences in the initial levels (T1) and change in these dimensions over time.

Effect sizes (Cohen's *d*) for the extent of change over the 2-year follow-up were estimated with the parameter estimates from the

[4] Inspection of the indicators of BA/P indicated no mean change between T2 and T3. Thus, the factor loadings for the BA/P slope were fixed at 0, 1, and 1.

longitudinal measurement models (i.e., factor means, variances, and intercorrelations). By Cohen's (1988) standards, the effect size for BA/P was small (.19), the effect sizes for the *DSM–IV* constructs were small to medium (range = .39 to .52), and the amount of change in N/BI approached a large effect (.70).

Results presented in Table 2 also show that, with the exception of BA/P, the Intercept and Slope factors were significantly correlated in each LGM. For the *DSM–IV* disorder constructs, these correlations were negative (range = −.42 to −.48), indicating that symptom reductions over time were most pronounced in patients with higher levels of initial (T1) symptoms (e.g., the higher the level of initial DEP severity, the greater the amount of DEP symptom decrease over follow-up).[5] Of interest, unlike the relationships seen for the *DSM–IV* constructs, the Intercept and Slope of N/BI were positively correlated ($r = .47$), indicating that patients with higher initial levels of N/BI tended to show fewer reductions in this dimension over time.

Finally, to evaluate the hypothesis that psychosocial treatment would be more effective in reducing symptoms of the *DSM–IV* disorder constructs than features of temperament, LGMs were conducted by regressing the Slopes of N/BI, DEP, SOC, and GAD onto their respective Intercepts and a treatment dummy code (0 = *no CARD treatment,* 1 = *received CARD treatment*). This analysis was not conducted for BA/P in view of the nonsignificant variance of this construct's Slope. In each instance, the Treatment → Slope path was statistically significant and in a direction indicating that, holding initial symptom levels constant, patients who received CARD treatment evidenced greater symptom reductions than untreated patients. Table 2 also provides standardized path coefficients that can be interpreted in the Cohen's *d* metric (i.e., holding initial status constant, the amount of standard deviation change in the Slope given an unstandardized unit change in the treatment covariate). As seen in Table 2, the effect size of treatment was largest for N/BI ($\gamma = -.98$).

Latent Growth Models of Parallel Processes

The next set of analyses entailed the directional relationships among the temperament and *DSM–IV* disorder constructs over time. It was predicted that the initial level of N/BI would predict individual differences in the extent of change in DEP, SOC, and GAD and that initial level of BA/P would predict change in DEP and SOC, but not GAD (cf. T. A. Brown et al., 1998). Conversely, it was expected that the initial levels of the *DSM–IV* disorder constructs would not predict change in the temperament constructs over time. These hypotheses were tested with parallel-process LGMs in which multiple growth processes were simultaneously estimated with regressions among the growth factors. The model parameterization is illustrated in Figure 2 with the use of N/BI and GAD. The residual variances of the parallel constructs at each time point (e.g., N/BI and GAD at T1) were specified to be intercorrelated (but held equal over time to foster model parsimony), in accord with the notion that the covariance of the constructs within a given assessment could not be fully accounted for by the underlying growth processes. Because excessive multicollinearity would exist otherwise (cf. intercept correlations in Figure 3; e.g., GAD, DEP, N/BI), these models were conducted separately for each *DSM–IV* construct. Moreover, for each disorder construct, the parallel LGMs were estimated with N/BI and BA/P alone and together (to examine the effect of each construct alone and holding the other constant).

The results of the parallel-process models are summarized in Figure 3. For space reasons, only the models that simultaneously estimated the effects of N/BI and BA/P are presented. To foster presentational clarity, the measurement model portion of the parallel-process LGMs has been omitted from this figure. As indicated in the caption for Figure 3, each of these models fit the data well. The predictions were partially supported in that initial levels of N/BI predicted variability in the rate of change in SOC and GAD in the expected direction (i.e., holding initial levels of these disorder constructs constant, higher initial N/BI was associated with less symptom improvement in SOC and GAD). Figure 4 depicts the nature of this effect on SOC and GAD (cf. Curran, Bauer, & Willoughby, 2004).[6] However, initial N/BI did not predict change in DEP, either when analyzed alone or with BA/P in the analysis (see Figure 4). Moreover, although initial BA/P significantly predicted less improvement in SOC when analyzed alone ($\gamma = -.18$, $p < .01$), this effect approached statistical significance ($\gamma = -.13$, $p = .07$) when N/BI was included in the model. BA/P did not predict change in DEP, either when analyzed alone or with N/BI.[7] Consistent with prediction, BA/P was not related to change in GAD.

As hypothesized, initial levels of GAD and DEP were not associated with changes over time in N/BI (because the BA/P Slope had a nonsignificant variance, it was not regressed onto any variables in the model). However, a significant SOC Intercept → N/BI Slope path was obtained, suggesting a bidirectional relationship of these constructs (i.e., the N/BI Intercept → SOC Slope path was also significant). Nevertheless, the direction of this relationship was counterintuitive ($\gamma = -.58$), indicating that persons with higher initial levels of SOC displayed greater decreases in N/BI, holding initial N/BI constant.

Do N/BI and BA/P Account for the Cross-Sectional and Temporal Covariance of DSM–IV Disorder Constructs?

A noteworthy finding of T. A. Brown et al. (1998) was the ability of negative affect and positive affect to account for virtually

[5] Interpretation of the Intercept–Slope correlations may be fostered by understanding that the random effect slopes are negatively signed for patients showing symptom reductions over follow-up (e.g., Slope = −10 indicates a 10-unit decrease in symptoms from T1 to T3). Thus, for an outcome such as DEP, a negative Intercept–Slope correlation ($r = -.42$; see Table 2) indicates that symptom reductions are more pronounced (the values of the random effect slopes become more negative) as the initial level of DEP increases (DEP intercept centered on T1).

[6] For the purposes of illustrating the positive associations between the disorder constructs and N/BI at intake (see Figure 3), Figure 4 is based on an equivalent model parameterization in which the N/BI Intercept correlations in Figure 3 are replaced by directional paths (e.g., N/BI Intercept → GAD Intercept).

[7] An attempt was made to replicate the findings of Kasch et al. (2002) with the use of analyses that more directly resembled those undertaken by these authors. Each of the three T2 indicators of depression was regressed onto the four T1 BIS/BAS scales (and the corresponding T1 depression indicator covariate). In all three analyses, none of the BAS scales accounted for significant variance in T2 depression.

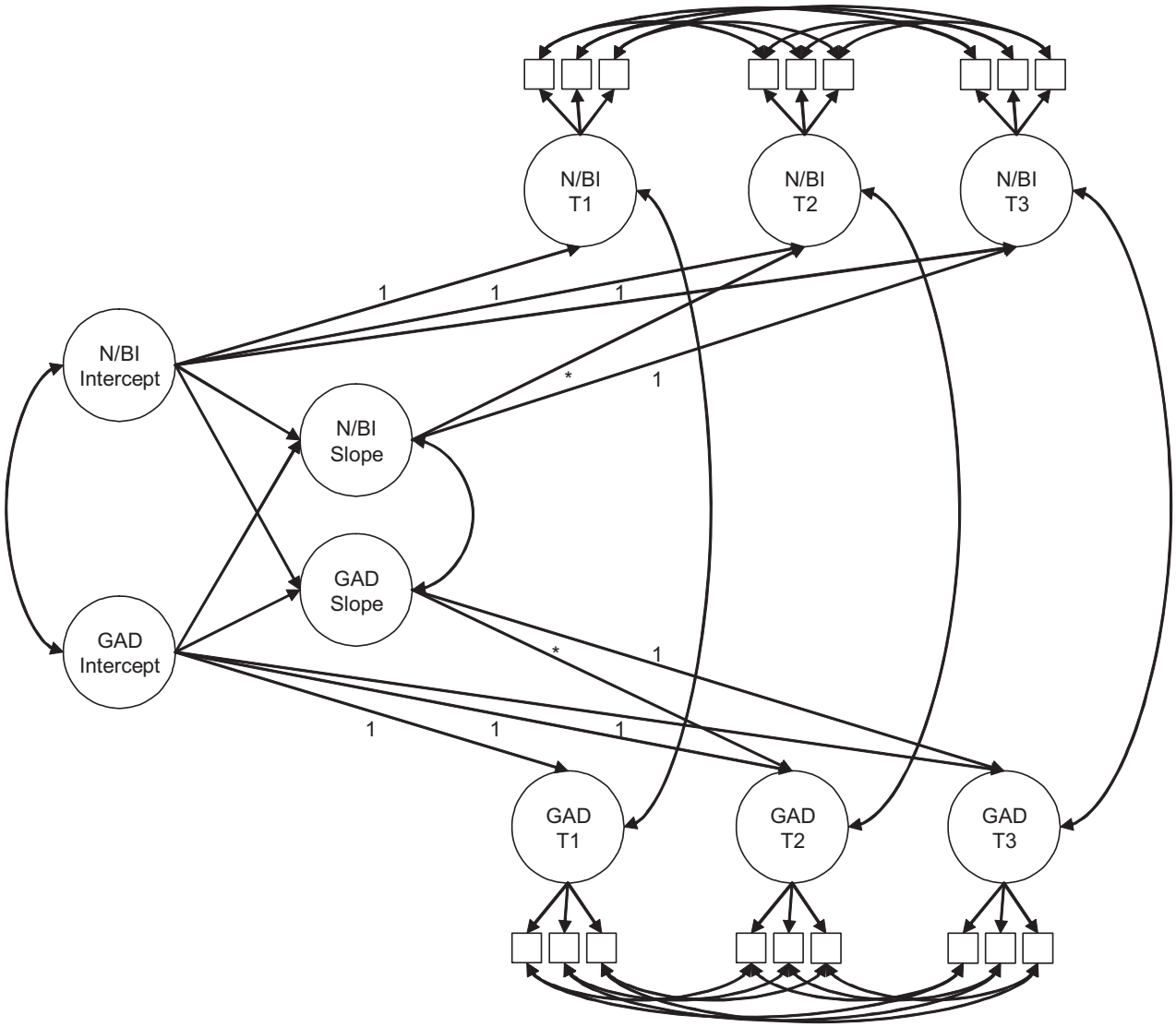


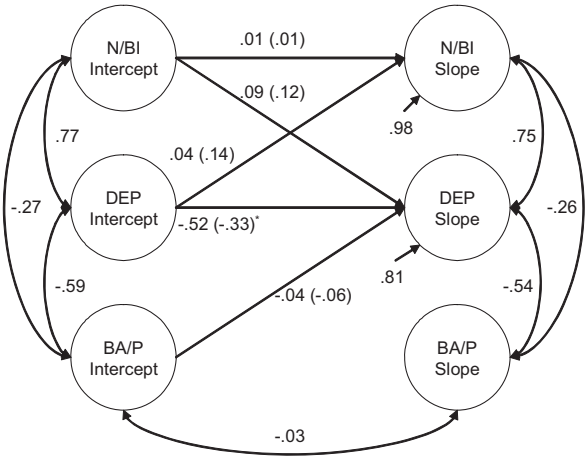
Figure 2. Example of parallel-process latent growth model. N/BI = neuroticism/behavioral inhibition; GAD = generalized anxiety disorder; T1 (Time 1) = intake; T2 (Time 2) = 12-month follow-up; T3 (Time 3) = 24-month follow-up. Asterisks represent freely estimated parameter.

all the considerable covariance among the major *DSM–IV* anxiety and mood disorder constructs. The present study attempted to replicate and extend this finding by evaluating whether (a) the overlap in the initial levels of DEP, SOC, and GAD could be explained by initial N/BI and BA/P and (b) the temporal covariance of these *DSM–IV* disorder constructs could be accounted for by change in N/BI (BA/P was not included in the second analysis because of its nonsignificant Slope variance).

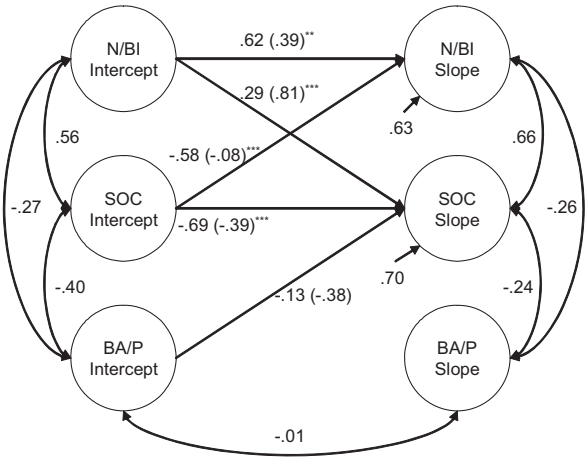
As seen in Table 3, the Intercepts and Slopes of the disorder constructs were significantly intercorrelated. In the first structural model, the Intercepts of the disorder constructs were regressed onto the Intercepts of N/BI and BA/P in accord with the structural relationships reported in T. A. Brown et al. (1998; the remaining growth factors were permitted to freely covary). As shown in the caption for Figure 5, this model fit the data well (Model A). Consistent with prediction, all three disorders were predicted significantly by N/BI, with DEP and GAD evidencing the strongest associations (γs = .65 and .68, respectively). Moreover, holding N/BI constant, BA/P predicted significant additional variance in DEP and SOC (γs = $-.39$ and $-.29$, respectively), but not GAD. All the covariance among the *DSM–IV* disorder constructs was explained by N/BI and BA/P, except for a trivial ($r = -.06$) yet statistically significant ($p < .05$) association between SOC and GAD.

In the second model (see Figure 5, Model B), the Slopes of the *DSM–IV* disorder constructs were regressed onto the Slope of N/BI. BA/P was not included because of the nonsignificant variance component of its Slope and its inability to account for unique variance in the disorder outcomes in previous analyses. This model also fit the data well (see caption for Figure 5). As seen in the

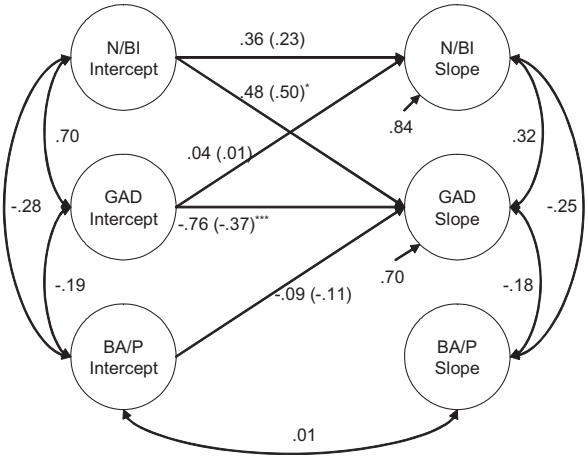
A.



B.



C.



figure, temporal change in DEP, SOC, and GAD was significantly related to change in N/BI (range of γs = .43 to .81). Of particular interest is the finding that all the temporal covariance of these *DSM–IV* disorder constructs was accounted for by change in N/BI; that is, when N/BI was specified as a predictor, the temporal overlap among disorder constructs was reduced to zero (i.e., all the correlated change in DEP, GAD, and SOC was explained by change in N/BI).

Discussion

Partial support was obtained for the prediction that dimensions of temperament would be more temporally stable than *DSM–IV* disorder constructs. Consistent with prior evidence in smaller clinical samples (e.g., Kasch et al., 2002; B. Meyer, Johnson, & Winters, 2001), the dimension of BA/P displayed a remarkable level of stability over the 2-year interval ($d = .19$, $r = .84$), despite the fact that over three quarters of participants received treatment during this time. The finding that the variance component of the BA/P was nonsignificant (see Table 2), although precluding further analyses (e.g., predictors of BA/P change), indicated this relative lack of change was uniform in the sample. However, of the five constructs examined, N/BI evidenced the greatest amount of temporal change ($d = .70$) and was the dimension associated with the largest treatment effect ($d = -.98$). Although at odds with the results of at least one investigation (Kasch et al., 2002) and strict trait conceptualizations of temperament, this finding is similar to a handful of studies that reported decreases in N/BI type measures following psychological or pharmacological treatment for major depression (e.g., Clark et al., 2003).

Although impact on BA/P was negligible, it is important to note that change in the five constructs was fostered by aspects of the study design (e.g., use of a heterogeneous clinical sample, T1 assessment occurred at intake, most patients received treatment). Furthermore, self-report (and interview-based) assessment of temperament is prone to mood-state distortion (cf. Clark et al., 2003; Widiger et al., 1999). As noted in the beginning of this article, the measurement of N/BI seems to consist of some combination of stable variance and variability attributable to generalized distress (i.e., more prone to mood-

Figure 3. Parallel-process models of temperament and *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) disorder constructs. Completely standardized and unstandardized (in parentheses) estimates are shown. A: Depression (DEP). B: Social phobia (SOC). C: Generalized anxiety disorder (GAD). N/BI = neuroticism/behavioral inhibition; BA/P = behavioral activation/positive affect. Overall fit of models: Model A, $\chi^2(223) = 493.30$, $p < .001$, standardized root-mean-square residual (SRMR) = .05, root-mean-square error of approximation (RMSEA) = 0.04 (90% confidence interval [CI] = 0.04–0.05, test of close fit [CFit] = .99), Tucker–Lewis index (TLI) = 0.96, comparative fit index (CFI) = .97; Model B, $\chi^2(229) = 453.90$, $p < .001$, SRMR = .06, RMSEA = 0.04 (90% CI = 0.03–0.05, CFit = 1.00), TLI = 0.97, CFI = .97; Model C, $\chi^2(231) = 436.56$, $p < .001$, SRMR = .05, RMSEA = 0.04 (90% CI = 0.03–0.04, CFit = 1.00), TLI = 0.96, CFI = .97. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

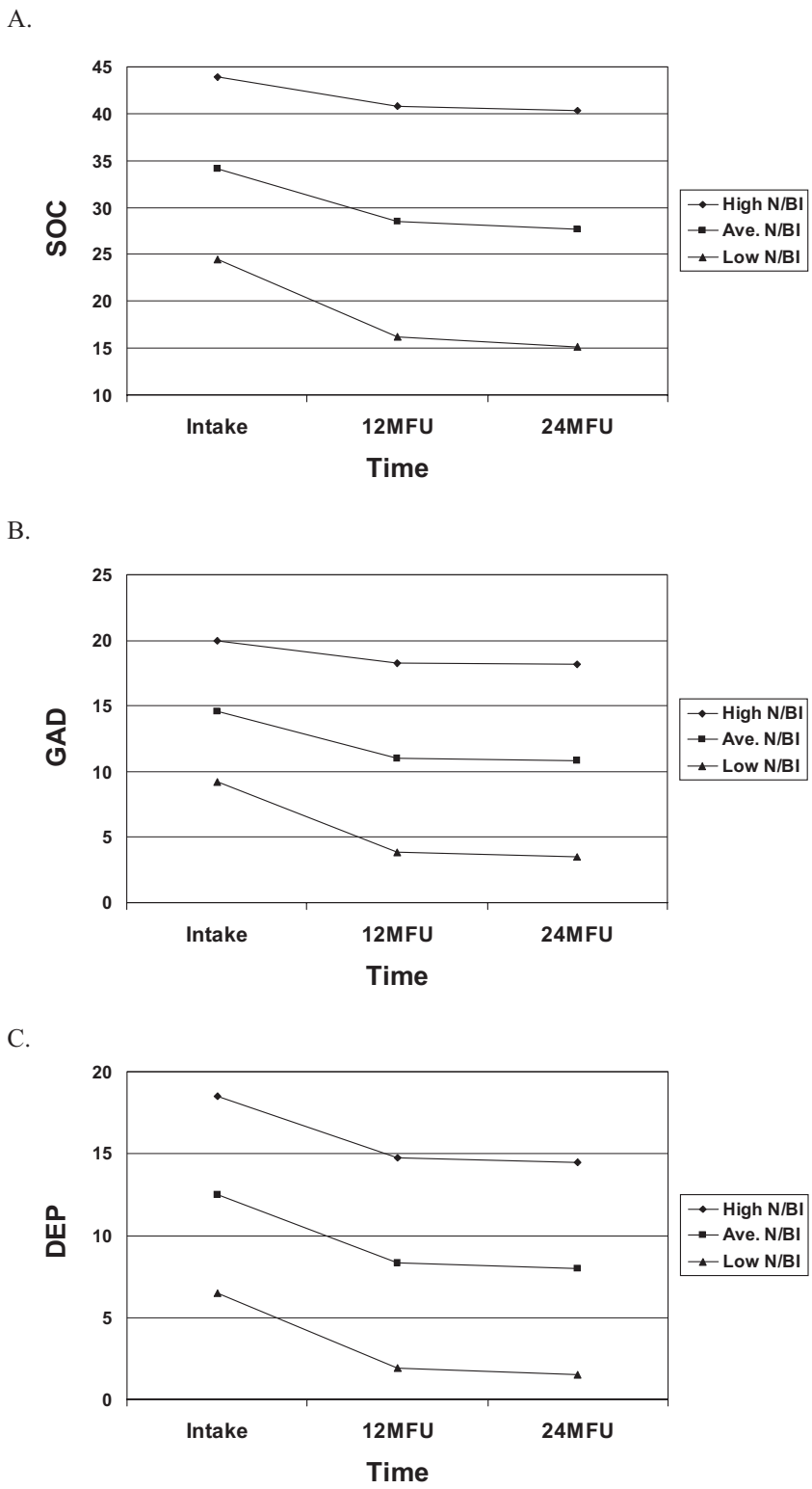


Figure 4. Model-implied trajectories of *Diagnostic and Statistical Manual of Mental Disorder* (4th ed.) disorder constructs as a function of neuroticism/behavioral inhibition. A: Social phobia (SOC). B: Generalized anxiety disorder (GAD). C: Depression (DEP). 12MFU = 12-month follow-up; 24MFU = 24-month follow-up; N/BI = neuroticism/behavioral inhibition; Ave. = average.

Table 3
Cross-Sectional and Temporal Correlations of DSM–IV Disorder Constructs

Construct	1	2	3	4	5	6
1. DEP_{INT}	—					
2. GAD_{INT}	.49	—				
3. SOC_{INT}	.47	.26	—			
4. DEP_{SLP}	−.43	−.09	−.27	—		
5. GAD_{SLP}	.13	−.40	−.03	.30	—	
6. SOC_{SLP}	−.03	.08	−.47	.54	.33	—

Note. *DSM–IV* = *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.); DEP = depression; GAD = generalized anxiety disorder; SOC = social phobia; INT = intercept; SLP = slope.

state distortion, subject to greater temporal fluctuation).[8] In other words, the considerable covariance between N/BI and the emotional disorders is due partly to a temperamental component (e.g., N/BI acts as a trait vulnerability dimension), but also to the distress associated with having a disorder (mood-state distortion). Presumably, the latter aspect is less stable and more apt to covary with temporal fluctuations in the severity of disorders (cf. Figure 5, Model B).

Nevertheless, it is important to underscore the various manners in which N/BI operated differently from the *DSM–IV* disorder constructs. For instance, LGMs of each *DSM–IV* disorder construct revealed inverse relations between the Intercept and Slope (*r*s = −42 to −.48); that is, higher initial disorder severity was associated with greater change over time. This could be interpreted along the lines of a regression to the mean effect, often seen in LGM studies of psychopathological processes (e.g., Curran, Stice, & Chassin, 1997). However, the Intercept and Slope of N/BI were positively correlated (r = .47), indicating that patients with higher initial levels of N/BI tended to show fewer reductions in this dimension over time, and patients with lower initial levels of N/BI tended to evidence greater decreases. Thus, unlike the *DSM–IV* disorders, the stability of N/BI increased as a function of initial severity. This suggests that the influence of mood-state distortion and/or general distress on the measurement of N/BI is most pronounced at the lower end of its continuum—it is the lower range of N/BI that is less temporally stable and more apt to covary with temporal change in disorder severity. Conversely, the stability of more severe initial N/BI may be indicative of a genetically based vulnerability dimension that is resistant to natural or treatment-induced remission (cf. T. A. Brown et al., 1995).

Although it is difficult to disentangle traitlike (e.g., vulnerability) and statelike (e.g., general distress) variance from measures of temperament, a multivariate analytic approach is useful in this endeavor (e.g., simultaneously modeling multiple latent dimensions of temperament and psychopathology; see Clark et al., 2003, for an alternate approach). Accordingly, the parallel-process LGMs (see Figure 3) addressed the unique contribution of initial levels of N/BI and BA/P in the prediction of temporal change in *DSM–IV* disorder constructs, holding initial levels of the disorders constant (and vice versa). In addition to the latent variable framework, the use of initial status of *DSM–IV* disorders as covariates may remove some of the general distress component from N/BI (and BA/P) in the prediction of disorder outcomes. Consequently, in two instances (i.e., SOC and GAD as outcomes), higher initial levels of N/BI were associated with less change in the *DSM–IV* construct (see Figure 4). Moreover, lower BA/P predicted poorer outcome of SOC, although this effect approached statistical significance (p = .07) when N/BI was included in the analysis. Although no temporal relations were obtained for DEP (see below), these results are in line with some earlier work and theory that N/BI and BA/P have directional temporal effects on Axis I psychopathology (Gershuny & Sher, 1998; Kasch et al., 2002; B. Meyer et al., 2001; but see Clark et al., 2003).

The measurement of the *DSM–IV* disorders contained some stable variance. This seemed particularly evident in the SOC latent variable, whose univariate stability was second only to BA/P in terms of test–retest correlations (e.g., T2–T3 r = .94) and mean change (d = .39). Such findings could be taken as further evidence of the illusory boundary between Axis I psychopathology and personality (Widiger & Samuel, 2005). However, it is interesting to consider the results of the multivariate LGMs (i.e., N/BI and BA/P are added as covariates), in which possibly some of this traitlike variance was partialed out of initial *DSM–IV* severity. In each analysis, the "regression to the mean" type relationships seen for the *DSM–IV* construct in the LGMs became more pronounced. For example, in the LGM, the correlation between the Intercept and Slope of SOC was −.48 (see Table 2), but this association increased to −.69 in the parallel-process LGM (standardized regressive path; see Figure 3). This change was most dramatic for GAD; that is, Intercept–Slope relationships were −.42 and −.76 in the univariate and multivariate models, respectively. Such findings lend some support for the distinction of dimensions of temperament and *DSM–IV* psychopathology. For instance, although the cross-sectional overlap of N/BI and GAD was considerable (Intercept r = .70), simultaneous longitudinal modeling indicated that these constructs operate and interact with one another differently over time (e.g., N/BI augmented a "regression to the mean effect" and predicted GAD change but not vice versa). These results also underscore the value of examining the nature of the relationships between temperament and psychopathol-

[8] Measures of N/BI may be differentially prone to mood-state distortion. Temporal effect size calculations of each indicator of N/BI revealed that the PANAS–N evidenced the greatest change (d = .82), and BIS the least (d = .50; but higher than BIS change reported in Kasch et al., 2002). This is consistent with findings that measures using single adjectives as items (e.g., PANAS) are most susceptible to mood-state distortion (e.g., Widiger et al., 1999).

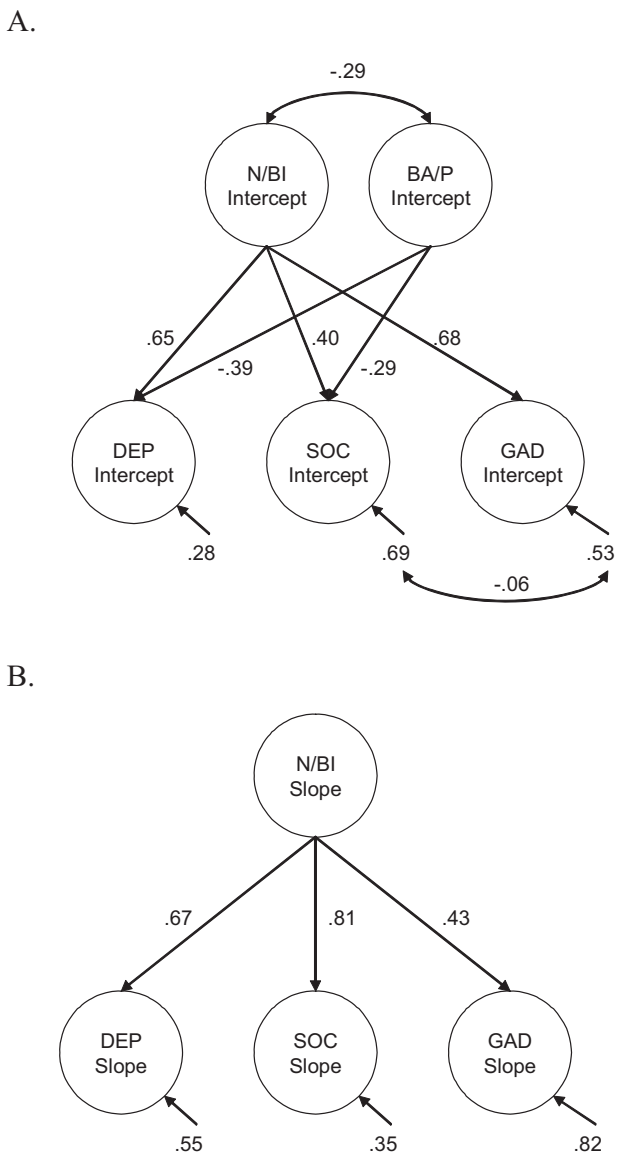


Figure 5. Dimensions of temperament as higher order factors explaining the covariance of *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) disorder constructs. Completely standardized estimates are shown; all corresponding unstandardized coefficients are significant ($ps < .05$). A: Cross-sectional covariance (intake). B: Temporal covariance (change over 24 months). Overall fit of models: Model A, $\chi^2(767) = 1{,}639.58$, $p < .001$, standardized root-mean-square residual (SRMR) = .06, root-mean-square error of approximation (RMSEA) = 0.04 (90% confidence interval [CI] = 0.04–0.05, test of close fit [CFit] = 1.00), Tucker–Lewis index (TLI) = 0.94, comparative fit index (CFI) = .95; Model B, $\chi^2(553) = 1{,}272.19$, $p < .001$, SRMR = .05, RMSEA = 0.05 (90% CI = 0.04–0.05, CFit = .96), TLI = 0.95, CFI = .95. N/BI = neuroticism/behavioral inhibition; BA/P = behavioral activation/positive affect; DEP = depression; SOC = social phobia; GAD = generalized anxiety disorder.

ogy in longitudinal designs in a multivariate analytic framework (e.g., use of constructs defined by multiple indicators, simultaneous modeling of multiple constructs; cf. Clark et al., 1994).

Consistent with prediction and conceptual models, initial levels of the *DSM–IV* disorders did not predict increases in temperament over time. However, one directional effect was found: Higher initial SOC was associated with greater reductions in N/BI. Although this finding is inconsistent with theory, it appears to have some connection with prior research. T. A. Brown et al. (1995) found that the presence of *DSM–III–R* (American Psychiatric Association, 1987) social phobia at pretreatment was a significant predictor of favorable response to treatment of panic disorder, a result that was recently replicated in a large multicenter outcome trial (Allen et al., 2003). Similar findings were reported by Chambless, Renneberg, Goldstein, and Gracely (1992). In an attempt to interpret these counterintuitive results, Chambless et al. (1992) speculated that the relationship between social phobia and favorable outcome was mediated by treatment compliance (prompted by fear of negative evaluation by the therapist). To the extent this explanation is viable, one should also consider the impact social anxiety may have on the measurement of temperament and psychopathology. That is, the measurement of these dimensions in socially anxious patients, especially after treatment, may be prone to the influence of social desirability (i.e., self-imposed pressure to embellish symptom improvement because of fears of being perceived as a "bad patient"). Along these lines, it is interesting to note that, unlike models involving GAD and major depressive disorder, the magnitude of the effect of N/BI intercept on slope increased ($\gamma = .62$) in the parallel-process LGM with SOC beyond that observed at the zero-order level ($r = .47$). Thus, removal of a social desirability element may augment the previously observed association of high initial N/BI and less change.

Although the current findings support the discriminant validity of dimensions of temperament and *DSM–IV* psychopathology (e.g., good model fit for the Figure 5 solutions, in which latent variables of temperament and *DSM–IV* constructs were modeled simultaneously), certain study design aspects may foster this differentiation. Specifically, temperament was measured by questionnaires with trait-type instructions, whereas the *DSM–IV* constructs were assessed with questionnaires and clinical ratings, most involving brief time frames (e.g., past month). Nonetheless, several findings suggest that method effects did not unduly influence the results (e.g., SOC was second to BA/P in terms of temporal stability; N/BI evidenced the most temporal change). However, it is also possible that the differential breadth of coverage of the temperament and *DSM–IV* constructs impacted the general pattern of directional effects obtained (i.e., N/BI → *DSM–IV* constructs, but not vice versa). For instance, it might be argued that the *DSM–IV* disorders reflect narrowly defined subcomponents of the broad dimensions of N/BI and BA/P that, if aggregated, would operate in a manner similar to temperament constructs (cf. the continuity explanation of temperament and psychopathology).[9]

Despite considerable methodological differences (e.g., use of a parallel-process LGM framework, substantially different indicator

[9] I am grateful to a reviewer for raising this possibility, which unfortunately could not be tested in the current data set because of the multivariate overlap (multicollinearity) involving N/BI with the *DSM–IV* disorder constructs at Time 1 (i.e., precluded a model in which N/BI, SOC, DEP, and GAD were simultaneous predictors of temporal change in temperament and *DSM–IV* psychopathology).

sets), the results presented in Figure 5 replicate and extend the findings of T. A. Brown et al. (1998). For instance, the cross-sectional relationships among the dimensions of temperament and *DSM–IV* disorders are quite similar (e.g., N/BI–DEP $r = .77$; cf. Table 3 in the current study and in the T. A. Brown et al., 1998, study). In the multivariate model, N/BI and BA/P accounted for virtually all the T1 covariance among the highly overlapping disorder constructs (see Figure 5, Model A). Moreover, the pattern of relationships was in accord with theory and prior evidence; for example, although N/BI was also predictive of SOC ($\gamma = .40$), it evidenced its strongest relationships with GAD and DEP (γs = .68 and .65, respectively). In addition, the specificity of the influence of BA/P on DEP and SOC was upheld (i.e., fit diagnostics indicated model fit would not be improved by the addition of a BA/P → GAD path). Although the discriminant validity of GAD and unipolar depression has been questioned because of high phenotypic overlap (T. A. Brown et al., 1994), such findings may lend further support for the differentiation of these conditions. Namely, the consistent finding that BA/P is relevant to DEP but not GAD is in accord with the conceptual position for the existence of multiple dimensions of risk, which have differential relevance to the various emotional disorders (i.e., the vulnerability factors for GAD and DEP are neither singular nor entirely overlapping; e.g., T. A. Brown & Barlow, 2002; Clark, 2005; Mineka et al., 1998). This is also in agreement with recent behavioral genetic evidence (e.g., Eley & Brown, 2006) indicative of an additive genetic structure that parallels the higher order models obtained in the current study and in the T. A. Brown et al. (1998) study.

The temporal covariance among the three *DSM* disorder constructs was not extreme (range of *r*s = .30 to .54; see Table 3), a finding that lends further evidence in favor of the discriminant validity of these selected disorder constructs; that is, change in one disorder dimension was not collinear with change in another (less than 30% shared variance). Of particular note is that temporal (co)variance of these disorder constructs was strongly related to, and fully accounted for, by change in N/BI (see Figure 5, Model B). The correlational nature of these findings precludes firm conclusions about the direction of these effects. Although the finding of N/BI reductions after treatment are more common (e.g., Clark et al., 2003) than not (Kasch et al., 2002), the reasons for these decreases are not clear. It is possible that N/BI is therapeutically malleable (cf. treatment effect size in Table 2), and this in fact mediates the extent of change in the emotional disorders. Alternatively or additively, a reduction in disorder severity is associated with a decrease in generalized distress, a feature shared by the emotional disorders, and is partially reflected in the measurement of N/BI. Finer grained methodologies are required to explicate the nature of these relationships (e.g., therapy outcome-based designs entailing multiple assessments of temperament and disorder features over the active treatment phase). In any case, these findings extend the extant literature by demonstrating, in a longitudinal context, the role of N/BI as a unifying construct in accounting for the covariance among the emotional disorders.

Although the current study represents an initial large-scale ($N = 606$, multiple indictor assessment) examination of the temporal relationships of temperament and psychopathology in a clinical sample, certain limitations should be kept in mind in the consideration of study findings. For instance, the assessment battery did not include a measure of life stress. Given evidence that the effects of vulnerability on psychopathology are potentiated by stressful life events (e.g., Kendler, Kuhn, & Prescott, 2004), the influence of N/BI and BA/P on the course of the *DSM–IV* disorder constructs may have been underestimated. Another limitation pertains to the use of a clinical sample. Although temporal stability and directional effects could be examined, use of a patient sample precluded a more comprehensive evaluation of the various explanations for the relationships between temperament and psychopathology (e.g., predispositional, complication/scar). Whereas it is important to examine temporal relations in the context of clinical disorders (e.g., do broader personality constructs predict the temporal course of psychopathology?), it may be that the ability to detect directional effects in patient samples is limited. Although the present study obtained several directional effects, perhaps temperament has its strongest effects on psychopathology as a predispositional influence. That is, once a disorder has emerged, other factors may be far more influential to its course and complications (e.g., environmental aspects such as social support and access to treatment).

Moreover, at the psychometric level, the overlap of psychopathology and temperament (N/BI, in particular) is exacerbated in clinical samples by generalized distress. Although the associations observed among dimensions of temperament and *DSM–IV* psychopathology did not exceed conventional guidelines for unacceptable discriminant validity (cf. T. A. Brown, 2006; Kenny, 1979), general distress likely contributed to the considerable overlap observed at the T1 assessment. In some cases (e.g., DEP–N/BI $r = .77$; cf. T. A. Brown et al., 1998), this overlap imposed restrictions on the amount of unique variance that could contribute to the prediction of individual differences in change in *DSM–IV* psychopathology (in addition to strong effects between the disorder's Intercept and Slope). From a clinical perspective, it seems plausible that depression might be associated with higher levels of general distress than most anxiety disorders, which may further account for the differential T1 overlap among Intercepts seen in Figure 3 (e.g., SOC–N/BI $r = .56$). Along these lines, Clark et al. (2003) conjectured that the inconsistent findings regarding the predictive effects of pretreatment temperament on depression treatment outcome may be due to varying degrees of trait and state variance in temperament measures across study samples (i.e., higher levels of unstable "state" variance mask stable variance in temperament and its predictive effects in clinical samples). Study differences in the proportion of trait versus state variance are likely influenced by the nature of the sample (e.g., community vs. clinic-based, outpatient vs. inpatient) and the measures of temperament used (cf. Footnote 8). Thus, in addition to clinical samples, longitudinal examinations of the etiological role of N/BI and BA/P should continue to involve twin, community, and at-risk samples using multiple, judiciously selected indicators of temperament. Regardless of the sample and measures used, such studies should address the issue of mood-state effects (i.e., trait vs. state variance) in the methodological design and interpretation of findings.

References

Allen, L. B., Barlow, D. H., White, K. S., Gorman, J. M., Shear, M. K., & Woods, S. W. (2003, November). Comorbidity as a moderator and outcome measure in patients receiving cognitive–behavioral treatment for panic disorder with agoraphobia. In D. H. Barlow (Chair), *The*

multicenter clinical trials on panic disorder: Mediators, moderators, and mechanisms of action. Symposium conducted at the meeting of the Association for Advancement of Behavior Therapy, Boston, MA.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112,* 545–557.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

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

Andrews, G. (1996). Comorbidity in neurotic disorders: The similarities are more important than the differences. In R. M. Rapee (Ed.), *Current controversies in the anxiety disorders* (pp. 3–20). New York: Guilford Press.

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment, 10,* 176–181.

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.

Bollen, K. A., & Curran, P. J. (2005). *Latent curve models: A structural equation approach*. New York: Wiley.

Boyce, P., & Parker, G. (1985). Neuroticism as a predictor of outcome in depression. *Journal of Nervous and Mental Disease, 173,* 685–688.

Brown, E. J., Turovsky, J., Heimberg, R. G., Juster, H. R., Brown, T. A., & Barlow, D. H. (1997). Validation of the Social Interaction Anxiety Scale and the Social Phobia Scale across the anxiety disorders. *Psychological Assessment, 9,* 21–27.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Brown, T. A., Antony, M. M., & Barlow, D. H. (1995). Diagnostic comorbidity in panic disorder: Effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology, 63,* 408–418.

Brown, T. A., & Barlow, D. H. (2002). Classification of anxiety and mood disorders. In D. H. Barlow (Ed.), *Anxiety and its disorders: The nature and treatment of anxiety and panic* (2nd ed., pp. 292–327). New York: Guilford Press.

Brown, T. A., & Barlow, D. H. (2005). Categorical vs. dimensional classification of mental disorders in *DSM–V* and beyond. *Journal of Abnormal Psychology, 114,* 551–556.

Brown, T. A., Barlow, D. H., & Liebowitz, M. R. (1994). The empirical basis of generalized anxiety disorder. *American Journal of Psychiatry, 151,* 1272–1280.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the *DSM–IV* anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110,* 585–599.

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., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy, 35,* 79–89.

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. (2005). A psychometric reanalysis of the Albany Panic and Phobia Questionnaire. *Behaviour Research and Therapy, 43,* 337–355.

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-Sills, L. A., Liverant, G., & Brown, T. A. (2004). Psychometric evaluation of the Behavioral Inhibition/Behavioral Activation Scales (BIS/BAS) in large clinical samples. *Psychological Assessment, 16,* 244–254.

Carver, C. S., & 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.

Chambless, D. L., Renneberg, B., Goldstein, A., & Gracely, E. J. (1992). MCMI-diagnosed personality disorders among agoraphobic outpatients: Prevalence and relationship to severity and treatment outcome. *Journal of Anxiety Disorders, 6,* 193–211.

Clark, L. A. (2005). Temperament as a unifying basis for personality and psychopathology. *Journal of Abnormal Psychology, 114,* 505–521.

Clark, L. A., Vittengl, J., Kraft, D., & Jarrett, R. B. (2003). Separate personality traits from states to predict depression. *Journal of Personality Disorders, 17,* 152–172.

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

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

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2004). Testing main effects and interactions in latent curve analysis. *Psychological Methods, 9,* 220–237.

Curran, P. J., & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 107–135). Washington, DC: American Psychological Association.

Curran, P. J., Stice, E., & Chassin, L. (1997). The relation between adolescent alcohol use and peer alcohol use: A longitudinal random coefficients model. *Journal of Consulting and Clinical Psychology, 65,* 130–140.

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM–IV: Lifetime Version (ADIS–IV–L)*. New York: Oxford University Press.

Eley, T. C., & Brown, T. A. (2006). *Phenotypic and genetic/environmental structure of anxiety and depressive disorder symptoms in adolescence*. Manuscript submitted for publication.

Eysenck, H. J. (1981). *A model for personality*. New York: Springer.

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

Fanous, A., Gardner, C. O., Prescott, C. A., Cancro, R., & Kendler, K. S. (2002). Neuroticism, major depression, and gender: A population-based twin study. *Psychological Medicine, 32,* 719–728.

Geerts, E., & Bouhuys, N. (1998). Multilevel prediction of short-term outcome of depression: Nonverbal interpersonal processes, cognitions, and personality traits. *Psychiatry Research, 79,* 59–72.

Gershuny, B. S., & Sher, K. J. (1998). The relation between personality and anxiety: Findings from a 3-year prospective study. *Journal of Abnormal Psychology, 107,* 252–262.

Gray, J. A. (1987). *The psychology of fear and stress* (2nd ed.). Cambridge, England: Cambridge University Press.

Hancock, G. R., Kuo, W. L., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling, 8,* 470–489.

Hettema, J. M., Prescott, C. A., & Kendler, K. S. (2004). Genetic and environmental sources of covariation between generalized anxiety disorder and neuroticism. *American Journal of Psychiatry, 161,* 1581–1587.

Hettema, J. M., Prescott, C. A., Myers, J. M., Neale, M. C., & Kendler, K. S. (2005). The structure of genetic and environmental risk factors for anxiety disorders in men and women. *Archives of General Psychiatry, 62,* 182–189.

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

Joyce, P. R., Mulder, R. T., & Cloninger, C. R. (1994). Temperament predicts clomipramine and desipramine response in major depression. *Journal of Affective Disorders, 30,* 35–46.

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.

Kendler, K. S., Kuhn, J., & Prescott, C. A. (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *American Journal of Psychiatry, 161,* 631–636.

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley-Interscience.

Kenny, D. A., & Zautra, A. (2001). Trait–state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 243–263). Washington, DC: American Psychological Association.

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.

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.

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.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28,* 487–495.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49,* 377–412.

Muthén, L. K., & Muthén, B. O. (1998–2005). *Mplus 3.13* [Computer software]. Los Angeles: Author.

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. (2005). Analysis of longitudinal studies with missing data using covariance structure modeling with full-information maximum likelihood. *Structural Equation Modeling, 12,* 493–505.

Sato, T., Hirano, S., Narita, T., Kusunoki, K., Kato, J., Goto, M., et al. (1999). Temperament and character inventory dimensions as a predictor of response to antidepressant treatment in major depression. *Journal of Affective Disorders, 56,* 153–161.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Viken, R. J., Rose, R. J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology, 66,* 722–730.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to the anxiety and depressive disorders. *Journal of Abnormal Psychology, 97,* 346–353.

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., & Samuel, D. B. (2005). Diagnostic categories or dimensions: A question for *DSM–V*. *Journal of Abnormal Psychology, 114,* 494–504.

Widiger, T. A., & Seidlitz, L. (2002). Personality, psychopathology, and aging. *Journal of Research in Personality, 36,* 335–362.

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 February 1, 2006
Revision received October 18, 2006
Accepted October 23, 2006 ■