Psychological Medicine, Page 1 of 10.
doi:10.1017/S0033291706009706 Printed in the United Kingdom

State affective instability in borderline personality disorder assessed by ambulatory monitoring

ULRICH W. EBNER-PRIEMER[1]*, JANICE KUO[2], NIKOLAUS KLEINDIENST[1], STACY S. WELCH[2], THOMAS REISCH[3], IRIS REINHARD[4], KLAUS LIEB[5], MARSHA M. LINEHAN[2] AND MARTIN BOHUS[1]

[1] *Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health, University of Heidelberg, Germany*; [2] *Department of Psychology, University of Washington, Seattle WA, USA*; [3] *University Hospital of Social and Community Psychiatry, Bern, Switzerland*; [4] *Division of Biostatistics, Central Institute of Mental Health, University of Heidelberg, Germany*; [5] *Department of Psychiatry and Psychotherapy, University of Freiburg, Germany*

ABSTRACT

Background. Although affective instability is an essential criterion for borderline personality disorder (BPD), it has rarely been reported as an outcome criterion. To date, most of the studies assessing state affective instability in BPD using paper-pencil diaries did not find indications of this characteristic, whereas in others studies, the findings were conflicting. Furthermore, the pattern of instability that characterizes BPD has not yet been identified.

Method. We assessed the affective states of 50 female patients with BPD and 50 female healthy controls (HC) during 24 hours of their everyday life using electronic diaries.

Results. In contrast to previous paper-and-pencil diary studies, heightened affective instability for both emotional valence and distress was clearly exhibited in the BPD group but not in the HC group. Inconsistencies in previous papers can be explained by the methods used to calculate instability (see Appendix). In additional, we were able to identify a group-specific pattern of instability in the BPD group characterized by sudden large decreases from positive mood states. Furthermore, 48% of the declines from a very positive mood state in BPD were so large that they reached a negative mood state. This was the case in only 9% of the HC group, suggesting that BPD patients, on average, take less time to fluctuate from a very positive mood state to a negative mood state.

Conclusion. Future ambulatory monitoring studies will be useful in clarifying which events lead to the reported, sudden decrease in positive mood in BPD patients.

INTRODUCTION

According to DSM-IV (APA, 2000) and ICD-10 (WHO, 1992), affective instability is an essential criterion for borderline personality disorder (BPD). The World Health Organization classification (WHO, 1992) also lists BPD under the category of emotionally unstable personality disorders. The significance of this criterion is further emphasized in light of the evidence suggesting that affective instability is the BPD criterion most strongly associated with suicide attempts (Yen *et al.* 2004). However, despite the central importance of this criterion, the tremendous challenge of operationally defining affective instability (Goodman *et al.* 2003; MacKinnon & Pies, 2006) has prevented the field from establishing a standard state measure of affective instability. Consequently, this criterion has rarely been reported as an outcome in BPD treatment or pharmacotherapy studies (for an overview, see Lieb *et al.* 2004).

* Address for correspondence: Ulrich W. Ebner-Priemer, Ph.D., Central Institute of Mental Health, Department of Psychosomatic Medicine and Psychotherapy, Postfach 12 21 20, 68072 Mannheim, Germany.
(Email: ulrich.ebner-priemer@zi-mannheim.de)

To date, there are three published cross-sectional studies assessing state affective instability in BPD (Cowdry *et al.* 1991; Stein, 1996; Farmer *et al.* 2004) and one study investigating psychiatric patients with affective instability, most of whom were diagnosed with BPD (Woyshville *et al.* 1999). All four studies assessed affective instability using paper-and-pencil diaries with repeated measures over time and calculated instability by applying seven different methods. Two of these studies did not find evidence of heightened affective instability in BPD compared to healthy controls (Cowdry *et al.* 1991) or to other personality disorders (Farmer *et al.* 2004), whereas in the other two studies findings depended strongly on the methods used to calculate instability (Stein, 1996; Woyshville *et al.* 1999). In particular, Woyshville *et al.* (1999) found significantly *increased* affective instability in BPD patients using the mean squared successive difference (MSSD) method, but significantly *reduced* instability in the same data sample using fractal dimension calculations. Stein (1996) reported increased instability of unpleasant affect when using a variation of the mean absolute successive difference (MASD) method, but when using the autocorrelation analysis method in the same data set, significant differences between BPD patients and healthy subjects were not found.

We suggest that these different findings are a result of the different methods used to calculate instability. Larsen (1987) proposed three general components that comprise affective instability: amplitude, frequency and temporal dependency. Amplitude determines whether changes are large or small; frequency determines whether changes are rare or frequent; and temporal dependency accounts for the sequence in time (i.e. which self-report was made first, which was made second, which was made third, and so on). However, most of the methods used previously in BPD did not cover all components (see Appendix for detailed information). The amplitude component was particularly neglected in most of the methods, which means that these methods did not differentiate between time series with large or small changes. Consistent with Larsen (1987), we propose that the magnitude of change is an essential component of instability in general, as well as in BPD patients. Only MSSD and MASD cover all three components (see Appendix for detailed information). MSSD and MASD consistently revealed heightened affective instability in previous studies (Stein, 1996; Woyshville *et al.* 1999), whereas the other methods (such as within-subject standard deviation, random variability ratio, fractal dimension, power spectral density and autocorrelation analysis) did not find indications of heightened affective instability in BPD compared to healthy controls (HC) or other personality disorders.

Another major limitation of the current literature is that most of the studies do not explain the pattern of affect instability in any detail. Although having a general knowledge about whether heightened instability exists in BPD is of great importance, perhaps what is even more compelling is the pattern of instability that characterizes this group, such as a pattern of high fluctuations from extremely positive mood states to negative mood states.

To determine whether affective instability is heightened in BPD compared to psychologically HC and to identify the pattern that specifically characterizes instability in BPD, we repeatedly assessed valence and distress during 24 hours by an ambulatory monitoring approach (Fahrenberg & Myrtek, 2001; Ebner-Priemer, 2006), also known as ecological momentary assessment (Stone *et al.* 2002).

METHOD

Subjects

The data were collected across two sites. Fifty female patients meeting criteria for BPD and a comparison group of 50 female psychologically HC were selected as participants. Forty-two per cent of the patients were investigated at the University of Washington, Seattle ($n=21$ BPD, $n=21$ HC) and 58% of the patients were assessed at the University of Freiburg, Germany ($n=29$ BPD, $n=29$ HC). All patients met DSM-IV criteria for BPD, which was assessed by the appropriate section of the International Personality Disorder Examination (IPDE; Loranger, 1999). Axis I co-morbidity was assessed using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First *et al.* 1997). Participants with a history of schizophrenia, bipolar disorder or current alcohol/drug abuse were excluded. Trained psychologists (Freiburg) and Masters-level clinical assessors (Seattle) administered all of the diagnostic instruments. The Seattle HC were recruited through advertisements in Seattle and the Freiburg HC were selected randomly from the national resident register of the City of Freiburg. Exclusion criteria for the control group included the diagnosis of BPD, as assessed by the IPDE, any current or past Axis I disorder (SCID-I), current psychotherapy or use of current medication, which was self-reported, and other Axis II disorders (SCID-II; First *et al.* 1996). The last exclusion criterion only applies to the German sample. Twenty per cent of the patients were not taking psychotropic medications. Of the 80% of patients on medication, 65% were on antidepressants, 32% were on antipsychotics, and 30% were on hypnotics. The breakdown of co-morbid Axis I disorders in the BPD group was as follows: major depressive disorder (present: 36%), post-traumatic stress disorder (PTSD) (present: 60%), anxiety disorders without PTSD (present: 60%), eating disorders (past: 50%), and substance abuse (past: 60%). All patients were part of existing psychotherapy programs in dialectical behavior therapy (DBT; Linehan, 1993). The age of BPD patients (mean = 31·3 years, S.D. = 8·1) and HC (mean = 27·7 years, S.D. = 6·8) differed significantly ($t = -2{\cdot}44$, df = 98, $p = 0{\cdot}016$). However, the main outcome variable in this study, instability in affect or distress, was not correlated with age (MASD: valence $r = -0{\cdot}06$, $p = 0{\cdot}55$; distress $r = -0{\cdot}02$, $p = 0{\cdot}84$; MSSD: valence $r = -0{\cdot}06$, $p = 0{\cdot}58$; distress $r = -0{\cdot}13$, $p = 0{\cdot}21$). All subjects were paid for participating in the study and gave written informed consent. The study was approved by the respective ethical review committees from the Universities of Freiburg and the University of Washington.

Momentary assessment of emotions and distress

Over a 24-h period, each participant carried a minicomputer that emitted a prompting signal every 10 to 20 min. Following the prompt, the software program MONITOR (Fahrenberg *et al.* 2001) displayed questions regarding the participant's current emotions (e.g. 'How did you feel just before the beep?') with a list of possible answers: happy, interest, anxious, angry, sad, shame, disgust, emotion but can't name it, and no emotion. This list was derived from studies defining basic emotions (Linehan, 1993). After selecting the emotion, participants rated the intensity of their emotion on an 11-point Likert scale. They were also asked about the occurrence of any second emotion (using the same list as above, minus the first reported emotion) and its corresponding intensity (on an 11-point Likert scale). The frequency and intensity of the participant's current emotions were reported in a previous paper (Ebner-Priemer *et al.* in press). Finally, participants were also asked to rate their current intensity of distress. All responses were automatically time-stamped by the software program. Prompting was achieved by emitting three signals, each with the duration of 5 s with an inter-signal interval of 40 s. If the participant failed to respond at all (within 340 s), the trial was recorded as missed. All participants were trained in the use of the equipment and told how to turn the device off before going to sleep.

Data acquisition and preprocessing

To compute a single index of valence, we used the common method of multiplying the negative emotions by -1 and the positive emotions by $+1$ at each measurement occasion (Zelenski & Larsen, 2000; Eaton & Funder, 2001). For example, a rating of happy with an intensity of 3 would result in a valence score of $+3$, whereas a rating of anxious with an intensity of 5 would result a score of -5. 'No emotion' was given the valence score of zero and 'emotion, but can't name it' was assigned as missing data. We did not include valence ratings of the second emotion in our main analysis because of its low frequency occurrence (see Table 1) and we cross-checked our main findings with a valence algorithm using a mean value of both emotions (mean valence = mean of valence of emotion 1 and emotion 2).

To assess ecological validity (Fahrenberg & Myrtek, 2001), we asked participants during the post-monitoring interview, which occurred after the conclusion of the entire procedure, if the device and the monitoring procedure altered their behavior (e.g. whether they experienced higher attention to emotions or whether they found the device to be burdensome, etc.). Responses for burden and reactivity (Table 1) were low in both groups (five-point Likert scale: 0 = not at all, 1 = somewhat, 2 = considerably,

Table 1. *Description of data acquisition, burden and reactivity*

	BPD Mean (S.D.)	HC Mean (S.D.)	p*
Duration of data acquisition (h)	23·5 (0·88)	23·3 (1·45)	0·572
Period of shut-down during night (h)	9·65 (1·34)	8·44 (1·91)	<0·001
Number of prompts	52·3 (7·25)	55·9 (9·14)	0·015
Number of ratings made			
Distress	49·1 (8·4)	53·7 (9·0)	0·007
First emotion	38·2 (10·3)	42·8 (11·2)	0·035
Second emotion	17·4 (12·1)	9·5 (8·3)	<0·001
Number of calculated changes			
Distress	45·2 (9·6)	50·1 (9·9)	0·009
First emotion	28·7 (12·2)	33·4 (13·5)	0·077
Second emotion	10·0 (10·4)	3·8 (5·5)	<0·001
Reactivity due to monitoring procedure			
Changed behavior?	0·42 (0·65)	0·49 (0·70)	0·609
Higher attention to emotions?	1·49 (1·42)	1·19 (1·18)	0·256
Higher attention to physical symptoms?	1·11 (1·14)	0·88 (0·98)	0·296
Burden due to monitoring procedure			
Self-reports too frequent?	1·22 (1·11)	1·02 (1·14)	0·381
Burden of device?	1·76 (1·00)	1·55 (0·94)	0·286
Unpleasantness of self-reports?	1·18 (1·06)	0·61 (0·76)	0·003

BPD, Borderline personality disorder; HC, healthy controls; S.D., standard deviation.
* Wilcoxon test.

3 = mostly, 4 = very much). There is one single group difference (Unpleasantness of self-reports), although the significantly higher ratings of the BPD patients were rated fairly low on the Likert scale.

As indicated in Table 1, the duration of the data acquisition did not differ between the two groups, although the BPD patients did shut down the minicomputer for a longer period at night. Consequently, their number of prompts was lower than that of the HC group. Furthermore, the 'number of ratings made' was lower than the 'number of prompts'. This was caused by (1) requests being ignored, (2) the item 'emotion, but can't name it' not having a corresponding valence score, and (3) the occurrence of just one emotion, which leads to missing values for 'valence emotion 2'. Finally, the 'number of calculated changes' was lower than the 'number of ratings made' because two consecutive ratings were necessary to calculate a change.

Data analysis

If the independent samples were in line with the assumption of normality, two-sample t tests were conducted to examine the interval data. In case of heteroscedasticy (tested by F tests), approximate t statistics were calculated according to Satterthwaite's approximation. For independent interval data, which were not normally distributed, Wilcoxon tests were used. Statistical analysis was performed using SAS version 8.2 (SAS Institute Inc, Cary, NC, USA). A p value of 0·05 (two-sided) was considered statistically significant.

RESULTS

Heightened affective instability in BPD

The BPD group exhibited higher instability values of valence and distress compared with the HC group [MSSD for valence BPD = 25·5 (22·9), HC = 13·8 (11·4), $t = 3{\cdot}23$, df = 72, $p = 0{\cdot}002$; MSSD for distress BPD = 4·3 (3·4), HC = 1·9 (1·9), $t = 4{\cdot}31$, df = 77, $p < 0{\cdot}0001$; MASD for valence BPD = 3·2 (1·8), HC = 2·4 (1·1), $t = 2{\cdot}49$, df = 82, $p = 0{\cdot}0147$; MASD for distress BPD = 1·4 (0·6), HC = 0·7 (0·5), $t = 5{\cdot}65$, df = 95, $p < 0{\cdot}0001$]. To demonstrate that these findings were not caused arbitrarily by our analysis strategy, we re-examined the data using alternative strategies. First, we double-checked the group differences using non-parametric statistics. This did not change our original results, as the non-parametric statistics also

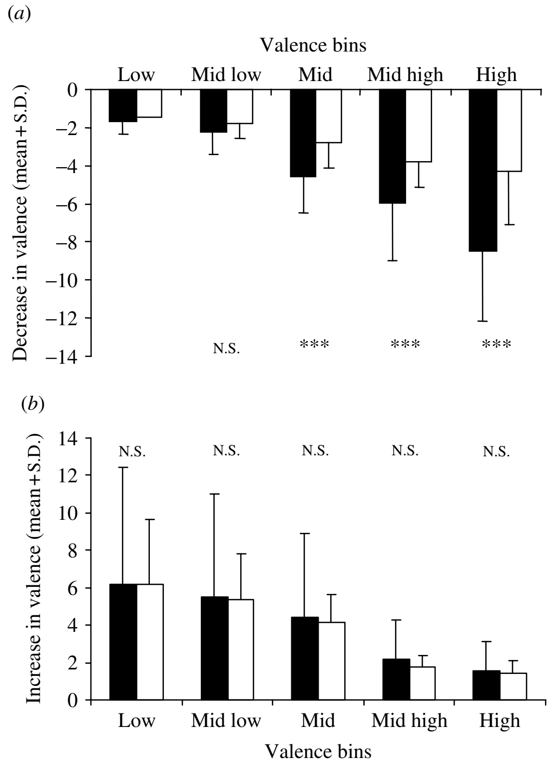

FIG. 1. Valence changes in relation to momentary valence: (*a*) decreases and (*b*) increases in valence in relation to each previous corresponding valence rating. *** $p \leqslant 0{\cdot}001$; N.S., no significant difference. ■, Borderline personality disorder; □, healthy controls.

suggested that the BPD group exhibited heightened instability (MSSD: Wilcoxon: valence $p = 0{\cdot}0133$, distress $p < 0{\cdot}0001$; MASD: Wilcoxon: valence $p = 0{\cdot}0490$, distress $p < 0{\cdot}0001$). Second, we used another approach, 'mean valence', to compute a single index of valence, as described earlier. Again, this method led to the same result, indicating more instability in BPD (MSSD: $t = 2{\cdot}38$, df = 78, $p = 0{\cdot}0197$).

The pattern of affective instability that characterizes BPD

To identify the pattern that characterizes instability in BPD patients, we decomposed the time series into point-by-point changes ($t_{i+1} - t_i$), which resulted in multiple decreases and increases in valence for every subject. Because of floor and ceiling effects, the magnitude of the increase and decrease in scores was heavily dependent on momentary ratings. For example, during times of very high valence (i.e. a very positive mood), a further increase in valence was not possible, and similarly, during periods of very low valence (i.e. a very negative mood), a large decrease was not possible. We therefore aggregated increases and decreases (successive differences) according to their momentary valence. Because low momentary valence (negative mood) was rare in the HC sample and high momentary valence (positive mood) was rare in the BPD sample, we aggregated momentary valence ratings into five momentary valence bins: low valence [−10, −7], mid-low valence [−6, −3], mid-valance [−2, +2], mid-high valence [+3, +6] and high valence [+7, +10]. That is, all increases starting at a momentary valence of −10, −9, −8, −7 were aggregated into the low valence bin; all increases starting at a momentary valence of −6, −5, −4, −3 were aggregated into the mid-low valence bin, etc.

Fig. 1(*a*) depicts the average decreases in valence in relation to the aggregated momentary valence. For example, if a patient changed from high valence (+9) at time point *t* to medium valence (valence = 0) at time point *t* + 1, this decrease in valence would be illustrated in Fig. 1(*a*) (decrease in valence) in the right black bar (high valence bin). This decrease (−9) would in fact be the same as the mean decrease in this bin for the patient group. Because of the floor and ceiling effects mentioned earlier, decreases in valence were smaller in low valence bins and larger in high valence bins. When comparing the two groups, BPD patients showed significantly higher decreases in the high valence bins compared to HC [high valence bin: $t(54) = 4{\cdot}84$, $p < 0{\cdot}0001$; mid-high valence bin: $t(53{\cdot}6) = 4{\cdot}19$, $p = 0{\cdot}0001$; mid-valence bin: $t(63{\cdot}4) = 4{\cdot}38$, $p < 0{\cdot}0001$; mid-low valence bin: $t(32{\cdot}2) = 1{\cdot}35$, $p = 0{\cdot}186$; low valence bin: too few subjects for comparison]. These results suggest that a positive mood is less stable in BPD compared with HC. These results were confirmed using non-parametric statistics (high valence bin: $p < 0{\cdot}0001$; mid-high valence bin: $p = 0{\cdot}0006$; mid-valence bin: $p < 0{\cdot}0001$; mid-low valence bin: $p = 0{\cdot}526$; low valence bin: too few subjects for comparison).

We further examined the large decreases from the high valence bin in the BPD group by subdividing all decreases from the high valence bin (valence $\geqslant 7$) into those that remained in a positive valence state (valence $\geqslant 0$) and those that decreased into a negative valence state (valence <0). For the HC group, only 9% of the declines from a high valence bin (positive mood) were so large that they reached a negative valence state (valence <0), whereas 91% of the declines remained in a positive valence state (valence $\geqslant 0$). By contrast, for the BPD group, 48% of the declines from a very high valence bin (positive mood) were so large that they reached a negative valence state (valence <0), whereas 52% of the declines remained in a positive valence state (valence $\geqslant 0$).

Figure 1(*b*) depicts the average increases in valence in relation to the aggregated momentary valence. Because of the floor and ceiling effects, mean increases in valence were larger in low valence bins and smaller in high valence bins for both groups. No group differences were demonstrated [low valence bin: $t(40)=0{\cdot}03$, $p=0{\cdot}973$; mid-low valence bin: $t(78{\cdot}7)=0{\cdot}22$, $p=0{\cdot}825$; mid-valence bin: $t(68)=0{\cdot}76$, $p=0{\cdot}448$; mid-high valence bin: $t(37)=1{\cdot}62$, $p=0{\cdot}114$; high valence bin: $t(27)=0{\cdot}62$, $p=0{\cdot}541$]. Once again, we reanalyzed the data using non-parametric statistics, and the results remained the same (low: $p=0{\cdot}773$, mid-low: $p=0{\cdot}769$, mid: $p=0{\cdot}503$, mid-high: $p=0{\cdot}158$, high: $p=0{\cdot}333$).

DISCUSSION

The hypothesis that BPD patients exhibit heightened affect instability was supported by our data, which indicated that the BPD group had higher instability values for both valence and distress compared with the HC. The demonstrated heightened affect instability is in accordance with the WHO definition of BPD as an 'emotionally unstable personality disorder' and also supports previous findings reported by Koenigsberg *et al.* (2002) and Henry *et al.* (2001), who retrospectively assessed instability using the affective lability scale (Harvey *et al.* 1989). It should be noted, however, that assessing symptoms retrospectively with questionnaires can result in exaggeration, as has been documented several times in patient samples (Margraf *et al.* 1987; Herman & Koran, 1998; Gendreau *et al.* 2003; Stone *et al.* 2004; Ebner-Priemer *et al.* 2006). Stone *et al.* (2004) also demonstrated that the retrospective assessment of change is much more problematic. Similarly, Larsen (1987) posits that 'to study variability, the researcher must, by definition, observe the object over time' (p. 1195).

However, previous studies that used paper-and-pencil diaries to assess the patients over time either did not to find heightened instability in BPD or reported inconsistent findings (Cowdry *et al.* 1991; Stein, 1996; Woyshville *et al.* 1999; Farmer *et al.* 2004). These discrepancies may be explained by the mathematical methods used to calculate instability (random variability ratio, autocorrelation analysis, fractal dimension, percentage power bins of the power spectral density), which do not cover the amplitude component. Therefore, these methods cannot determine whether the fluctuations in mood are large or minimal. Consistent with Larsen (1987), we propose that the magnitude of change is an essential component of instability in general, as well as in BPD patients. Analyses of previous studies (Stein, 1996; Woyshville *et al.* 1999) using our proposed methods (MASD, MSSD) to calculate instability revealed heightened instability in BPD as well.

Although paper-and-pencil diaries offer several benefits, including the assessment of the individual over time in their everyday environment, there are also disadvantages, such as the problems that arise with issues of compliance. For example, in a paper-and-pencil diary study, Stone *et al.* (2002) found that only 11% of reports were complete and consistent with the time schedule. Electronic diaries with time stamps circumvent these problems, and are therefore considered the gold standard in characterizing daily life experience (Kahneman *et al.* 2004).

Furthermore, we were able to demonstrate that affective instability in BPD is characterized by sudden large decreases in positive mood states. Specifically, BPD patients exhibited significantly higher decreases in the high valence bins compared to HC. Indeed, 48% of the declines from a high valence bin in BPD were so large that they reached a negative valence state. This was the case in only 9% of the HC group, suggesting that BPD patients, on average, take less time to fluctuate from a high valence state to a negative valence state. Such sudden, large decreases from a very positive mood displayed by the BPD group are consistent with our clinical impressions of this population and with current theories of BPD, including the biosocial theory of Linehan (1993). This pattern has not yet been reported, which is not surprising given that a very high sampling frequency is needed to capture such rapid affective changes. This was made possible by the ambulatory monitoring method used here. However, these results need to be replicated to determine if the chosen analytic approach favored one pattern of results over another and to further examine if the identified pattern was specific to our sample.

The limitations of this study should be noted. First, while this study used a categorical approach to assess emotions, we acknowledge that applying a dimensional approach might have led to different findings. However, the distress scale was assessed using a Likert scale and analysis of the distress data still suggested that BPD individuals displayed heightened instability. Second, we only assessed emotional experience and did not ask participants to report emotionally relevant events or daily life stressors during the monitoring. Therefore, the events that led to the enormous decreases in valence in the BPD group could not be identified and also could not be compared between groups. Clearly, this should be investigated in future studies. Third, as this was a female sample, the generalizability of the findings is constrained. However, having solely a female sample also reduced heterogeneity, which may be useful given the literature on sex differences on emotion. Furthermore, the multi-site nature of the study perhaps helped to increase generalization. Fourth, age differed significantly between groups, but instability in affect and distress were not correlated with age. Fifth, the possibility exists that a 24-h assessment may be too short to capture a reliable picture of instability in BPD patients. Although this does not pose a problem for group comparisons, it may be problematic when examining single cases. Although using a longer assessment period would be desirable, it would also enhance participant burden. A similar concern is whether displaying questions at an average of every 15 min substantially affects the participants' typical, everyday life. However, our post-monitoring interview revealed that, for both BPD patients and HC, the assessment procedure caused minimal unpleasantness and minimal alteration in their attention to emotions. Sixth, DBT includes psychoeducation about affective instability, which in of itself could sensitize the patients to their mood states and may contribute to an artificial group difference. We therefore compared patients before the start of DBT (the German group) with patients who were examined in the middle of ongoing DBT (the US group) and analyses did not show any significant differences for affective instability between the two groups (data available upon request). Therefore, our data do not suggest that sensitization to mood states occurred in response to treatment. Nevertheless, caution is warranted because possible effects of cohort (nationality) and duration of treatment participation (before and in the middle respectively) are mixed. Seventh, whereas it is an empirically valuable approach to randomly select participants from the city register, the benefits of this approach might be hampered by a high rate of refusal. Typically, the rate of refusal is above 60% in our studies. Therefore, we cannot exclude the possibility that we included an atypically compliant group. However, we propose that a random selection from the city register is still beneficial compared to selection from advertisements that recruit own staff or students from psychology departments. Finally, it should be noted that there are also disadvantages of the MASD and MSSD methods, particularly the parametric approach they use. Further research is necessary to develop new non-parametric methods to calculate instability. We also agree with others (Schwartz & Stone, 1998; Bolger *et al.* 2003) that, in general, ambulatory monitoring data should be analyzed with multi-level models. However, to our knowledge, multi-level modeling of instability has not yet been conducted, and even authors recommending the use of multi-level models in ambulatory monitoring use 'single level' analyses for calculating instability (Stone *et al.* 2005).

Heightened affect instability was clearly supported by our data, in contrast to previous paper-and-pencil diary studies, and a specific pattern of instability with sudden large decreases from positive mood states was identified. Future ambulatory monitoring studies will be useful in clarifying which events lead to the reported, sudden decrease in mood in BPD patients and, furthermore, may be an effective assessment tool to evaluate treatment response.

APPENDIX

Identifying appropriate mathematical methods to calculate instability

To clarify which of the seven previously used methods in BPD studies addressed the three general components in affective instability (amplitude, frequency and temporal dependency; Larsen, 1987), we selected an original time series from a BPD patient and modified the components of instability. Figure A 1(*a*) depicts the original time series (black line), a time series with a factor 2 reduced amplitude (gray line), and the center (dotted line). Figure A 1(*b*) describes the time series with a manipulated frequency, where the frequency of the gray line (manipulated) is half of that of the black line (original). Accordingly, only half of the original data series can be plotted, as indicated by the dotted line. Figure A 1(*c*) displays a redistribution of the time points in the time series. Both time series (original and manipulated) contain the same data but the temporal order is different.

We calculated instability values using all seven previously used methods for the time series (one original, three manipulated) using SAS (SAS Institute Inc., Cary, NC, USA). For the random variability ratio method, we applied the formula cited by Cowdry *et al.* (1991). For the fractal dimension method, we used a QuickBASIC algorithm published by Sevcik (1998) and translated the procedure into a SAS algorithm. We verified our findings for the fractal dimension method using the original algorithm from Woyshville *et al.* (1999), which was transformed into an excel program.

If a method covers the amplitude, the frequency or the temporal dependency component of the manipulated time series [Fig. A 1(*a*–*c*)] should reveal lower instability values than the original time series. Most of the methods did not address all three components (Table A 1). As illustrated, the random variability ratio, the autocorrelation analysis, the fractal dimension, and the percentage power bins of the power spectral density methods did not cover the amplitude component. A reduction in amplitude

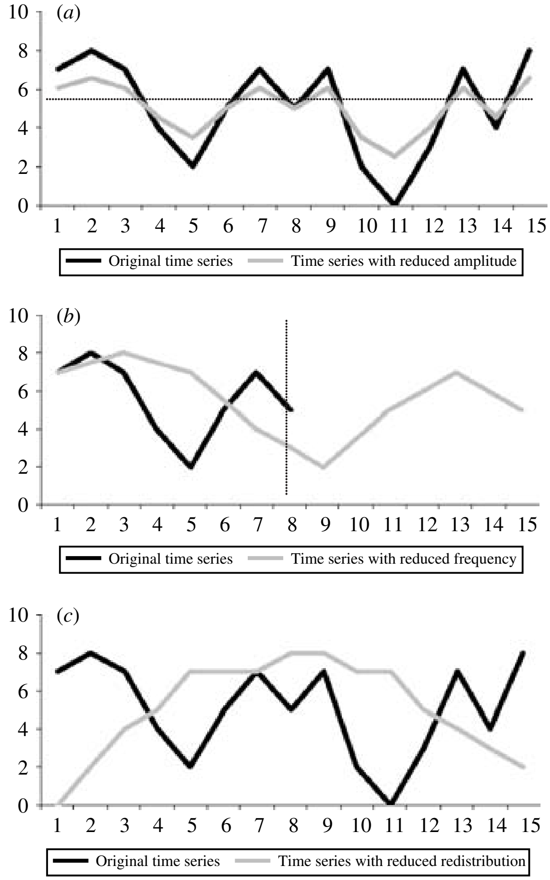

FIG. A 1. Original and manipulated time series with regard to amplitude, frequency and temporal dependency.

did not correspond with reduced instability values in these methods, suggesting that they cannot differentiate between time series with large or small changes. All the methods covered frequency. The standard deviation and the total power of power spectral density methods did not cover temporal dependency. As seen in Table A1, 'redistribution of time points' did not correspond with reduced instability values in these methods. Although the absolute power bins of the power spectral density method covered temporal dependency, the redistribution of time points led to complex results; absolute power in the mid- and high-frequency bins decreased, whereas absolute power in the low-frequency bin increased. That is, low-frequency instability was enhanced and high-frequency instability was reduced by the redistribution of

Table A 1. *Instability values of original and manipulated time series*

	Original	Reduced amplitude[a]	Reduced frequency[a]	Redistribution[a]
Standard deviation	2·49	1·25	1·82	**2·49**
MASD	2·64	1·32	1·00	1·00
MSSD	8·21	2·05	1·14	1·57
Random variability ratio	0·94	**0·94**	0·49	1·36
Autocorrelation[b]	0·27	**0·27**	0·80	0·67
Fractal dimension				
D	1·47	**1·47**	1·28	1·24
$\mathbf{D}_{corr}$	1·68	**1·68**	1·49	1·44
Power spectral density				
Total	6·92	1·73	3·71	**6·92**
Absolute power bin				
Low	4·01	1·00	3·47	**6·61**
Mid	2·37	0·59	0·20	**0·26**
High	0·54	0·14	0·04	**0·05**
Percentage power bin				
Low	57·9	**57·9**	93·4	95·5
Mid	34·3	**34·3**	5·5	3·8
High	7·8	**7·8**	1·1	0·7

MASD, Mean absolute successive difference; MSSD, mean squared successive difference; *D*, fractal dimension score; D_{corr}, fractal dimension score corrected for small *n*.

Figures in bold represent components of instability that are not fulfilled.

[a] Should lead to a reduced instability value compared with original.

[b] Higher autocorrelation values represent less instability.

time points. Using this approach, one would have to decide which frequencies are the typical frequencies that appropriately depict instability in BPD. Unfortunately, such information is still lacking. Therefore, only MASD and MSSD cover all components – amplitude, frequency and temporal dependency.

ACKNOWLEDGEMENTS

This research was supported by grants (Bo 1487/4-1) from the German Research Society (DFG), the Royalty Research Fund, University of Washington, and the Borderline Personality Disorder Research Foundation (BPDRF). We thank J. Fahrenberg, M. Myrtek, F. Foerster and P. Hüttner from the Psychophysiology Research Unit, University of Freiburg, Germany for methodological and technical support, B. Krumm from the Division of Biostatistics, Central Institute of Mental Health, Germany and S. Witte from the Department of Medical Biometry, University of Heidelberg, Germany for statistical support, M. Woyshville for extensive consultation and software on fractal dimension, and T. Thielgen, L. Detterer, S. Csontos and S. Platzbecker for data collection.

DECLARATION OF INTEREST

None.

REFERENCES

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn), *Text Revision* (*DSM-IV-TR*). American Psychiatric Association: Washington, DC.

Bolger, N., Davis, A. & Rafaeli, E. (2003). Diary methods: capturing life as it is lived. *Annual Review of Psychology* **54**, 579–616.

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.

Eaton, L. G. & Funder, D. C. (2001). Emotional experience in daily life: valence, variability, and rate of change. *Emotion* **1**, 413–421.

Ebner-Priemer, U. W. (2006). *Ambulatory Psychophysiological Monitoring – New Perspectives and Applications* [in German]. Peter Lang: Frankfurt.

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

Ebner-Priemer, U. W., Welch, S. S., Grossman, P., Reisch, T., Linehan, M. M. & Bohus, M. (in press). Psychophysiological ambulatory assessment of affective dysregulation in borderline personality disorder. *Psychiatry Research*.

Fahrenberg, J., Huettner, P. & Leonhart, R. (2001). MONITOR: acquisition of psychological data by a hand-held PC. In *Progress in Ambulatory Assessment: Computer-assisted Psychological and Psychophysiological Methods in Monitoring and Field Studies* (ed. J. Fahrenberg and M. Myrtek), pp. 93–112. Hogrefe & Huber: Seattle.

Fahrenberg, J. & Myrtek, M. (2001). *Progress in Ambulatory Assessment: Computer-assisted Psychological and Psychophysiological* *methods in Monitoring and Field Studies*. Hogrefe & Huber: Seattle.

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., Gibbon, M. & Williams, J. B. W. (1996). *User's Guide for the Structured Clinical Interview for DSM-IV Personality Disorders* (*SCID-II*). American Psychiatric Press: Washington, DC.

First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W. & Benjamin, L. S. (1997). *User's Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders* (*SCID-I*) – *Clinical Version*. American Psychiatric Press: Washington, DC.

Gendreau, M., Hufford, M. R. & Stone, A. A. (2003). Measuring clinical pain in chronic widespread pain: selected methodological issues. *Best Practice and Research*. *Clinical Rheumatology* **17**, 575–592.

Goodman, M., Weiss, D. S., Koenigsberg, H. W., Kotlyarevsky, V., New, A. S., Mitropoulou, V., Silverman, J. M., O'Flynn, K. & Siever, L. J. (2003). The role of childhood trauma in differences in affective instability in those with personality disorders. *CNS Spectrums* **8**, 763–770.

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. M. & 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.

Herman, S. & Koran, L. M. (1998). In vivo measurement of obsessive-compulsive disorder symptoms using palmtop computers. *Computers in Human Behavior* **14**, 449–462.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N. & Stone, A. A. (2004). A survey method for characterizing daily life experience: the day reconstruction method. *Science* **306**, 1776–1780.

Koenigsberg, H. W., Harvey, P. D., Mitropoulou, V., Schmeidler, J., New, A. S., Goodman, M., Silverman, J. M., Serby, M., Schopick, F. & Siever, L. J. (2002). Characterizing affective instability in borderline personality disorder. *American Journal of Psychiatry* **159**, 784–788.

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

Lieb, K., Zanarini, M. C., Schmahl, C. G., Linehan, M. M. & Bohus, M. (2004). Borderline personality disorder. *Lancet* **364**, 459–461.

Linehan, M. M. (1993). *Cognitive Behavioral Treatment of Borderline Personality Disorder*. Guilford Press: New York.

Loranger, A. W. (1999). *International Personality Disorder Examination* (*IPDE*)*: DSM-IV and ICD-10 modules*. Psychological Assessment Resources: Odessa.

MacKinnon, D. F. & Pies, R. (2006). Affective instability as rapid cycling: theoretical and clinical implications for borderline personality and bipolar spectrum disorders. *Bipolar Disorders* **8**, 1–14.

Margraf, J., Taylor, C. B., Ehlers, A., Roth, W. T. & Agras, W. S. (1987). Panic attacks in the natural environment. *Journal of Nervous and Mental Disease* **175**, 558–565.

Schwartz, J. E. & Stone, A. A. (1998). Strategies for analyzing ecological momentary assessment data. *Health Psychology* **17**, 6–16.

Sevcik, C. (1998). A procedure to estimate the fractal dimension of waveforms. *Complexity International* (www.complexity.org.au/ci/vol05/sevcik/sevcik.html). Accessed 17 May 2006.

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

Stone, A. A., Broderick, J. E., Shiffman, S. S. & Schwartz, J. E. (2004). Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. *Pain* **107**, 61–69.

Stone, A. A., Schwartz, J. E., Broderick, J. E. & Shiffman, S. S. (2005). Variability of momentary pain predicts recall of weekly pain: a consequence of the peak (or salience) memory heuristic. *Personality and Social Psychology Bulletin* **31**, 1340–1346.

Stone, A. A., Shiffman, S. S., Schwartz, J. E., Broderick, J. E. & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *British Medical Journal* **324**, 1193–1194.

WHO (1992). *The ICD-10 Classification of Mental and Behavioral Disorders. Clinical Descriptions and Diagnostic Guidelines*. World Health Organization: Geneva.

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

Yen, S., Shea, M. T., Sanislow, C. A., Grilo, C. M., Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Zanarini, M. C. & Morey, L. C. (2004). Borderline personality disorder criteria associated with prospectively observed suicidal behavior. *American Journal of Psychiatry* **161**, 1296–1298.

Zelenski, J. M. & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: a state and trait perspective from experience sampling data. *Journal of Research in Personality* **34**, 178–197.