Affective Instability: Measuring a Core Feature of Borderline Personality Disorder With Ecological Momentary Assessment

Article *in* Journal of Abnormal Psychology · August 2008

DOI: 10.1037/a0012532 · Source: PubMed

7 authors, including:

Timothy J Trull
University of Missouri
198 PUBLICATIONS **9,738** CITATIONS
SEE PROFILE

Seungmin Jahng
Sungkyunkwan University
19 PUBLICATIONS **1,086** CITATIONS
SEE PROFILE

Phillip Karl Wood
University of Missouri
136 PUBLICATIONS **6,492** CITATIONS
SEE PROFILE

David Watson
University of Notre Dame
245 PUBLICATIONS **74,284** CITATIONS
SEE PROFILE

Journal of Abnormal Psychology
2008, Vol. 117, No. 3, 647–661

0021-843X/08/$12.00 DOI: 10.1037/a0012532

Affective Instability: Measuring a Core Feature of Borderline Personality Disorder With Ecological Momentary Assessment

Timothy J. Trull, Marika B. Solhan, Sarah L. Tragesser, Seungmin Jahng, Phillip K. Wood, and Thomas M. Piasecki
University of Missouri—Columbia

David Watson
University of Iowa

Ecological momentary assessment (EMA; Stone & Shiffman, 1994) was used to characterize and quantify a dynamic process—affective instability in borderline personality disorder (BPD). Sixty outpatients (34 with BPD and affective instability; 26 with current depressive disorder but not with BPD or affective instability) carried electronic diaries for approximately 1 month and were randomly prompted to rate their mood state up to 6 times a day. Results indicated that BPD patients (a) did not report significantly different mean levels of positive or negative affect; (b) displayed significantly more variability over time in their positive and negative affect scores; (c) demonstrated significantly more instability on successive scores (i.e., large changes) for hostility, fear, and sadness than did patients with depressive disorders; and (d) were more likely to report extreme changes across successive occasions (≥90th percentile of change scores across participants) for hostility scores. Results illustrate different analytic approaches to quantifying variability and instability of affect based on intensive longitudinal data. Further, results suggest the promise of electronic diaries for collecting data from individuals in their natural environment for purposes of clinical research and assessment.

Keywords: borderline personality disorder, ecological momentary assessment, affective instability, electronic diaries

The purpose of this study was to use an innovative methodology, *ecological momentary assessment* (EMA; Stone & Shiffman, 1994), to characterize and quantify affective instability. Affective instability is a dynamic process that is associated with several major forms of psychopathology, including borderline personality disorder (BPD; American Psychiatric Association [APA], 2000; Linehan, 1993). Across numerous conceptualizations and definitions of BPD, the affective features that show this instability include depression or dysphoria; irritability; anger or hostility; and anxiety, panic, or fear (APA, 2000; Gunderson, 2001). These extreme shifts in mood typically last from a few hours to a few days and may occur as a result of interpersonal stresses or identity crises, for example. Although these affects (i.e., dysphoria, irritability, anxiety) clearly overlap with those characteristic of other mental disorders (e.g., major depressive disorder [MDD]), what distinguishes the affective instability of BPD is reactivity to environmental stimuli (versus an insidious onset), as well as a transient, fluctuating course. Several studies with questionnaire-based measures of affective instability have suggested that BPD patients can be distinguished from other personality-disordered or bipolar II patients on the basis of mood shifts from anger and anxiety to euthymia (e.g., Henry et al., 2001; Koeningsberg et al., 2002).

In addition to being a symptom of BPD, affective instability (or emotional dysregulation) may in fact be the driving force behind many additional behaviors seen in the disorder. For example, Linehan (1993) has postulated that in BPD, emotional vulnerability and inability to regulate emotions lead to maladaptive attempts to regulate intense affective states or to control problematic outcomes associated with these affective states. Impulsive behavior (including suicidal behavior) may be seen as a maladaptive solution to painful negative affect, identity disturbance may result from a lack of emotional consistency and predictability, and disturbed interpersonal relationships may be the product of the difficulty in regulating emotional states and impulses, as well as the inability to tolerate painful stimuli. As another example, it has been proposed that those with BPD may be especially vulnerable to developing substance use disorders because alcohol or drugs may be used to cope with negative affective states. Therefore, affective instability can be quite dysfunctional in its own right and may well contribute to other symptoms and features of BPD.

Measuring and characterizing the experience of affective instability have proved challenging, however. Traditional measures of affective instability rely on respondents' retrospective recall and subjective assessment of affective variability or reactivity on interview or questionnaire items. For example, structured interviews for this BPD criterion ask respondents to judge whether they have often had strong mood shifts within a day over the last 2 to 5 years. The Affective Lability Scale (Harvey, Greenberg, & Serper, 1989) is perhaps the questionnaire measure most frequently used by clinical researchers to quantify affective

Timothy J. Trull, Marika B. Solhan, Sarah L. Tragesser, Seungmin Jahng, Phillip K. Wood, and Thomas M. Piasecki, Department of Psychological Sciences, University of Missouri—Columbia; David Watson, Department of Psychology, University of Iowa.

This research was supported in part by National Institute of Mental Health Grant MH-69472 and by funding from the Borderline Personality Disorder Research Foundation.

Correspondence concerning this article should be addressed to Timothy J. Trull, Department of Psychological Sciences, University of Missouri—Columbia, 106C McAlester Hall, Columbia, MO 65211. E-mail: trullt@missouri.edu

instability in BPD and other psychiatric disorders (e.g., Henry et al., 2001; Koeningsberg et al., 2002). The Affective Lability Scale asks respondents to rate how well a list of statements characterizes the respondent (e.g., "One minute I can be feeling OK, and then the next minute I'm tense, jittery, and nervous."). However, no time frame is given, nor are any thresholds provided (i.e., How many times must this experience occur in order for it to be characteristic?). Thus, it is left to respondents to decide on the time frame for the responses (e.g., Over their lifetime? Over the past year?) as well as the threshold for endorsement.

Retrospective recall is suspect for a number of reasons (e.g., Hufford, Shiffman, Paty, & Stone, 2001). Memory of past events is influenced by cognitive processes used to reconstruct past events. Individuals are more likely to recall or report experiences that seem more personally relevant, that occurred more recently, that stand out as significant or unusual, or that are consistent with their current mood state. As a result, an individual may be biased when recalling past events or experiences, and such biases are particularly likely when an individual is asked to aggregate moods or experiences over time.

These findings are relevant to both retrospective reports of typical mood state and experience and frequency of past life events. Evidence suggests that retrospective report of mood by questionnaire, for example, will be primarily influenced by the moment of peak affect intensity and the affect at the end of the assessment period (i.e., the peak-and-end rule; Fredrickson, 2000). Further, personality traits like neuroticism may influence one's beliefs about experience, including emotional experience, suggesting an additional source of bias in the retrospective report of mood by individuals or people high in neuroticism (Robinson & Clore, 2002). For example, Ebner-Priemer et al. (2006) found that people with BPD, compared with control participants, exhibited a negative recall bias (more intense negative affect and less intense positive affect) when real-time ambulatory assessment mood data for a 24-hr period were compared with retrospective ratings for this period. When individuals are additionally asked about their reactivity in mood, it seems unlikely that they could correctly recall and assess fluctuations in mood over time, given that they are limited in their ability to report past mood states accurately. In summary, there are multiple reasons that retrospective report of mood and of mood change or instability is limited. The use of real-time data collection methods, like EMA, may be especially valuable in the assessment of mood state and mood instability.

EMA

EMA (Stone & Shiffman, 1994) is well-suited to address these limitations, and this method shows great promise as a means to capture dynamic processes important to psychopathology. In EMA, ambulatory data collection methods (often diaries) are used to minimize the need for retrospective reporting. The defining characteristics of EMA are that assessments are both ecological (experiences are measured in the participant's natural environment) and momentary (assessments capture information about immediate or near immediate experiences and require minimal retrospection). A series of immediate reports can be statistically aggregated to summarize daily experience without relying on participants' memory. EMA is a conceptual strategy encompassing a wide array of specific techniques. EMA studies may vary with respect to the manner of which experiences are sampled (e.g., time sampling, event sampling, or a combination), the response channels assessed (e.g., subjective states, discrete behaviors, physiological measures, cognitive performance), and the platform for data collection (e.g., paper-and-pencil diaries, ambulatory physiological monitoring, palmtop computers). An important forerunner, the experience sampling method (ESM; Csikszentmihaly & Larson, 1987) has been considered a subtype of EMA in which time-based schemes are used to sample experience, and assessments are focused on private, subjective states (Stone & Shiffman, 2002). However, some scholars consider the ESM and EMA labels to be interchangeable (Scollon, Kim-Prieto, & Diener, 2003).

EMA has many advantages over traditional research designs when investigators are interested in characterizing dynamic, clinically important, psychological processes (Piasecki, Hufford, Solhan, & Trull, 2007). The ecological nature of the assessments is a clear advantage—processes such as mood can be studied in individuals' natural habitats, where individuals are subject to the many environmental and interpersonal factors that typify everyday life but that cannot be recreated in the laboratory. Further, EMA methods can not only sample the process of interest (e.g., mood) but can also sample the characteristics of the environment (e.g., location, time of day, presence of interpersonal conflict) that change over time and that may be important for explaining variation. Finally, compared with traditional forms of assessment, EMA is simpler and less subject to bias and forgetting.

EMA and ESM Studies of Affective Instability in BPD

To date, only a limited number of empirical studies have been conducted with EMA or ESM to assess mood instability in BPD patients. In general, studies that compared mood or aversive tension score variability between BPD patients and healthy control participants report significantly greater variability among BPD patients (Cowdry, Gardner, O'Leary, Leibenluft, & Rubinow, 1991; Ebner-Priemer et al., 2007; Stein, 1996; Stiglmayr, Grathwol, & Bohus, 2001; Woyshville, Lackamp, Eisengart, & Gilliland, 1999), especially for negative mood states. However, in two studies (Farmer, Nash, & Dance, 2004; Russell, Moskowitz, Zuroff, Sookman, & Paris, 2007), a significant difference in mood variability between BPD participants and healthy control participants was not found. Two of these studies also compared the mood variability of BPD patients with that observed in other psychiatric patient groups (Cowdry et al., 1991; Stein, 1996). In both of these studies, it was found that the variability of mood scores for BPD patients was significantly greater than that found for other patient groups (e.g., major depression, anorexia nervosa), although Stein (1996) found this was the case only for positive affective states. Finally, a recent study by Links et al. (2007) examined affective instability only within BPD patients and reported that the amplitude of the affective states was negatively correlated with the association of successive mood ratings, implying that more extreme mood ratings were less predictable over time than were less extreme mood ratings.

Overall, most of these studies suggest some form of affective disturbance is common in those diagnosed with BPD, and this affective experience can be distinguished from that of other individuals by its variability. Despite these important findings, however, there are a number of limitations. First, it was not always clear whether studies were assessing affective instability as commonly conceptualized. As we discuss below, some of the previously used indices of affective instability (e.g., standard deviations

of scores over time) are limited in what they tell us about mood fluctuations, and they may not adequately capture all components of affective lability (e.g., frequency of change, variability, temporal dependency, amplitude; Larsen, 1987).

In addition, many of the studies were limited methodologically in their ability to describe and capture the nature of affective instability. For example, some studies used a relatively short period of time to assess moods of BPD participants (e.g., 1 day to 1 week). Given the *Diagnotic and Statistical Manual of Mental Disorders* (4th ed., text rev. [*DSM-IV-TR*]; APA, 2000) definition of affective instability, these short assessment periods may not be sufficient to capture the multiple instances of mood instability that characterize BPD. Another methodological limitation is that the studies with paper-and-pencil diaries were not able to document that the patients did indeed complete the assessments at the time they were prompted to do so. In a related vein, several studies reported relatively low compliance rates, leaving open the possibility that the patients may not have been reporting all the mood fluctuations they experienced, for example. Missing data points are especially problematic when characterizing processes like affective instability that are defined by successive scores from occasion to occasion. Finally, and most important, almost all previous studies have used healthy controls as a comparison group. It is of greater interest to demonstrate that the mood patterns that characterize BPD patients with affective instability are distinguishable from those presented by other near-neighbor diagnostic groups also characterized by affective disturbance.

The Current Study

In the present study, we used EMA methodology to examine affective instability in BPD. We hypothesized that (a) patients diagnosed with BPD would not differ from those with major depression or dysthymia in mean levels of positive or negative affect; (b) these two groups of patients would, however, differ in the variability of negative affect scores (i.e., the overall amount of dispersion of scores over the course of the study); (c) BPD patients would show more instability in moment-to-moment negative affect scores over time (i.e., changes in scores from one assessment to the next); and (d) a higher probability of large, acute changes in negative affect scores would characterize BPD patients.

Method

Participants

This study targeted two groups of psychiatric outpatients. The BPD group ($n = 34$) included psychiatric outpatients who presented with a *DSM-IV-TR* BPD diagnosis and the presence of significant features of affective instability. The second group, the major depressive disorder/dysthymic disorder (MDD/DYS) group ($n = 26$), included psychiatric outpatients who met criteria for either a current *DSM-IV-TR* MDD or a current DYS diagnosis and who did not meet criteria either for BPD in general or for the specific feature of affective instability. Recruitment took place at four local psychiatric outpatient clinics. Potential participants were made aware of the study through flyers in the waiting rooms, at intake sessions, and through their assigned therapists or doctors. Interested participants completed a form giving permission to research staff to contact the participants about the study, to review their medical records, and to consult with their treating therapists about the patients' current diagnosis. If it appeared the patients might be eligible, they were invited to a screening session (which included both Axis I and Axis II diagnostic interviews) to establish diagnostic eligibility. Patients with current psychosis, with current mania, with a history of significant head trauma or neurological dysfunction, or with a current history of severe substance abuse or a past history of severe and sustained substance dependence were excluded. Those that were eligible for the study were then scheduled for an orientation session (see below). All potential participants were paid $20 for completing the screening interview.[1]

Table 1 presents demographic characteristics of the sample, by group. For the entire sample, the average age of participants was 34.98 years ($SD = 12.25$), and the majority of participants were women (88.3%; 11.7%, men); White non-Hispanic (86.7%; 10.0%, African American; 1.7%, Hispanic; 1.7%, Asian American); and single, divorced, or separated (66.7%; 18.3%, married; 15.0%, cohabitating). Participants reported a family income of $25,000.00 or less (71.7%; 18.3%, $25,001 to $50,000; 10.0%, $50,001 or above). All participants but 1 reported taking at least one type of psychotropic medication at the beginning of the assessment period, primarily antidepressants (72.1%) or anxioloytics (51.1%). Fifty percent of the sample reported being currently employed full- or part-time. Most participants reported at least one previous psychiatric hospitalization (63.3%). As for current Axis I disorders, the most frequently diagnosed anxiety disorders were generalized anxiety disorder (36.7%), social phobia (35.0%), and posttraumatic stress disorder (31.7%); the most frequently diagnosed mood disorders were MDD (55.0%) and DYS (23.3%). A smaller percentage of participants had current alcohol abuse or dependence (3.3%) or current cannabis abuse or dependence (3.3%). The most common Axis II disorders, other than BPD, were avoidant (36.7%), obsessive compulsive (13.3%), and antisocial (11.7%) personality disorders.

Measures

Psychiatric diagnoses. To establish diagnostic eligibility for the study, all patients completed the Structured Clinical Interview for *DSM-IV* Axis I Disorders (First, Spitzer, & Williams, 1995) and the Structured Interview for *DSM-IV* Personality (Pfohl, Blum, & Zimmerman, 1994). Audio recordings of the Structured Clinical Interview for *DSM-IV* Axis I Disorders and Structured Interview for *DSM-IV* Personality interviews from a randomly selected sample of 14 participants were reviewed and scored independently by an alternative interviewer who served as a reliability checker. Agreement was excellent for the presence or absence of affective instability ($\kappa = 1.0$), a diagnosis of MDD/DYS ($\kappa = 1.0$), a

[1] To examine whether our final sample was representative of all those that we interviewed with BPD and all those we interviewed with MDD/DYS, we examined the diagnostic interview data for the 53 individuals who we interviewed but who were not eligible for our study due to the exclusionary criteria. Of these, 1 person received a current diagnosis of major depression but met the diagnostic criterion for affective instability, 3 individuals received a diagnosis of BPD but did not meet the affective instability diagnostic criterion, and 11 others met the affective instability diagnostic criterion but did not receive a diagnosis of BPD.

Table 1
Demographic Characteristics and Affect Scores for Currently Depressed (MDD/DYS) and Borderline Personality Disorder (BPD) Participants

	MDD/DYS				BPD			
Characteristic	*M*	*SD*	*SE*	%	*M*	*SD*	*SE*	%
Age	37.69	12.16			33.26	12.41		
Women				76.9%*				97.1%*
Men				23.1%				2.9%
White, non-Hispanic				84.6%				90.6%
Other race or ethnicity				15.4%				9.4%
Single, divorced, or separated				69.2%				64.7%
Employed full- or part-time				36.0%				60.6%
Family income less than $25,000				73.1%				70.6%
Previous hospitalization				46.2%*				76.5%*
Antidepressant medication				90.0%				65.4%
Anti-anxiety medication				60.0%				44.0%
Current Axis I diagnoses								
Generalized anxiety				30.8%				42.4%
Social phobia				38.5%				33.3%
Posttraumatic stress				26.9%				36.4%
Major depression				80.8%*				36.4%*
Dysthymic disorder				38.5%*				12.1%*
Alcohol use disorder				0.0%				6.1%
Cannabis use disorder				3.8%				3.0%
Current Axis II disorder								
Avoidant				23.1%				48.5%
Obsessive compulsive				11.5%				15.2%
Antisocial				3.8%				18.2%
Dependent				3.8%				12.1%
PANAS scores from electronic diary assessment								
Positive Affect item score	2.12		0.12		2.25		0.11	
Negative Affect item score	1.55		0.11		1.70		0.10	
Hostility item score	1.36		0.08		1.54		0.07	
Fear item score	1.61		0.13		1.73		0.12	
Sadness item score	1.79		0.13		1.77		0.11	

Note. For MDD/DYS, $n = 26$; for BPD, $n = 34$. *ns* range from 58–60. MDD/DYS = major depressive disorder/dysthymic disorder; BPD = borderline personality disorder.
* Indicates two groups differed, $p < .05$.

diagnosis of BPD ($\kappa = .85$), and the number of BPD symptoms present (intraclass correlation coefficient = .96).

EMA. Mood descriptor items from the Positive and Negative Affect Schedule (PANAS; Watson & Clark, 1999) and the expanded form (PANAS-X) (Watson & Clark, 1999) were presented to each participant on the electronic diary (ED) during each momentary assessment over 28 days. The PANAS (Watson & Clark, 1999) consists of two 10-item scales, one for positive affect and one for negative affect. Respondents consider each mood descriptor and then provide a rating that reflects the extent to which she or he felt this way (1–5; 1 = *very slightly or not at all*, 5 = *extremely*) in the designated time period. Although a number of time frames are possible, we asked respondents to rate the extent to which they felt this way since the last prompt. In addition to items from the PANAS positive and negative affect scales, we administered several additional mood items from the PANAS-X (Watson & Clark, 1999) so that we would be able to also calculate scores for the following negative affect subscales: Hostility (six items); Fear (six items); and Sadness (five items). These latter three mood subscales were used to characterize specific affects relevant to BPD and to the *DSM-IV-TR* definition of affective instability.

Procedures

Participants were issued a Palm Zire 31 handheld computer to record their affects, experiences, and behaviors six times a day over a 28-day period. The EDs were loaded with Pendragon Forms software (Version 4.0; Pendragon Software Corp., Libertyville, IL), which displayed each question to participants along with a series of check boxes corresponding to possible responses; respondents answered questions by simply tapping the Palm stylus on the box representing the best answer. Each ED was loaded with software custom-written for this project that allowed flexible configuration of random prompts. Each participant's typical rising and retiring times were entered. The software stratified the waking hours into six equal intervals and then randomly selected one moment within each interval to deliver a prompt. Prompts could not occur within 20 min of each other, however. When a prompt was triggered, the ED emitted an audible beep that repeated every 5 min for up to 10 min. Beeping was terminated if participants initiated the ED recording or if they failed to respond to the prompt within the 10-min period. Collected data were time stamped to determine whether participants responded to prompts in a timely manner and to

scale affective variability more accurately as a function of time, as described in the measures below.

At the initial orientation and training session, research staff explained the use of the ED devices. Research staff demonstrated the audible prompt, explained how to initiate responding when a prompt was sounded, and emphasized the importance of timely responding to ED prompts. In addition, research staff demonstrated and observed the use of the stylus to respond to each question on the ED assessment battery. Research staff answered any questions participants had about either the devices or the assessments. Finally, research staff informed participants that compliance would be checked at data download visits. Although timely responding was emphasized, staff explained that delayed responding might be unavoidable in certain situations (e.g., when driving a car). Participants were trained to respond as soon as was safely possible in such circumstances. At the end of the training session, participants were scheduled for individual weekly data download visits. The EDs were programmed so that the first random prompt occurred the evening of the training, and staff called participants to make sure that all had gone well.

At the weekly data download visits, research staff met one-on-one with participants and downloaded the past week's recorded data from the ED device. These visits also provided an opportunity to check ED recording compliance and reinforce good recording practices. Uploaded ED recordings from the past week were scanned for missed prompts or unusually short or long responses to individual assessment items. If necessary, staff queried participants about the circumstances surrounding any problems identified in the data. At each weekly data download session, participants were paid $45.

Overview of Statistical Modeling

Several features of the ED data are important to note: (a) Each participant may have a different total number of usable assessments; (b) assessment times within a day were randomly selected; and (c) these times vary randomly across people. The data for each participant are therefore unbalanced in terms of the number of observations and time intervals between observations. These data features necessitate the use of multilevel modeling (MLM) when examining mean levels of mood and when comparing groups on indices of variability and instability. Further, in some analyses, it was necessary to adjust successive differences according to the time intervals so that valid comparisons could be made.

We conducted a series of models that probed distinct conceptualizations of affective instability. First, basic multilevel models were used to characterize group differences in mean levels and overall dispersion of affect scores. Next, a model accounting for systematic, time-bound variation in affect was conducted, permitting examination of group differences in unpredictable or residual affect variation. These time-detrended data were then used to determine whether BPD patients showed larger successive differences (i.e., changes from one measurement to the next). Finally, the successive difference scores computed in this model were used to test whether BPD patients were more likely to report large-amplitude acute changes (i.e., successive differences at or above the 90th percentile of all those observed). Thus, the analyses were cumulative, with each taking advantage of the results of prior models to test refined conceptualizations of affective instability. To facilitate comprehension of the series, we interleave descriptions of the theoretical motivation and computation of each model with presentation of model findings in the Results section.

Results

Preliminary Considerations

We first examined the rate of compliance with prompts for completing mood ratings on the ED. The maximum number of possible mood ratings that could be completed each day was 6; As a result, 168 ratings were possible for those who participated for 28 days (28 days × 6 prompts per day). However, 8 participants did not complete the full 4 weeks of assessment (due to their own time constraints), and some individuals completed mood ratings on more than 28 days (due to the scheduling of the last data download). In addition, some responses were not usable due to incomplete ratings or to assessments that were completed without a prompt (participants sometimes thought they heard the prompt when it did not occur). In all, the BPD patients produced 5,170 complete sets of mood ratings (M = 152.06), and the MDD/DYS patients contributed 3,951 complete sets of mood ratings (M = 151.96). The mean lag time between prompt and participant response for each group did not differ significantly (BPD M = 2.88 min, MDD/DYS M = 2.40 min), $F(1, 9061) = 0.35$, *ns*. Compliance rate was calculated by dividing the number of completed and usable sets of mood ratings by the total number of prompts while the patient was in the study. Average compliance rates were .87 (SD = .07) for the BPD group and .88 (SD = .07) for the MDD/DYS group, Wald's $\chi^2(1) = 1.46$, ns.

Mean and Variance Differences in Mood Between BPD and MDD Patients

Overview. We hypothesized that the two groups of patients would not differ in overall levels of positive or negative affect. However, we predicted that BPD patients would show more variability in negative affect scores than would the MDD/DYS patients. Means and standard errors for all affect scores by group are shown in Table 1.

In order to test our hypotheses, we used MLM, which is appropriate when different numbers of observations for each participant are possible, as in the present study.

The multiple equation form of the model was

$$\text{Level 1: } Y_{ti} = b_{0i} + e_{ti}, \text{ and}$$

$$\text{Level 2: } b_{0i} = \gamma_{00} + \gamma_{01}\text{Group}_i + u_{0i}, \quad (1)$$

where Y_{ti} indicates t_{th} assessment for person i, $e_{ti} \sim N(0, \sigma_k^2)$ and $u_{0i} \sim N(0, \tau^2)$. The random effect b_{0i} represents the mean of each individual and the fixed effects γ_{00} and γ_{01} indicate the group mean of MDD/DYS and the group mean difference between BPD and MDD/DYS, respectively. The same (homogenous) variance model (i.e., positing that the variance of affect scores for the two groups was equal) assumes $\sigma_1^2 = \sigma_2^2$, whereas the different (heterogeneous) variance model assumes $\sigma_1^2 \neq \sigma_2^2$, where σ_1^2 is the variance of the MDD/DYS group and σ_2^2 is the variance of the BPD group. The significance of γ_{01} in Equation 1 provides a statistical test of group differences in mean affect scores. A chi-square test of the difference in the log likelihood of the models, with and without the assumption of equal variances, was used to test the hypothesis of equal variances across groups.

Analyses of mood level and variability. As indicated in Tables 1 and 2, BPD patients and MDD/DYS patients did not differ in terms of their mean level (across all assessment occasions) of negative affect (i.e., the general Negative Affect [NA] composite, as well as the Sadness, Fear, or Hostility subscales) or positive affect. However, as indicated in Table 2, we found that BPD patients displayed significantly more variable affect scores across time than did the MDD/DYS patients. That is, for positive affect scores and for each negative affect score, the model that allowed different variances for the two groups provided a better fit to the data than did the model that specified equal variances. In each case, BPD patients displayed more variability in scores across time.

Difference in Affective Instability Between BPD and MDD Patients

Characterizing affective instability. Examining the affect variance scores provides an incomplete assessment of affective instability. As defined earlier, affective instability involves extreme and frequent fluctuations of mood over time, and these changes are relatively unpredictable (under a model specification without covariates) in the sense that they are believed to be influenced by external events. In addition, because mood changes characterizing affective instability are assumed to be unsystematic, in the sense that they are not affected by the systematic mean change over time, it is necessary to remove any temporally dependent or systematic variance in mood change across time before examining amplitude of change and frequency of change (Jahng, Wood, & Trull, in press; Tennen, Affleck, & Armeli, 2005; West & Hepworth, 1991).

Variance decomposition approaches to affect variability. To examine the trends that underlie affective variation over time, our first step was to decompose the total variance in mood scores for each person into variance accounted for by time factors and residual variance. We chose four possible time factors that might influence systematic change of the mean affect level: (a) trends due to time of day of assessment; (b) differences in affect due to assessments obtained on weekdays versus weekend days; (c) general trends of increase or decrease over all days of assessment; (d) and trends reflecting interaction of weekend versus weekday with time of day. To obtain variance accounted for by the time factors, multiple regressions were used for each individual because the trends in data may vary from person to person. Group mean difference in each variance component (i.e., variance accounted for by each time factor and residual variance) was tested by a generalized linear model with gamma error and log link function.[2] Because variance components are linked by log function, the ratio of the mean of each variance component of the two groups was tested against 1 (no difference) to investigate the mean difference in each variance component, and confidence intervals of the ratio were estimated.

Assuming that mood variation is influenced by external events or factors not included in the model, variance accounted for by each time factor can be seen as caused by external or unknown factors that vary systematically across the time unit specified by each time factor. For example, if an individual shows a systematic difference in negative affect between weekday and weekend, resulting in a significant amount of variance accounted for by the weekend factor, the individual may have experienced a systematic difference in the external events influencing negative affect between weekday and weekend. Accordingly, individual or group differences in variance accounted for by each time factor lead to two possible interpretations. The degree of the (time dependent) systematic change of underlying factors influencing affect may vary from one participant to another. For example, one participant may experience very different external events (influencing negative affect) on weekdays than on weekends, in comparison with another individual. As an alternative, assuming the time dependent systematic variations of unknown factors are the same across individuals or groups, the difference in the negative affect variance accounted for by time factors may reflect a greater reactivity to the external events influencing negative affect for one individual compared with another.

As seen in Table 3, significant mean differences in variance accounted for by the model including all time factors (total) were found in negative affect and positive affect (MDD/DYS model variance = .031 and BPD = .056, 95% confidence interval of the ratio [.31, .99] for negative affect; MDD/DYS = .065 and BPD = .101, 95% confidence interval of the ratio [.42, .99] for positive affect). Concerning individual time factors influencing negative affect scores, significant group mean differences were observed for the linear trend of time of day effect (negative affect, hostility), the quadratic trend of time of day (sadness), the interaction effect of linear trend of time of day by weekend (negative affect, hostility, fear), and the interaction effect of quadratic trend of time of day by weekend (negative affect, fear). The linear trend of day effect and interaction effect of linear trend of time of day by weekend in positive affect were significantly different for the groups. For residuals, the BPD group had greater variance in negative affect and fear than did the MDD/DYS group, which indicates that the BPD group has more variability in negative affect and fear than does the MDD/DYS group, over and beyond the systematic time effects specified above.

Temporal instability: Adjusted squared successive differences (ASSD). The variance decomposition into systematic change due to time factors and residual variance allowed us to focus the remaining analyses on the variance in mood scores that was not accounted for by time factors. This is preferred because common definitions of affective instability refer to mood variation and change that is less systematic or predictable from previous mood states. After detrending the data, it is possible to evaluate the instability of mood scores by conducting additional analyses on individuals' residual scores. These residual scores represent mood variability that cannot be accounted for by systematic time effects.

A useful index of instability is the squared successive difference. By squaring the differences in successive mood ratings, larger changes are weighted more; this is consistent with conceptualizations of affective instability in BPD.[3] If assessments are equally spaced in time, the successive difference has the same meaning across all occasions. However, if the time between assessments is randomly spaced, as in this study, successive differences with different time intervals indicate different amounts of change. Assuming positive autocorrelation in the level of affect

[2] The use of a generalized linear model with gamma error distribution and log link function is supported by the fact that for a normally distributed variable, variance components obtained from a general linear model (e.g., multiple regression) are types of scaled chi-square distributions, special cases of gamma distribution. See Jahng et al. (in press) for more details.

[3] The successive differences were calculated only for within-day fluctuations. Although the affective change between the last assessment of a day and the first assessment of the next day can be of interest, we focused only on the change of affect within day.

Table 2
Estimates of Fixed and Random Effects and Model Fits for Homogeneous and Heterogeneous Error Variance Models

	Model 1: Homogeneous		Model 2: Heterogeneous	
Affect score and variable	Estimate	*SE*	Estimate	*SE*
Negative Affect				
Fixed effect				
Intercept	1.549	0.109	1.549	0.109
Group (MDD = 0, BPD = 1)	0.148	0.145	0.148	0.145
Random effect				
Level 2 variance (τ^2)	0.306		0.306	
Level 1 variance				
σ_1^2 (MDD)	0.283		0.207	
σ_2^2 (BPD)	0.283		0.342	
Model fit -2 log-likelihood	14,692.8		14,424.3*	
Positive Affect				
Fixed effect				
Intercept	2.123	0.124	2.123	0.124
Group (MDD = 0, BPD = 1)	0.125	0.165	0.125	0.165
Random effect				
Level 2 variance (τ^2)	0.400		0.400	
Level 1 variance				
σ_1^2 (MDD)	0.449		0.354	
σ_2^2 (BPD)	0.449		0.522	
Model fit -2 log-likelihood	18,875.1		18,710.9*	
Hostility				
Fixed effect				
Intercept	1.357	0.082	1.357	0.082
Group (MDD = 0, BPD = 1)	0.186	0.109	0.186	0.109
Random effect				
Level 2 variance (τ^2)	0.172		0.171	
Level 1 variance				
σ_1^2 (MDD)	0.369		0.264	
σ_2^2 (BPD)	0.369		0.449	
Model fit -2 log-likelihood	17,041.9		16,739.1*	
Fear				
Fixed effect				
Intercept	1.613	0.131	1.613	0.131
Group (MDD = 0, BPD = 1)	0.113	0.175	0.113	0.174
Random effect				
Level 2 variance (τ^2)	0.447		0.446	
Level 1 variance				
σ_1^2 (MDD)	0.301		0.225	
σ_2^2 (BPD)	0.301		0.360	
Model fit -2 log-likelihood	15,273.3		15,037.3*	
Sadness				
Fixed effect				
Intercept	1.794	0.127	1.794	0.127
Group (MDD = 0, BPD = 1)	−0.022	0.169	−0.021	0.169
Random effect				
Level 2 variance (τ^2)	0.416		0.416	
Level 1 variance				
σ_1^2 (MDD)	0.476		0.403	
σ_2^2 (BPD)	0.476		0.532	
Model fit -2 log-likelihood	19,414.5		19,330.7*	

Note. For all intercepts, $ps < .0001$. For all group effects, $ps > .05$. MDD = major depressive disorder; BPD = borderline personality disorder.
* Chi-square deviance tests between Model 1 and Model 2 for all affect scores are significant at $p < .0001$, in each case favoring Model 2 (heterogeneous error variances).

Table 3
Ratio of the Means of Variance Components Accounted for by Time Factors for the BPD and the MDD/DYS Groups

Variance	Affect				
	Negative affect	Positive affect	Hostility	Fear	Sadness
Model					
Total (all time factors)					
MDD/DYS	0.031	0.065	0.033	0.045	0.087
BPD	0.056	0.101	0.051	0.075	0.099
Ratio[a]	0.55*	0.64*	0.64	0.60	0.88
95% CI	0.31, 0.99	0.42, 0.99	0.34, 1.21	0.32, 1.15	0.50, 1.56
Days (linear)					
MDD/DYS	0.015	0.014	0.011	0.025	0.047
BPD	0.023	0.034	0.021	0.034	0.049
Ratio	0.63	0.42*	0.54	0.72	0.97
95% CI	0.24, 1.67	0.20, 0.87	0.22, 1.32	0.30, 1.74	0.43, 2.18
Days (quadratic)					
MDD/DYS	0.007	0.018	0.010	0.009	0.015
BPD	0.012	0.014	0.015	0.020	0.023
Ratio	0.54	1.33	0.68	0.42	0.65
95% CI	0.23, 1.27	0.68, 2.63	0.29, 1.61	0.18, 1.01	0.33, 1.30
Weekend					
MDD/DYS	0.003	0.006	0.005	0.004	0.004
BPD	0.007	0.007	0.005	0.006	0.009
Ratio	0.45	0.83	1.00	0.63	0.48
95% CI	0.18, 1.12	0.39, 1.76	0.43, 2.33	0.27, 1.47	0.23, 1.02
Time of day (linear)					
MDD/DYS	0.001	0.009	0.001	0.002	0.007
BPD	0.004	0.014	0.003	0.004	0.006
Ratio	0.33*	0.61	0.44*	0.42	1.12
95% CI	0.15, 0.71	0.25, 1.45	0.20, 0.99	0.18, 1.02	0.50, 2.51
Time of day (quadratic)					
MDD/DYS	0.003	0.013	0.003	0.004	0.009
BPD	0.003	0.022	0.002	0.004	0.004
Ratio	1.11	0.58	1.13	0.94	2.46*
95% CI	0.48, 2.56	0.29, 1.15	0.46, 2.75	0.38, 2.32	1.07, 5.69
Time of Day (linear) × Weekend					
MDD/DYS	0.001	0.002	0.001	0.002	0.002
BPD	0.004	0.005	0.004	0.004	0.003
Ratio	0.24*	0.40*	0.24*	0.38*	0.73
95% CI	0.10, 0.60	0.18, 0.90	0.11, 0.54	0.16, 0.88	0.32, 1.66
Time of Day (quadratic) × Weekend					
MDD/DYS	0.001	0.003	0.002	0.001	0.002
BPD	0.002	0.005	0.002	0.002	0.005
Ratio	0.26*	0.68	0.90	0.42*	0.47
95% CI	0.11, 0.59	0.32, 1.43	0.38, 2.11	0.20, 0.90	0.20, 1.09
Residual					
MDD/DYS	0.176	0.303	0.239	0.174	0.305
BPD	0.287	0.425	0.390	0.293	0.435
Ratio	0.61*	0.71	0.61	0.59*	0.70
95% CI	0.39, 0.96	0.48, 1.05	0.36, 1.06	0.36, 0.98	0.45, 1.09

Note. Mean differences were tested by a generalized linear model with log link and gamma error distribution. Model total = model that includes all time factors. Residual = means of variance not accounted for by all time factors. MDD = major depressive disorder; BPD = borderline personality disorder; MDD/DYS = major depressive disorder/dysthymic disorder; CI = confidence interval.
[a] Ratio = (MDD/DYS)/BPD: Ratio of 1 represents no difference.
* $p < .05$.

between two successive time points (i.e., a level of affect at one time point has a similar value with a level of affect at an adjacent time point), an absolute successive difference with the longer time interval tends to be greater than that with shorter time interval. If this is the case, the difference in amount between two successive changes with different time intervals does not indicate a difference in instability but only indicates positive autocorrelation.

Figure 1A shows that this concern is valid for our data. In Figure 1A, the absolute successive differences of the detrended (i.e., residual) negative affect scores are plotted against corresponding time intervals. The solid line, a smoothed regression line obtained by locally weighted regression, indicates a linear increase in mean absolute successive difference. For the successive differences with different time intervals to be comparable, it is necessary to adjust the successive differences according

to length of time interval. One way to adjust the successive difference score for the amount of elapsed time is to divide the successive difference by its time interval, $t_{i+1} - t_i$, where t_i is the time at occasion i. The longer a time interval, given the same amount of change, the smaller the adjusted successive change.

This approach, however, may produce too high of a value of adjusted successive difference with the shorter time interval and too low of a value of adjusted successive difference with a longer time interval. Here, we propose an adjustment in the calculation of successive difference by dividing the difference by $[(t_{i+1} - t_i)/\text{median}(TI)]^{1/n}$, where t_i is the time at occasion i and median (TI) is the median of the time intervals for all observations.[4] Because the purpose of the adjustment is to produce successive differences that do not depend on the length of time intervals, the value n is chosen to make the means of absolute adjusted successive differences as equal as possible across different time intervals. Figure 1B presents the smoothed regression lines for the absolute value of adjusted successive difference, adjusted by different values for n, on the time intervals between two consecutive assessments. As shown, the expected value of adjusted absolute successive difference looks constant across time interval when $n = 9$. For more details about this procedure, such as mathematical properties of the adjustment and selection rules for n, see Jahng et al. (in press). These considerations resulted in the following equation for the ASSD:

$$\text{ASSD}_{ij} = \left\{ \frac{r_{(i+1)j} - r_{ij}}{[(t_{(i+1)j} - t_{ij})/\text{median}(TI)]^{\frac{1}{n}}} \right\}^2, \tag{2}$$

where r_{ij} and t_{ij} indicate the residual (from detrending) and time at occasion i for person j, respectively. In this study, the optimal values for n were 9 for detrended negative affect, 8 for predetrended negative affect, 6 for detrended positive affect, 5 for predetrended positive affect, 30 for detrended hostility, 20 for predetrended hostility, 9 for detrended fear, 9 for predetrended fear, 6 for detrended sadness, and 7 for pre-detrended sadness.[5]

We used generalized multilevel modeling with gamma distribution and log link to test the mean difference in ASSD for all PANAS affect scores. This approach was used because of the unbalanced number of observations across participants and the nonnormal distribution of the error terms. Further, the distribution of squared difference scores was best characterized as a gamma distribution. The multiple equation form was

$$\text{ASSD}_{ij}|\alpha_j,\beta_j \sim \gamma(\alpha_j, \beta_j),\ E(\text{ASSD}_{ij}|\alpha_j,\beta_j) = \alpha_j\beta_j = \mu_j,$$

$$\text{Var}(\text{ASSD}_{ij}|\alpha_j,\beta_j) = \alpha_j\beta_j^2,$$

$$\text{Level 1 link function: } \eta_j = \log(\mu_j),$$

$$\text{Level 1 structural model: } \eta_j = b_{0j}, \text{ and}$$

$$\text{Level 2 model: } b_{0j} = \gamma_{00} + \gamma_{01}\text{Group}_j + u_{0j},\ u_{0j} \sim N(0, \tau^2).$$

As before, we predicted that BPD patients would demonstrate greater affective instability, especially for negative affects, than would the MDD/DYS patients. As indicated in Table 4 (Group contrast), the BPD group showed more instability on hostility, fear, and sadness scores. However, there were no significant differences in instability of positive affect or general negative affect scores.

Temporal instability: Probability of acute change. Finally, we focused on the frequency of large, extreme changes in the participants' affect scores. We defined an acute change in affect as a successive change that equaled or exceeded the value for the 90th percentile of the total distribution of change for all participants in the study. In this way, we could compute how many acute changes occurred for each participant. However, because the number of assessments is randomly spaced, it was necessary as before to adjust the index for acute change (adjusted acute change [AAC]):

$$\text{AAC}ij = 1, \text{ if}$$

$$\frac{r_{(i+1)j} - r_{ij}}{[(t_{(i+1)j} - t_{ij})/\text{median}(TI)]^{\frac{1}{n}}} > k, \text{ and} \tag{3}$$

$$\text{AAC}_{ij} = 0, \text{ if}$$

$$\frac{r_{(i+1)j} - r_{ij}}{[(t_{(i+1)j} - t_{ij})/\text{median}(TI)]^{\frac{1}{n}}} < k,$$

where k is the 90th percentile of

$$\frac{r_{(i+1)j} - r_{ij}}{[(t_{(i+1)j} - t_{ij})/\text{median}(TI)]^{\frac{1}{n}}}.$$[6]

We used a multilevel logistic model to compare probability of AAC between the BPD group and the MDD/DYS. The multiple equation form was

$$\text{AAC}_{ij}|p_j \sim \text{Binomial}(1, p_j),\ E(\text{AAC}_{ij}|p_j) = p_j = \mu_j,$$

$$\text{Var}(\text{AAC}_{ij}|p_j) = p_j(1 - p_j),$$

$$\text{Level 1 link function: } \eta_j = \log\left(\frac{\mu_j}{1 - \mu_j}\right),$$

$$\text{Level 1 structural model: } \eta_j = b_{0j}, \text{ and}$$

$$\text{Level 2 model: } b_{0j} = \gamma_{00} + \gamma_{01}\text{Group}_j + u_{0j},\ u_{0j} \sim N(0, \tau^2).$$

As indicated in Table 4, a significant difference in the probability of AAC between the two groups was found only for hostility scores.

[4] We use the median instead of the mean because the median is robust to a skewed distribution due to the nonzero value of time interval. The median is based on the data provided by all participants. It seems unlikely that there is a difference in each individual's median of time intervals because the assessment time was randomly prompted. Nevertheless, we prefer to use median over all data points, across all participants, in order for the adjusted difference to be comparable across individuals.

[5] Although out preference is to analyze the detrended or residual scores for conceptual reasons, we also present results for the predetrended data in which variance due to systematic time factors was not partialled from each individual's scores (see Table 4).

[6] The value of ks in this sample were .49 for detrended negative affect, .50 for pre-detrended negative affect, .77 for detrended positive affect, .79 for predetrended positive affect, 51 for detrended hostility, .50 for predetrended hostility, .53 for detrended fear, .51 for pre-detrended fear, .69 for detrended sadness, and .75 for predetrended sadness.

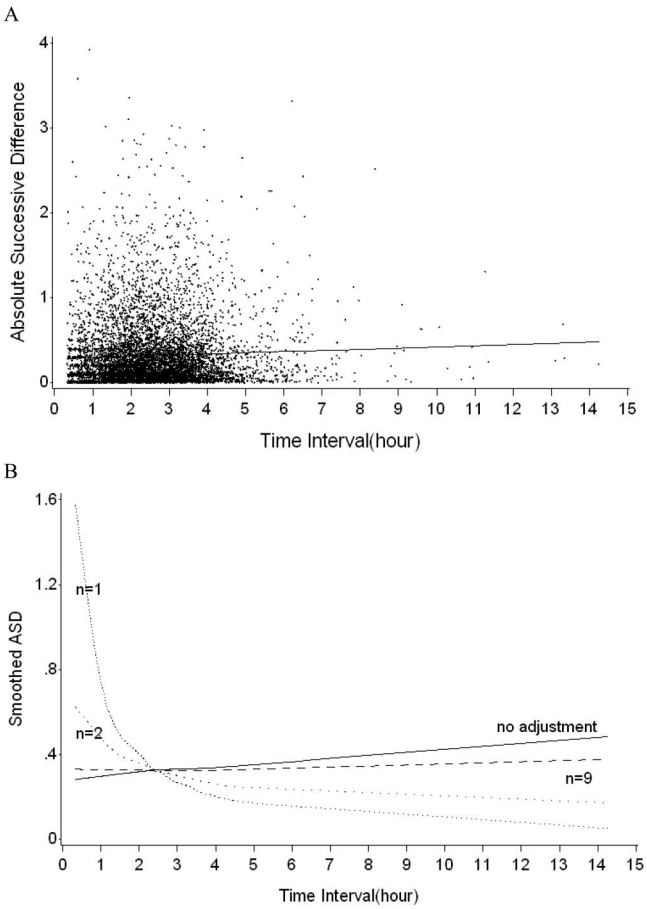

Figure 1. A: Absolute successive difference (ASD) of detrended negative affect across time interval between successive assessments. B: Smoothed regression line of unadjusted and adjusted absolute successive difference with different *n*s (see text for more details).

Thus far, the results we presented focus on group comparisons of various indices of affective variability and instability. However, it is important to remember that the calculation of these indices, as well as the method of detrending and adjustment for time interval, takes place at the level of the individual participant. To provide some sense of the pattern of raw scores and of the scores corrected for time factors and varying time assessment intervals, in Figures 2 and 3, we present two panels (A and B) representing NA scores from an MDD\DYS and BPD participant, respectively.

Figure 2 presents raw NA scores for each participant over the course of the study (solid lines). The two participants completed approximately the same number of assessments over the course of the study ($n = 157$ and $n = 156$, respectively), and their mean NA item score over all assessments was approximately equal ($M = 1.45$ and $M = 1.41$, respectively; dashed lines). In addition to presenting the raw data, Figure 2 presents the values of the squared successive differences (needles or bars above baseline). The BPD participant (Panel B) shows more variability in NA scores, including more and greater increases and decreases in NA, and this is reflected in the more frequent spikes seen in squared successive differences values compared with those for the MDD/DYS participant. The plus symbols (+) that appear above the solid line indicate that these successive increases in NA are greater than or equal to the 90th percentile of the total successive difference in NA scores in the entire sample. Specifically, the MDD/DYS participant has 18 such instances, whereas the BPD participant has 21.

Figure 3 presents the same indices for the same participants following a detrending of the raw data (i.e., removing time factors) and an adjustment of the size of the assessment intervals (in order to allow valid comparisons). Although the basic pattern of NA scores and of squared successive differences values remains the same, it is worth noting that the detrending and adjustment for time interval procedure results in half as many acute changes (+) for the MDD/DYS participant ($n = 9$), whereas the number for the BPD participant is not as drastically reduced ($n = 18$). These detrended and time interval-adjusted results suggest, for example, that many of the extreme increases in NA for this MDD/DYS participant can be accounted for by systematic time effects on NA scores as opposed to less predictable instability. An examination of the results of the regression analyses for this participant that assessed the influence of systematic time factors on the variance of NA scores indicates that there was a significant quadratic effect of time of day on this individual's NA scores.[7] In this case, higher NA scores were more typical in the middle of the day and lower both early and late in the day. This example highlights how a consideration of the influence of time effects on affect scores can provide clues as to the nature of some predictable affective variability and also allows a clearer focus on the variability and instability that may be triggered by unsystematic and less predictable events or encounters.

Discussion

The characterization and assessment of dynamic processes in psychopathology have proven challenging. In this study, we focused on characterizing affective instability as it occurs in the patient's natural environment. Although there is widespread agreement that affective instability is a core feature of BPD, few empirical studies have directly assessed features of affective instability in a natural context. Ours is the first study to use EMA via EDs to compare affective features of BPD with those seen in MDD/DYS. With the use of EMA methods to compare fluctuations in mood states among individuals with BPD and MDD/DYS, the present study addresses several important limitations in the previous literature on BPD affective instability. First, we used a psychiatric control group, the members of whom are characterized by elevated levels of negative affect. With few exceptions (e.g., Cowdry et al., 1991; Stein, 1996), previous studies have compared patients with BPD to normal control participants on measures of affective variability. By comparing BPD patients with those from the MDD/DYS group on various measures of affective instability, we were able to demonstrate that affective instability does appear to distinguish

[7] Regression results for both participants are available from Timothy J. Trull.

Table 4
Estimates of Fixed and Random Effects for Adjusted Squared Successive Difference (ASSD) Scores, With Log Link and Gamma Error Distribution, and Adjusted Acute Change (AAC) Scores, With Log Link and Binomial Error Distribution

	ASSD				AAC			
	Detrended		Predetrended		Detrended		Predetrended	
Variable	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
				Negative Affect				
Intercept	−2.01**	0.20	−1.98**	0.20	−2.57**	0.19	−2.59**	0.20
Group (MDD = 0, BPD = 1)	0.43	0.27	0.43	0.27	0.25	0.25	0.23	0.26
Level 2 variance	0.99		0.99		0.76		0.87	
Level 1 variance	6.16		6.43		0.92		0.91	
				Positive Affect				
Intercept	−1.21**	0.18	−1.14**	0.19	−2.59**	0.19	−2.53**	0.18
Group (MDD = 0, BPD = 1)	0.26	0.24	0.27	0.25	0.30	0.25	0.23	0.24
Level 2 variance	0.85		0.86		0.80		0.74	
Level 1 variance	3.94		4.18		0.89		0.90	
				Hostility				
Intercept	−2.00**	0.24	−1.99**	0.24	−2.70**	0.19	−2.70**	0.19
Group (MDD = 0, BPD = 1)	0.67*	0.32	0.68*	0.32	0.47*	0.25	0.47*	0.24
Level 2 variance	1.48		1.47		0.77		0.75	
Level 10 variance	8.69		9.04		0.91		0.91	
				Fear				
Intercept	−2.12**	0.23	−2.10**	0.23	−2.65**	0.21	−2.66**	0.21
Group (MDD = 0, BPD = 1)	0.59*	0.31	0.59*	0.31	0.28	0.28	0.30	0.27
Level 2 variance	1.34		1.35		0.97		0.94	
Level 1 variance	7.40		7.81		0.89		0.89	
				Sadness				
Intercept	−1.55**	0.20	−1.53**	0.20	−2.62**	0.19	−2.67**	0.19
Group (MDD = 0, BPD = 1)	0.55*	0.27	0.54*	0.27	0.35	0.24	0.38	0.25
Level 2 variance	1.02		1.03		0.74		0.81	
Level 1 variance	5.90		6.32		0.91		0.91	

Note. p values for group difference are calculated as a one-tailed test due to the directional hypotheses that the BPD group will exhibit more instability. For the intercept, the p values are calculated as a two-tailed test. MDD = major depressive disorder; BPD = borderline personality disorder.
* $p < .05$. ** $p < .0001$.

these two near-neighbor diagnostic groups despite the relatively high mean levels of negative affect that characterize both groups. That is not to say that those with MDD/DYS show little variability in their affect or in increases in negative affect from occasion to occasion. Rather, it is the degree of variability and instability, the frequency of mood changes, and the amplitude of these changes that separates the two groups. These findings, based on detrended data, are consistent with the conceptualization of affective changes in BPD being rather abrupt, large in magnitude, and likely brought on by external (and thus less predictable) events.

Second, we were able to capture different aspects of affective instability more satisfactorily by examining statistical indices of variability, frequency of change, amplitude, and temporal dependency (Larsen, 1987). Many previous studies focused assessment of affective instability on the variability of mood ratings and used the variance (standard deviation) of ratings within patient to operationalize this construct. Although simple standard deviations of scores over time are easily calculated, such measures are open to the criticism that they do not reflect necessarily variability, per se, but instead may represent a tendency of BPD patients to undergo systematic change in affect over time. Further, it is possible that their changing responses over time are an artifact of prolonged assessment. Yet another possibility is that individuals with BPD show greater variability in affect in response to their environment and that observed differences in variability are due to such reactivity. For example, the finding that the two groups differed in terms of variance in negative affect accounted for by a linear time of day effect by weekend trend might reflect that BPD participants are exposed to coworkers during daylight hours during the week, but spend time with romantic partners in the evenings over the weekend.

However, although our detrending of the individual series may be helpful in ruling out competing explanations of intra-individual variability over time, the variability differences of the detrended individual series may be due to factors other than actual variation in affect over time. One alternative, in particular, merits comment: reliability of assessment. The larger amount of variation in BPD individuals over time could simply reflect the possibility that these individuals provide less reliable data over time. In other words, those with BPD may not necessarily be more labile in their affect but may simply report their affective states less informatively or with less precision

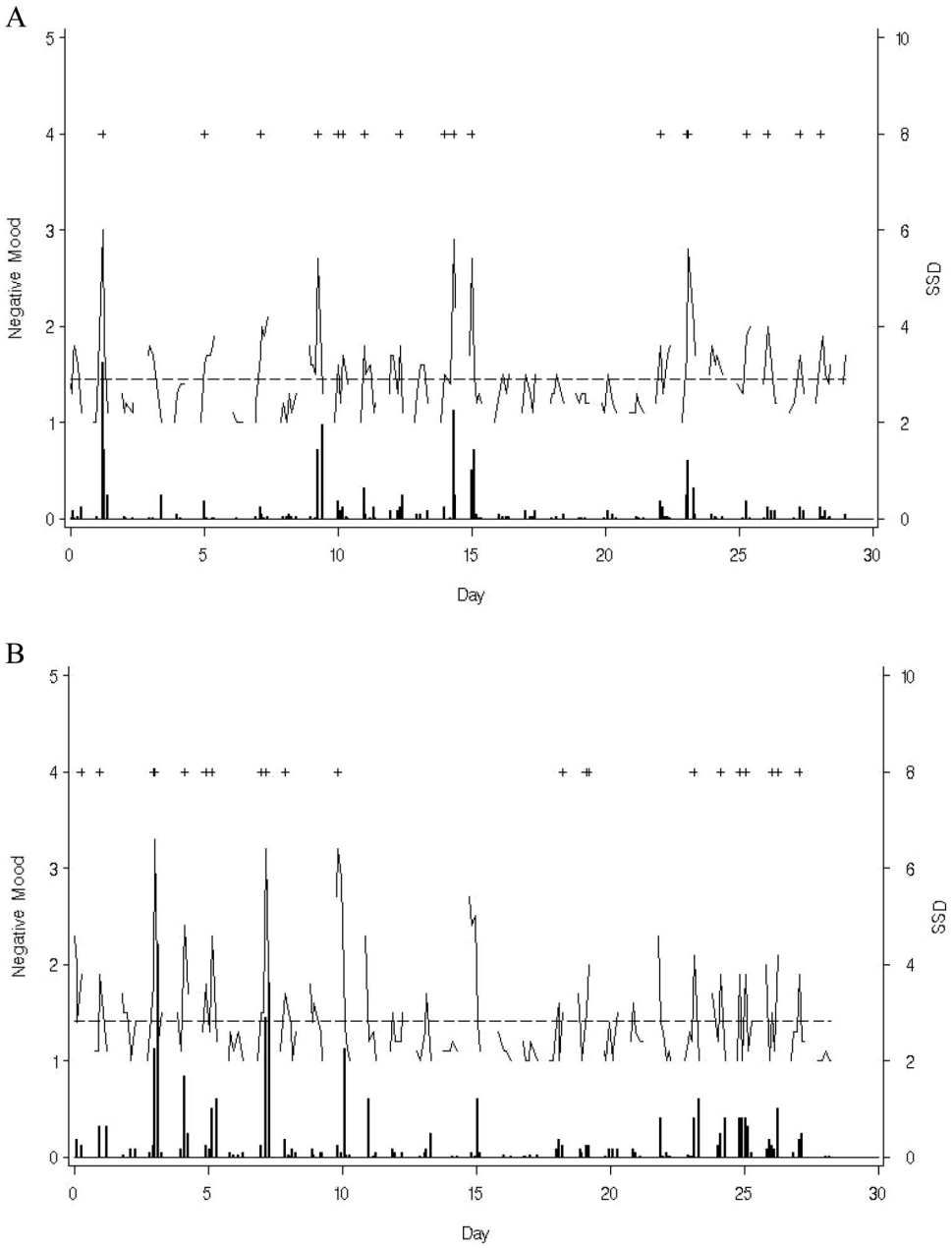

Figure 2. A: Raw Negative Affect (NA) scores for a study participant with current major depressive disorder (n = 157 measurement occasions), over the course of the study. B: Raw NA scores for a study participant with borderline personality disorder and affective instability (n = 156 measurement occasions), over the course of the study. Solid lines indicate raw NA scores (gaps between solid lines represent between-day assessments for which the successive differences are not considered in the analysis); dashed line indicates the mean NA score over all measurement occasions; bars above baseline indicate the squared successive difference (SSD) score between the current and the previous occasions; and plus symbols indicate those occasions in which the successive difference in NA was greater than or equal to the 90th percentile of the total successive difference in NA scores in the entire sample.

than other individuals. The methods proposed here can be extended to address alternative explanations like these (Jahng et al., in press), but some of these possibilities require the researcher to have multiple reports of affective state at a given measurement occasion and require the researcher to explicitly model time-bound relationships by modeling lagged as well as contemporaneous effects.

Our study also highlights a relatively new methodology for psychopathology research. ED methods, completed in patients' natural environments, offer a new approach to describing and

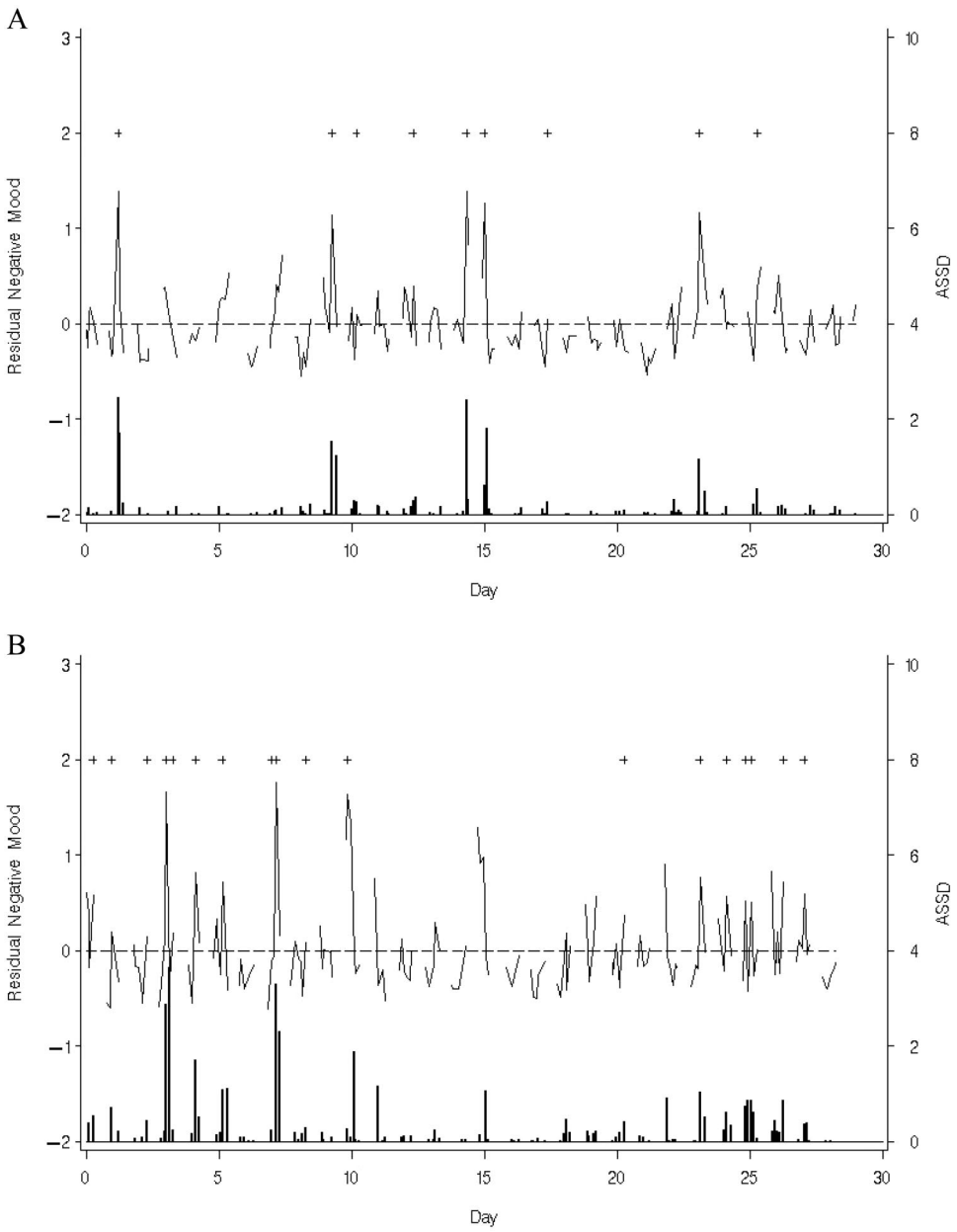

Figure 3. A: Detrended Negative Affect (NA) scores for the study participant with current major depressive disorder ($n = 157$ measurement occasions). B: Detrended NA scores for the study participant with borderline personality disorder and affective instability ($n = 156$ measurement occasions) referred to in Figure 1. Solid lines indicate residual NA scores after detrending (gaps between solid lines represent between-day assessments for which the successive differences are not considered in the analysis); dashed line indicates the mean residual NA score over all measurement occasions; bars above baseline indicate the adjusted squared successive difference score for residuals between the current and the previous occasions; and plus symbols indicate those occasions in which the adjusted successive difference in residual NA was greater than or equal to the 90th percentile of the adjusted successive difference in residual NA scores in the entire sample.

characterizing important dynamic processes and constructs in psychopathology, like affective instability in BPD. In the present case, this approach has several advantages over other traditional measurement approaches for affective instability or variability. Traditional approaches, like clinical interviews or questionnaires, require the individual to retrospect over an extended period of time (often several weeks or more) to describe mood patterns. This task is quite complex because it requires the individual to remember mood states accurately within each day over the targeted period of time, to remember

mood changes from this time period, and to decide the threshold for a mood change. The aggregation of this information is left up to the individual with very little (if any) guidance from the interview or questionnaire items. In contrast, ED approaches require that the individual reflect on and rate the degree to which a mood is present but do not require that the individual make decisions about when a mood change has occurred.

There are other advantages as well. Electronic diaries may be programmed to automate sophisticated experience sampling schemes, audibly prompting or beeping participants at random or quasi-random moments throughout each day. These prompts signal participants to complete a self-report assessment protocol programmed into the ED. Noncompliance, if it occurs, is recorded as such by the absence of a completed assessment at the programmed prompt time. Back-filled entries, if participants tried to make them, would be time-stamped and, thus, easily identified as questionable data. EDs may also improve data quality; compared with paper-and-pencil diaries, EDs are less vulnerable to illegible or out-of-range responses. Further, electronic records of the data entered into the EDs can be easily downloaded onto computers, making transcription and raw data entry unnecessary. Given the widespread use of cell phones and personal digital assistants (PDAs), respondents may find that completing ED assessments in their natural environment is less conspicuous than taking out a paper diary and completing it. Finally, when EDs are used, previous ratings are not visible and may be less likely to influence individuals' current ratings of moods or behaviors. For these and other reasons, we believe that EMA methods involving electronic diaries hold great promise for psychopathology research.

Two findings of this study deserve special comment. First, we found that in addition to negative affect scores, BPD patients exhibited more variability in positive affect scores than did those from the MDD/DYS group, but this was not true either in the case of instability or in the probability of acute change in positive affect scores. These latter two indices emphasize the amplitude or degree of change. Therefore, our findings suggest that what may distinguish the two diagnostic groups in terms of positive affect is the variation of positive mood but not frequent, extreme changes in positive affect.

Second, the last set of analyses revealed that hostility was the only negative affect in which the probability of acute change distinguished the groups. This finding is consistent with the clinical picture of BPD, especially the interpersonal elements and correlates of this diagnosis. Along these lines, Miller and Pilkonis (2006) found that trait affective instability (rather than their trait level of neuroticism and negative emotionality) showed the highest association with participants' level of aggression and interpersonal problems as reported by others, as well as with prospective reports of romantic problems. Those who more frequently have extreme spikes in their levels of hostility are more likely to have interpersonal conflict and to be seen by others as difficult. Although spikes of extreme sadness and fear also likely have some interpersonal sequelae, it seems the clinical picture of BPD patients' interpersonal conflicts and problems are often driven by feelings of hostility. On the other hand, the spikes in hostility might be the consequence of an interpersonal encounter or frustration. Future work is needed to shed light on whether affective changes (such as that in hostility) precede or follow certain experiences like interpersonal conflicts.

Several limitations of the present study should be acknowledged. First, because of the relatively small sample size and the number of different combinations of medications and substances used, we were unable to assess the effects of these variables on mood fluctuations. Further, we did not evaluate or control for the amount or type of psychological treatment participants received during the mood monitoring period. In order to assess potential effects on mood like these with sufficient statistical power, a larger number of participants will be needed. Second, in order to adequately test the hypothesis that environmental stimuli produce mood instability, future studies might incorporate both an event-contingent (enter data on the ED when experiencing certain environmental events, like rejection) and a signal-contingent (respond to random prompts) paradigm, to test the environmental reactivity model of affective instability more directly (Shiffman, 2007). There are advantages and limitations for each sampling strategy; in the case of characterizing the antecedents and consequences of affective instability, a combination of the two sampling strategies seems optimal.

In a related vein, it is important to highlight that we used the time frame "since the last prompt" for mood ratings rather than asking participants to rate their mood at that particular moment. Because other experiences and behaviors were queried and because this time frame was used, we did not want to require that the participants shift time frames for some questions because we thought this would be both burdensome and confusing (and, therefore lead to less reliable ratings). Admittedly, however, it is an empirical question as to whether we would have obtained similar evidence for affective instability with "right now" versus "since the last prompt" mood ratings.

Third, other studies with samples with different demographic characteristics (e.g., gender ratio, ethnicity) are needed to determine the generalizability of our results. Our participants were predominantly women and predominantly of White, non-Hispanic descent. The two groups we studied differed in terms of gender distribution and of previous psychiatric hospitalization. Concerning gender, it will be important to replicate our findings in other samples with greater representation of men, given that gender is associated with negative affect and neuroticism. Although our groups did not differ in terms of mean levels of negative affect, it is possible that the differences that we did observe in terms of the instability of certain negative affects may have been at least partially due to differences in the proportion of women in each group. Group differences for the instability of hostility scores were replicated when we performed supplemental analyses of women only. However, larger samples of patients representing both genders will be necessary to more definitively assess for gender effects in affective instability. Finally, future research is needed to determine the best quantitative approaches for analyzing times series data like these (Jahng et al., in press), to explore the overlap between EMA indices of affective instability and those from questionnaires and interviews, to explore the antecedents and consequences of affective instability, and to use these methods to characterize the affective profiles of those with near-neighbor diagnoses (but not a diagnosis of BPD) such as bipolar II disorder, depression, and anxiety disorder.

References

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Cowdry, R. W., Gardner, D. L., O'Leary, K. M., Leibenluft, E., & Rubinow, D. R. (1991). Mood variability: A study of four groups. *American Journal of Psychiatry, 148,* 1505–1511.

Csikszentmihaly, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease, 175,* 526–537.

Ebner-Priemer, U. W., Kuo, J., Kleindienst, N., Welch, S. S., Reisch, T., Reinhard, I., et al. (2007). State affective instability in borderline personality disorder assessed by ambulatory monitoring. *Psychological Medicine, 37,* 961–970.

Ebner-Priemer, U. W., Kuo, J., Welch, S. S., Thielgen, T., Witte, S., Bohus, M., & Linehan, M. (2006). A valence-dependent group-specific recall bias of retrospective self-reports. *Journal of Nervous and Mental Disease, 194,* 774–779.

Farmer, R. F., Nash, H. M., & Dance, D. (2004). Mood patterns and variations associated with personality disorder pathology. *Comprehensive Psychiatry, 45,* 289–303.

First, M. B., Spitzer, R. L., & Williams, J. B. W. (1995). *User's guide for the Structured Clinical Interview for DSM-IV Axis I Disorders.* New York: Biometrics Research Department, New York Psychiatric Institute.

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion, 14,* 577–606.

Gunderson, J. G. (2001). *Borderline personality disorder: A clinical guide.* Washington, DC: American Psychiatric Publishing.

Harvey, P. D., Greenberg, B. R., & Serper, M. R. (1989). The Affective Lability Scales: Development, reliability, and validity. *Journal of Clinical Psychology, 45,* 786–793.

Henry, C., Mitropoulou, V., New, A. S., Koenigsberg, H. W., Silverman, J., & Siever, L. J. (2001). Affective instability and impulsivity in borderline personality and bipolar II disorders: Similarities and differences. *Journal of Psychiatric Research, 35,* 307–312.

Hufford, M. R., Shiffman, S., Paty, J., & Stone, A. A. (2001). Ecological momentary assessment: Real-world, real-time measurement of subject experience. In J. Fahrenberg, & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 69–92). Seattle, WA: Hogrefe & Huber.

Jahng, S., Wood, P. K., & Trull, T. J. (in press). Analysis of affective instability in EMA: Indices using successive difference and group comparison via multilevel modeling. *Psychological Methods.*

Koeningsberg, H. W., Harvey, P. D., Mitropoulou, V., Schmeidler, J., Anotonia, S., Goodman, M., et al. (2002). Characterizing affective instability in borderline personality disorder. *American Journal of Psychiatry, 159,* 784–788.

Larsen, R. J. (1987). The stability of mood variability: A spectral analytic approach to daily mood assessments. *Journal of Personality and Social Psychology, 52,* 1195–1204.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford Press.

Links, P. S., Eynan, R., Heisel, M. J., Barr, A., Korzekwa, M., McMain, S., & Ball, J. S. (2007). Affective instability and suicidal ideation and behavior in patients with borderline personality disorder. *Journal of Personality Disorders, 21,* 72–86.

Miller, J. D., & Pilkonis, P. A. (2006). Neuroticism and affective instability: The same or different? *The American Journal of Psychiatry, 163,* 839–845.

Pfohl, B., Blum, N., & Zimmerman, M. (1994). *Structured Interview for DSM-IV Personality: SIDP-IV.* Iowa City, IA: Author.

Piasecki, T. M., Hufford, M. R., Solhan, M., & Trull, T. J. (2007). Assessing clients in their natural environments with electronic diaries: Rationale, benefits, limitations, and barriers. *Psychological Assessment, 19,* 25–43.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128,* 934–960.

Russell, J. J., Moskowitz, D. S., Zuroff, D. C., Sookman, D., & Paris, J. (2007). Stability and variability of affective experience and interpersonal behavior in borderline personality disorder. *Journal of Abnormal Psychology, 116,* 578–588.

Scollon, C., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies, 4,* 5–34.

Shiffman, S. (2007). Designing protocols for ecological momentary assessment. In A. S. Stone, S. Shiffman, A. A. Atienza, & L. Nebeling (Eds.), *The science of real-time data capture: Self-reports in health research* (pp. 27–53). New York: Oxford University Press.

Stein, K. F. (1996). Affect instability in adults with a borderline personality disorder. *Archives of Psychiatric Nursing, 10,* 32–40.

Stiglmayr, C., Grathwol, T., & Bohus, M. (2001). States of aversive tension in patients with borderline personality disorder: A controlled field study. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 135–141). Seattle, WA: Hogrefe & Huber.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment in behavioral medicine. *Annals of Behavioral Medicine, 16,* 199–202.

Stone, A. A., & Shiffman, S. (2002). Capturing momentary self-report data: A proposal for reporting guidelines. *Annals of Behavioral Medicine, 24,* 236–243.

Tennen, H., Affleck, G., & Armeli, S. (2005). Personality and daily experience revisited. *Journal of Personality, 73,* 1465–1484.

Watson, D., & Clark, L. A. (1999). *The PANAS-X: Manual for the Positive and Negative Affect Schedule—Expanded Form.* Retrieved February 23, 2008, from University of Iowa, Department of Psychology Web site: Http://www.psychology.uiowa.edu/Faculty/Watson/Watson.html

West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality, 59,* 313–338.

Woyshville, M. J., Lackamp, J. M., Eisengart, J. A., & Gilliland, J. A. M. (1999). On the meaning and measurement of affective instability: Clues from chaos theory. *Biological Psychiatry, 45,* 261–269.

Received June 7, 2007
Revision received February 25, 2008
Accepted March 10, 2008 ■