# Social feedback processing in borderline personality disorder

**C. W. Korn[1,2,3]\*, L. La Rosée[1,4], H. R. Heekeren[1] and S. Roepke[5]**

[1] *Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany*
[2] *Department of Psychiatry, Psychotherapy, and Psychosomatics, University of Zurich, Zürich, Switzerland*
[3] *Neuroscience Center Zurich, University of Zurich, Switzerland*
[4] *Neuro-Cognitive Psychology Department, Ludwig-Maximilians-Universität München, Munich, Germany*
[5] *Department of Psychiatry, Charité-Universitätsmedizin Berlin, Campus Benjamin Franklin, Germany*

**Background.** Patients with borderline personality disorder (BPD) show negative and unstable self- and other-evaluations compared to healthy individuals. It is unclear, however, how they process self- and other-relevant social feedback. We have previously demonstrated a positive updating bias in healthy individuals: When receiving social feedback on character traits, healthy individuals integrate desirable more than undesirable feedback. Here, our aim was to test whether BPD patients exhibit a more negative pattern of social feedback processing.

**Method.** We employed a character trait task in which BPD patients interacted with four healthy participants in a real-life social interaction. Afterwards, all participants rated themselves and one other participant on 80 character traits before and after receiving feedback from their interaction partners. We compared how participants updated their ratings after receiving desirable and undesirable feedback. Our analyses included 22 BPD patients and 81 healthy controls.

**Results.** Healthy controls showed a positivity bias for self- and other-relevant feedback as previously demonstrated. Importantly, this pattern was altered in BPD patients: They integrated undesirable feedback for themselves to a greater degree than healthy controls did. Other-relevant feedback processing was unaltered in BPD patients.

**Conclusions.** Our study demonstrates an alteration in self-relevant feedback processing in BPD patients that might contribute to unstable and negative self-evaluations.

## Self-evaluations in borderline personality disorder (BPD)

BPD is characterized by negative self-evaluations (APA, 2013). Compared to non-clinical controls, BPD patients report lower self-esteem and they evaluate themselves less positively on BPD-specific and unspecific self-schema scales (Rüsch *et al.* 2007; Roepke *et al.* 2011). According to a structural analysis of social behaviour, BPD patients are characterized by a high degree of self-blame and self-neglect in combination with reduced self-love (Klein *et al.* 2001). Recently, we have demonstrated that the self-concept of BPD patients contains a greater proportion of negative attributes compared to healthy controls and depressed patients (Vater *et al.* 2015). Specifically, BPD patients rate negative self-aspects as more important than positive self-aspects.

## Other-evaluations in BPD

The definition of BPD highlights a pattern of unstable and intense interpersonal relationships, which alternate between extremes of idealization and devaluation (APA, 2013). Studies assessing how BPD patients evaluate other persons indicate a complex picture: Several studies found a negative evaluation bias of others in BPD, whereas some studies found a general bias for more extreme evaluations of others (Roepke *et al.* 2013). Compared to healthy controls, BPD patients judge others as more malevolent and dangerous in Rorschach responses (Westen, 1990) and self-report questionnaires (Pretzer, 1990; Arntz *et al.* 2004). Additionally, BPD patients judge persons in short film clips as more negative and less positive (Barnow *et al.* 2009). In another study, BPD patients displayed more extreme evaluations of others in film clips with BPD-specific themes, but not in film clips with unspecific themes (Veen & Arntz, 2000). In a manipulated therapeutic interaction, BPD patients evaluated the interaction partner in a more extreme way compared to healthy controls, independent of conversation content (Arntz & ten Haaf, 2012).

* Address for correspondence: Dr C. W. Korn, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany.

(Email: christoph.w.korn@gmail.com)

## Social interactions in BPD

While self- and other-evaluations of BPD patients have been subject to considerable research, only a few studies have investigated how BPD patients interact with others in social settings – possibly due to the difficulty of setting up these interactions. One line of research has focused on social exclusion in a virtual ball-tossing game. BPD patients feel more easily excluded compared to healthy controls, regardless of whether or not this is actually the case (Staebler *et al.* 2011; Renneberg *et al.* 2012). Another line of research used multi-player economic games to investigate how trust develops over multiple economic transactions (King-Casas *et al.* 2008; Unoka *et al.* 2009; Franzen *et al.* 2011). Initially, BPD patients reciprocate as much trust as healthy controls, but they fail to repair breaches of trust (King-Casas *et al.* 2008). Thus, some evidence from the laboratory supports deficient social interactions in BPD but these results remain largely silent on how altered information processing could be related to the integration of social feedback and its subsequent effects on negative self- and other-evaluations.

## Social feedback processing in healthy individuals

In contrast to BPD patients, healthy individuals show a variety of (self-related) positivity biases. They tend to rate themselves high on positive and low on negative character traits (Taylor & Brown, 1988; Leary, 2007; Alicke & Sedikides, 2009). Despite the impressive amount of literature on positivity biases, a central proposition of their early conceptualization has been relatively underexplored. Taylor and colleagues (Taylor & Brown, 1988; Taylor *et al.* 1989) posited that a series of 'cognitive filters' distort information processing towards the positive. We have recently addressed this lacuna by demonstrating that healthy individuals display a prominent positivity bias when receiving social feedback about character traits. They update their self- and other-evaluations more after receiving desirable feedback than after undesirable feedback (Korn *et al.* 2012, 2014*a*). This finding suggests that positive self- and other-evaluations in healthy persons could be developed and maintained by positively biased social feedback processing.

## Overview and hypotheses

Despite the well-described negative self- and other-evaluations of BPD patients, it has not been addressed how these evaluations *change* in response to incoming information and it is thus unclear how BPD patients develop and maintain their negative or extreme evaluations. Here, we adapted the *character trait task* of our

previous studies (Korn *et al.* 2012, 2014*a*). Our main hypothesis stated that BPD patients would show a more negative pattern than healthy controls when receiving self-relevant feedback. Such a more negative pattern could result from *reduced* integration of *desirable* feedback and/or *enhanced* integration of *undesirable* feedback.

Additionally, we tested the following hypotheses. First, BPD may show altered processing of other-related feedback. Second, feedback processing may be especially altered for traits that are of high relevance for BPD. Third, building on the literature summarized above, we expected BPD patients to entertain more negative self-and other views than healthy controls. Fourth, healthy participants may perceive BPD patients as more negative than they perceive other healthy participants. To address the last point, we included an additional group of healthy individuals who processed feedback about BPD patients.

## Method

### Participants

Demographic and clinical characteristics of the participants are presented in Tables 1 and 2. All participants were female.

Twenty-four participants with a diagnosis of BPD according to DSM-IV criteria (APA, 2000) were recruited through a ward specializing in personality disorders at the Department of Psychiatry, Charité-Universitätsmedizin Berlin. DSM-IV Axis II diagnoses were assessed using the Structured Clinical Interview (SCID-II; First *et al.* 1997) and DSM-IV Axis I diagnoses were assessed using the Mini-International Neuropsychiatric Interview (MINI; Sheehan *et al.* 1998). Exclusion criteria comprised schizophrenia, schizoaffective disorder, bipolar disorder, depressive disorder with psychotic features, anorexia, alcohol or drug abuse, and dependence in the last 6 months. Data were excluded from two BPD patients [one had an IQ score 2 standard deviations (s.d.) below the rest of the whole group and reported difficulties understanding the task, one reported during debriefing that she had known about the experimental manipulation from another participant]. Thus, the analysed sample included 22 BPD patients (mean age 28.0 years, s.d. = 7.1, range 18–39).

Healthy controls were recruited through online advertisement and flyers. A trained psychologist (L.L.R.) assessed healthy controls during a telephone interview, using a questionnaire that included the screening questions of the SCID-II and the MINI. We only invited persons who did not currently fulfil DSM-IV criteria, did not report a history of mental disorders, were not under the influence of psychotropic

**Table 1.** *Demographic and clinical characteristics of participants*

| Characteristic | BPD patients | Healthy controls group A (ConA) | Healthy controls group B (ConB) | Effects | $p$ values |
|---|---|---|---|---|---|
| $n$ | 22 | 42 | 39 | – | – |
| Age, years | 28.0 (7.1) | 27.2 (7.0) | 27.0 (6.2) | $F_{2,100} = 0.16$ | $p > 0.8$ |
| Family status: single, $n$ (%) | 12 (55) | 26 (62) | 20 (51) | $\chi^2_2 = 0.97$ | $p > 0.6$ |
| Education, years | 11.0 (2.2) | 11.0 (1.0) | 11.3 (0.9) | $F_{2,100} = 0.63$ | $p > 0.5$ |
| Verbal IQ (WST)[a] | 94.4 (7.8) | 98.6 (9.7) | 100.1 (10.6) | $F_{2,99} = 2.47$ | $p = 0.089$ |
| GSI SCL-90-R | 1.6 (0.5) | 0.3 (0.2) | 0.3 (0.3) | $F_{2,100} = 143.27$ | $p < 0.001$ |
| | | | | BPD $v$. A: $t_{24.6} = 11.9$ | $p < 0.001$ |
| | | | | BPD $v$. B: $t_{28.8} = 10.7$ | $p < 0.001$ |
| | | | | A $v$. B: $t_{67.9} = -1.5$ | $p > 0.1$ |
| BSL-95[b] | 2.0 (0.6) | – | – | – | – |
| BDI | 36.0 (10.7) | 4.4 (3.9) | 5.2 (4.7) | $F_{2,100} = 217.20$ | $p < 0.001$ |
| | | | | BPD $v$. A: $t_{24.0} = 13.4$ | $p < 0.001$ |
| | | | | BPD $v$. B: $t_{25.6} = 12.8$ | $p < 0.001$ |
| | | | | A $v$. B: $t_{79} = -0.8$ | $p > 0.4$ |
| Rosenberg self-esteem scale, sum score | 7.0 (4.1) | 25.7 (3.3) | 24.9 (4.0) | $F_{2,100} = 204.22$ | $p < 0.001$ |
| | | | | BPD $v$. A: $t_{62} = -19.7$ | $p < 0.001$ |
| | | | | BPD $v$. B: $t_{59} = -16.6$ | $p < 0.001$ |
| | | | | A $v$. B: $t_{79} = -1.0$ | $p > 0.3$ |
| | | | | A $v$. B: $t_{62.0} = -0.2$ | $p > 0.8$ |
| Perceived similarity to other[c] | 3.7 (1.4) | 3.6 (1.6) | 3.4 (1.4) | $F_{2,96} = 0.33$ | $p > 0.7$ |
| Perceived surprise about social feedback[c] | 5.3 (1.6) | 4.8 (1.3) | 4.7 (1.9) | $F_{2,96} = 2.80$ | $p > 0.3$ |
| Winning position in Monopoly game, median [minimum, maximum] | 3 [1, 5] | 3 [1, 5] | 3 [1, 5] | $H_2 = 0.37$ | $p > 0.8$ |

BPD, Borderline personality disorder; control group A (rated another control participant in other condition); control group B (rated BPD patient in other condition); Family status: single, comparison with logistic regression; WST, Wortschatztest, a verbal IQ test implemented in the HAWIE-R (German adaptation of the Wechsler Adult Intelligence Scale; Schmidt & Metzler, 1992); GSI, General Symptom Index (Franke, 1995; Staebler *et al.* 2011); SCL-90-R, Symptom Checklist-90 – Revised; BSL-95, Borderline Symptom List (Bohus *et al.* 2007); BDI; Beck Depression Inventory (Hautzinger *et al.* 1994); Rosenberg self-esteem scale (Rosenberg, 1965); Perceived similarity to other: participants' rating of how similar they perceived themselves to person they rated in other condition (from 1 *not at all similar* to 8 *very similar*); Perceived surprise about social feedback (from 1 *not at all surprised* to 8 *very surprised*); Winning position in Monopoly game (1 = winner, 5 = loser), comparison with Kruskal–Wallis test.

Data are given as mean (standard deviation) and were compared using 1 × 3 ANOVAs – unless otherwise specified.

[a] Data from one participant in ConA were not collected on the WST.

[b] Patients completed the BSL-95 during hospitalization. Data from two patients were not collected.

[c] Data from one patient, two participants in ConA, and one participant in ConB were not collected.

**Table 2.** *Clinical characteristics of BPD patients*

| Psychiatric hospitalization and co-morbidities | *n* (%) |
|---|---|
| Hospitalization | 6 (27) |
| Depressive disorder | 13 (59) |
| Post-traumatic stress disorder | 8 (36) |
| Any substance abuse/dependency | 8 (36) |
| Eating disorder | 11 (50) |
| Social phobia | 1 (5) |
| Panic disorder | 3 (14) |
| Obsessive-compulsive disorder PD | 1 (5) |
| Narcissistic PD | 1 (5) |
| Obsessive-compulsive PD | 2 (9) |
| Avoidant PD | 4 (18) |
| Histrionic PD | 1 (5) |
| Dependent PD | 2 (9) |
| Paranoid PD | 1 (5) |
| Medication, *n* (%) | |
| No medication | 1 (5) |
| Antidepressant | 11 (50) |
| Antipsychotic | 2 (9) |
| Antiepileptic | 2 (9) |
| Antidepressant and antipsychotic | 3 (14) |
| Antidepressant and neuroleptic | 1 (5) |
| Antidepressant and antiepileptic | 2 (9) |

BPD, Borderline personality disorder; PD, personality disorder.

medication, had no neurological disease, and who had no history of head trauma. In addition, we decided *a priori* to exclude healthy participants for whom Beck Depression Inventory (BDI; Beck *et al.* 1988) scores at the day of testing indicated moderate or severe depression. Data were excluded from four healthy controls (three scored at or above the cut-off of 19 in the BDI, one did not answer 88% of the trials). Hence, the analysed sample comprised a total of 81 healthy controls (mean age 27.1 years, S.D. = 6.6, range 18–45).

The study was approved by the Ethics Committee of the Charité-Universitätsmedizin Berlin (EA4/053/13). Participants gave written informed consent and were debriefed carefully. In accordance with the ethical guidelines, inpatients were not paid. All other participants were paid (€24).

### Task: part 1 – social interaction and first ratings

Fig. 1 gives an outline of the experimental layout, which followed our previous studies (Korn *et al.* 2012, 2014*a*). Participants interacted in groups of five consisting of one BPD patient and four healthy controls. To get to know each other, they played a table version of the popular board game Monopoly (Hasbro) for 45 min. We made sure that participants

did not know each other beforehand. Healthy controls did not know that one of their fellow participants had a psychiatric diagnosis and BPD patients were not told that the other four participants were healthy controls. To make sure that participants did not discover each other's diagnostic status, participants were only allowed to talk about topics related to the game and testing was not conducted at the clinic but at the Freie Universität Berlin.

Participants were told that the aim of the study was to gain insight into the process of how people get to know each other. They were informed that they were going to rate each other on a number of trait adjectives after the game and that these ratings were going to be shown anonymously to the respective fellow players. All participants were given name tags and we made sure that participants memorized all names. The winner of the game and all other positions in the game were determined. BPD patients and healthy controls did not differ in their winning positions (Table 1).

Afterwards, each player rated three out of four other players on 40 positive and 40 negative trait adjectives. We used the same adjectives as in our previous studies (Supplementary Table S2). Ratings were on a Likert scale from 1 (*this trait does not apply to the person at all*) to 8 (*this trait applies fully to the person*). These and all following ratings were recorded on PCs using the Matlab toolbox Cogent 2000 (http://www.vislad.ucl.ac.uk/cogent.php). Each person was rated in a separate block. During each trial, participants saw one of the 80 adjectives as well as the name of the person to rate. They had up to 10 s to respond. At the end of part 1, participants had not yet rated themselves and the fourth other player. To make participants believe that they would be receiving the veridical feedback, the respective files were distributed on all computers with a USB stick.

In nine of the total 24 groups, one of the invited healthy control participants did not arrive for the experiment and research assistants of similar age acted as participants during the social interaction (in one group two research assistants took part). Data from the research assistants were not analysed.

### Task: part 2 – receiving social feedback

In part 2, participants rated themselves (target condition: self) and the fourth participant (target condition: other), whom they had not rated in part 1. BPD patients rated themselves and one healthy control participant. Healthy controls were randomly assigned to two groups: ConA rated themselves and another healthy control. ConB rated themselves and the BPD patient. For our main hypothesis, BPD patients were compared with ConA. We compared ConA with ConB to address the supplementary question whether

healthy participants would perceive of BPD patients differently than other healthy controls. Importantly, participants received social feedback for themselves and the other person in the form of mean trait ratings. Participants believed that this feedback was based on the mean trait ratings of the interaction partners from part 1. Specifically, participants believed that feedback for the fourth other participant came from the remaining three interaction partners. However, in reality feedback was determined by an algorithm that reliably created a sufficient number of trials in which feedback was better (desirable) or worse (undesirable) with respect to participants' own ratings.

In each of the 160 trials, participants initially saw a cue whether this trial was about themselves or the other person (1 s). Then one of the 80 adjectives was presented and participants had to imagine how much this trait applied to themselves or to the other person (2 s). Thus, participants saw each of the 80 adjectives twice, once in the self- and once in the other condition. The question 'How much does this trait adjective apply to you/to this person' prompted participants to indicate their rating (maximum 10 s). After a fixation cross (750 ms), participants saw what they believed to be the mean rating of three participants from part 1 (2 s). This social feedback consisted of a pseudo-randomized number ranging from 1 to 8. The number was rounded to one decimal (since this number was supposed to be the mean of three integer ratings this decimal could be 0, 3, or 7). After a second fixation cross (750 ms), a new trial began. Trials for self and other were randomly intermixed. Within a group of five interacting participants, the patient always rated a control participant in the other condition. Two of the four control participants in a group, always rated another control participant (ConA) and the other two control participants always rated the BPD patient (ConB).

Directly afterwards, participants rated themselves and the fourth person again on the 80 traits. Thereby, we measured how much participants changed (i.e. updated) their self- and other-ratings after receiving social feedback. Participants rated themselves and the other person in two blocks (one for self and one for other), which were randomized for order. Participants had up to 6 s to respond for each trait.

### Additional tasks

After part 2, participants' memory for the feedback was tested. For all 80 traits participants had to type in the feedback they recollected (i.e. a number between 1 and 8 with one decimal such as 1). Participants recollected the feedback in two separate blocks (one for self and one for other), which were randomized for order. They had up to 12 s to respond.

To assess whether BPD patients perceived the adjectives differently, participants rated all 80 adjectives on social positivity on a scale from 1 (*not positive at all*) to 8 (*very positive*; Table 3).

### Questionnaires

To provide a detailed description of our sample, we collected the two most commonly used measures of symptom severity: General Symptom Index Symptom Checklist-90 – Revised (GSI SCL-90-R; Franke, 1995) and Borderline Symptom List (BSL-95; Bohus *et al.* 2007). Additionally, participants completed widely used measures for trait self-esteem (Rosenberg self-esteem scale; Rosenberg, 1965) and current mood state (Multidimensional Mood State Questionnaire; see Supplementary Table S1) (Steyer *et al.* 1997). The purpose for collecting all these measures was to explore whether they correlated with the expected effect of interest.

Participants provided two overall ratings: one rating indicating how similar they thought the person they had rated in the other condition was to them on a Likert scale from 1 (*not similar at all*) to 8 (*very similar*) and another rating indicating how surprised they were about the feedback they had received on a Likert scale from 1 (*not surprised at all*) to 8 (*very surprised*) (Table 1).

Following the character trait task, participants completed another independent task that will be reported elsewhere. The whole testing session lasted around 2.5 h.

### Task conditions and data analysis

See Fig. 1 for illustration. Unless otherwise indicated, we first compared all three groups using mixed-design ANOVAs including the three-level between-subject factor group (BPD/ConA/ConB). Provided that the relevant effects in the overall ANOVA were significant, we addressed specific hypotheses using follow-up ANOVAs or *t* tests. We specifically tested for differences between BPD patients and ConA to assess how the behaviour of BPD patients differed from controls. ConA was the relevant control group because, similar to BPD patients, they rated healthy participants in the other condition. ConB rated BPD patients in the other condition, which may create spillover effects even for the self-condition. To explore whether BPD patients were perceived differently from controls, we tested for differences between ConA and ConB.

Ratings for negative trait adjectives were reverse-coded. Specifically, ratings for negative traits were subtracted from 9 (e.g. if the original rating for unpleasant was 1 this number was transformed into 8 for the analyses).

For the analyses of update behaviour, trials were grouped according to feedback desirability. Feedback ratings were classified as 'desirable' when they were more
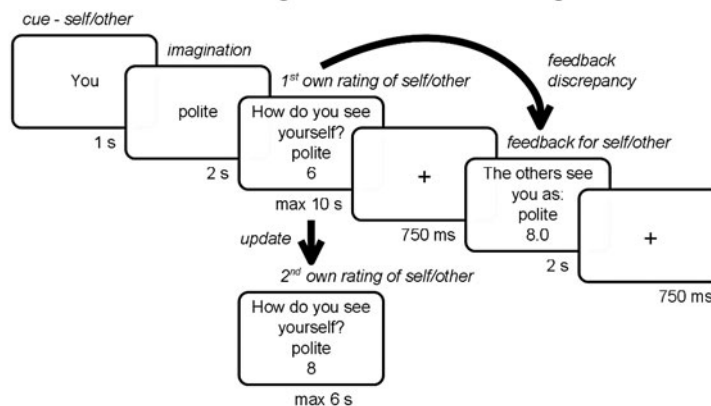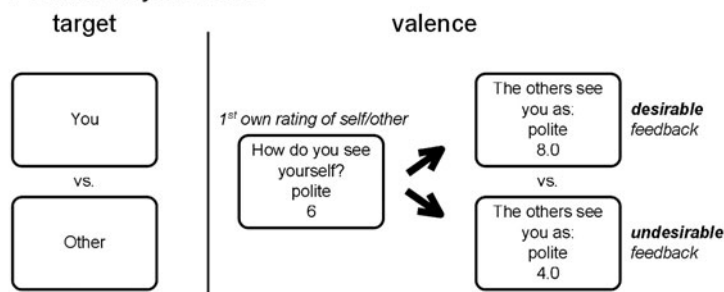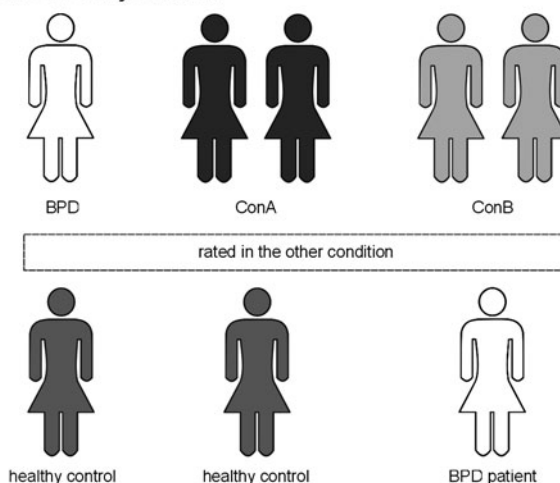
**Fig. 1.** For legend see next page.

positive than participants' initial rating. For positive (negative) adjectives, desirable feedback indicated that the feedback rating was numerically higher (lower) than the initial rating (e.g. positive traits: initial rating for polite was 6 and feedback was 8.0; e.g. negative traits: initial rating for aggressive was 3 and feedback was 1.0). Conversely, *un*desirable feedback ratings were more 'negative' than participants' own initial ratings.

As in our previous studies, the differences between participants' own first ratings and the feedback ratings they received, which we labelled feedback discrepancies, were manipulated using a random number generator.

$$\text{Feedback discrepancy} = \text{abs}(\text{feedback rating} - \text{first own rating}) \quad (1)$$

In an intermediate step, we calculated an update term for each participant and condition.

$$\text{Update} = \text{second own} - \text{first own rating} \quad (2)$$

Mean updates were scaled by the respective mean feedback.

$$\text{Relative absolute mean update} = \text{abs}(\text{mean update}/\text{mean feedback discrepancy}) \quad (3)$$

These relative absolute mean updates are thus percentages of how much participants adjusted their ratings toward the feedback and constitute the metric of interest analysed to assess updating behaviour.

Trials with adjectives for which participants failed to respond in time for the first or second rating were excluded. The number of these trials did not differ across the three groups ($F_{2,100} = 1.24$, $p > 0.2$) (Table 3). Furthermore, trials with a feedback discrepancy of zero were excluded from analyses involving updates since these trials could not be clearly assigned to either receiving desirable or undesirable feedback (Table 3).

Memory errors were calculated as the absolute differences between the recollected number and the actual feedback rating.

$$\text{Absolute memory error} = \text{abs}(\text{feedback rating} - \text{recollection offeedback rating}) \quad (4)$$

### Expert ratings of stimuli

Ten experts with professional contact to BPD patients rated the trait adjectives for BPD relevance. These expert ratings were intended for an analysis investigating whether BPD patients may show a particularly strong alteration of feedback processing for character traits that are highly relevant (*v.* less relevant) for BPD. The experts included three psychiatrists, four psychologists, and four nurses (seven female; mean age 40.0 years, S.D. = 12.0; mean of estimated number of treated BPD patients 396, S.D. = 600, range 30–2000; mean number of years working with BPD patients 6.05, S.D. = 4.53, range 0.5–14). The experts rated the 80 character traits on self-relevance (in comparison to healthy persons, how much do the following traits apply to female BPD patients in general?) and on other relevance. The scales ranged from −3 (*applies very much less*) to +3 (*applies very much more*). We calculated mean ratings per item and performed median splits on the items using absolute mean ratings. Using absolute means accommodated for the fact that both negative and positive ratings indicated high relevance and ratings close to the midpoint of the scale (i.e. zero) indicated low relevance (see Supplementary Table S2).

### Results

#### Participants' characteristics

We compared demographic and clinical characteristics across the three participant groups: BPD patients, healthy controls from group A (ConA) and group B

---

**Fig. 1.** Task outline. (*a*) In part 1, participants interacted in groups of five consisting of one borderline personality disorder (BPD) patient and four healthy controls by playing the board game Monopoly for 45 min. Afterwards, each participant rated three fellow players on 80 trait adjectives on a scale from 1 (*trait does not apply at all*) to 8 (*trait applies fully*). Participants did not rate themselves and one other person. (*b*) In part 2, participants provided first ratings for themselves and one and other participant (i.e. the 'fourth' participant, whom they had not rated in part 1). Importantly, participants received social feedback in the form of mean trait ratings, which they believed to be based on the mean trait ratings of part 1 but which were determined by an adaptive algorithm. The differences between participants' first own ratings were conceptualized as feedback discrepancies. After receiving feedback, participants rated themselves and the other person a second time, which allowed the calculation of update scores. (*c*) The experimental design comprised two within-subject factors for the main analyses on update scores. As described above, the feedback target was either the self or one other person. Feedback was desirable (i.e. feedback rating was higher than first own rating for a positive trait adjective) or undesirable (i.e. feedback rating was smaller than first own rating for a positive trait adjective). Negative trait adjectives were reverse-scored. (*d*) The between-subject factor group had three levels BPD/ConA/ConB. In all three groups, participants rated themselves in the self-condition. The two healthy control groups differed in the other condition: ConA rated another healthy control. ConB the BPD patient. For the main hypotheses, BPD patients were compared with ConA. We compared ConA with ConB to explore whether BPD patients were perceived differently from controls.

**Table 3.** *Task variables*

| Task variable | BPD patients | Healthy controls group A (ConA) | Healthy controls group B (ConB) |
|---|---|---|---|
| *n* trials excluded: missing answers (of 160 trials in total for self and other) | 11.2 (10.2) | 9.2 (10.3) | 7.4 (6.8) |
| *n* trials excluded for update analyses: zero feedback discrepancies (of 160 trials in total for self and other) | 16.9 (4.1) | 19.2 (4.6) | 19.1 (4.4) |
| 1st ratings – self | 5.4 (0.6) | 6.1 (0.5) | 6.0 (0.6) |
| 1st ratings – other | 5.4 (0.7) | 5.8 (0.7) | 5.5 (0.7) |
| 2nd ratings – self | 5.6 (0.5) | 6.4 (0.4) | 6.2 (0.5) |
| 2nd ratings – other | 5.6 (0.8) | 5.9 (0.8) | 5.7 (0.7) |
| Feedback – self | 5.5 (0.3) | 5.8 (0.3) | 5.8 (0.3) |
| Feedback – other | 5.5 (0.3) | 5.7 (0.3) | 5.5 (0.4) |
| Update – desirable – self | 0.38 (0.20) | 0.49 (0.33) | 0.32 (0.22) |
| Update – undesirable – self | 0.22 (0.31) | 0.07 (0.19) | 0.10 (0.18) |
| Update – desirable – other | 0.39 (0.21) | 0.41 (0.37) | 0.38 (0.29) |
| Update – undesirable – other | 0.20 (0.32) | 0.20 (0.25) | 0.13 (0.18) |
| Memory – desirable – self[a] | 2.2 (0.4) | 2.4 (0.5) | 2.2 (0.5) |
| Memory – undesirable – self[a] | 1.7 (0.5) | 1.9 (0.5) | 1.7 (0.4) |
| Memory – desirable – other[a] | 2.4 (0.6) | 2.5 (0.5) | 2.2 (0.5) |
| Memory – undesirable – other[a] | 2.0 (0.5) | 2.0 (0.4) | 1.9 (0.4) |
| Positivity rating – positive words | 6.6 (0.5) | 6.8 (0.6) | 6.7 (0.6) |
| Positivity rating – negative words | 2.1 (0.4) | 2.0 (0.4) | 2.1 (0.5) |

Data are given as mean (standard deviation). Ratings are given on an 8-point Likert scale (from 1 *trait does not apply at all* to 8 *trait applies fully*). Updates are given as relative absolute mean updates. 'Memory' refers to absolute memory errors. Participants rated the social positivity of the character trait words on an 8-point Likert scale (from 1 *not positive at all* to 8 *very positive*).

[a] Data from one patient were not collected.

(ConB). ConA rated another healthy control in the other condition whereas ConB rated a BPD patient. All statistics are reported in Table 1. The three groups did not differ in age, family status, years of education, or verbal IQ. We found significant differences between BPD patients and controls in the expected directions for GSI, BDI, and trait self-esteem. As expected due to the random group assignment, no differences emerged between ConA and ConB. Additionally, the three groups did not differ in the perceived similarity to the other person, perceived surprise about the social feedback, and winning positions in the Monopoly game. For participants' mood state see Supplementary Table S1.

### *Updating overall*

To test for the influence of BPD on feedback processing, we performed a 2 (target: self/other) × 2 (valence: desirable/undesirable) × 3 (group: BPD/ConA/ConB) ANOVA on relative absolute mean updates (Table 3). Crucially, the triple interaction reached significance ($F_{2,100} = 3.27$, $p < 0.05$, $\eta_P^2 = 0.06$). Additionally, the main effect of valence was highly significant ($F_{1,100} = 63.90$, $p < 0.001$, $\eta_P^2 = 0.39$), which indicated that overall participants updated more for desirable than for undesirable feedback. All other

effects were not significant (all $p$'s $> 0.1$). Based on the significant triple interaction, we deconstructed this ANOVA with respect to the factor target and report separate ANOVAs for self- and other-related updating in BPD patients and healthy controls from ConA.

### *Updating for self*

We conducted a 2 (valence: desirable/undesirable) × 2 (group: BPD/ConA) ANOVA on relative absolute mean updates for self. Our main hypotheses was confirmed by the significant interaction of valence and group ($F_{1,62} = 7.17$, $p < 0.01$, $\eta_P^2 = 0.10$). Additionally, the main effect of valence reached significance ($F_{1,62} = 35.05$, $p < 0.001$, $\eta_P^2 = 0.36$), indicating that participants overall updated more for desirable than for undesirable self-related feedback. The main effect of group was not significant ($p > 0.7$).

Following up on the significant interaction of valence and group, we separately compared desirable and undesirable self-related updating scores between BPD patients and healthy controls (ConA). BPD patients updated significantly more after receiving undesirable feedback ($t_{29.1} = 2.05$, $p < 0.05$), but no difference emerged for desirable feedback ($p > 0.1$).

Thus, self-relevant feedback processing differed between BPD patients and healthy controls, with BPD patients showing enhanced updating for undesirable feedback.

## Updating for other

Following up on the significant triple interaction, we also addressed whether BPD patients and ConA differed from each other in their updating pattern in the other condition using a 2 (valence: desirable/undesirable) × 2 (group: BPD/ConA) ANOVA on relative absolute mean updates. Only the main effect of valence ($F_{1,62} = 13.47$, $p < 0.001$, $\eta_P^2 = 0.18$) reached significance. This main effect indicated that overall participants updated more for desirable than for undesirable other-related feedback. The main effect of group and the interaction did not reach significance (all $p$'s > 0.8).

To explore whether healthy participants process social feedback concerning BPD patients differently than feedback concerning other healthy participants, we additionally compared relative absolute mean updates in the other condition between ConA and ConB in a 2 (valence: desirable/undesirable) × 2 (group: ConA/ConB) ANOVA. Again, only the main effect of valence ($F_{1,79} = 26.27$, $p < 0.001$, $\eta_P^2 = 0.25$) was significant, indicating higher updates for desirable than for undesirable other-related feedback. The interaction and the main effect of group did not reach significance (all $p$'s > 0.2).

Taken together, we found no significant influences of the factor group on other-related feedback processing.

## Correlation between BPD patients' self-related updating, symptom severity, and mood

To explore the relationship between undesirable feedback processing and BPD patients' state and trait variables, we correlated the difference in relative absolute mean updates for desirable *v.* undesirable feedback with symptom severity (GSI SCL-90-R and BSL-95 scores), trait self-esteem, and current mood state (the three sub-scores of the Multidimensional Mood State Questionnaire ultimately that were assessed post-task). None of the correlations reached significance (all $p$'s > 0.1) (even without correcting for multiple comparisons). This may be due to the notorious heterogeneity of BPD symptoms. Measures of symptom severity combine a variety of symptoms, most of which may not be directly related to social feedback processing.

## Updating for self- and BPD-relevant traits

We tested whether altered self-related updating in BPD patients was particularly pronounced for traits that are relevant for BPD symptomatology. We split traits into those more or less self-relevant for BPD (as rated by 10 experts on BPD; see Supplementary Table S2) and performed a 2 (group: BPD/ConA) × 2 (valence: desirable/ undesirable) × 2 (BPD relevance: high/low) ANOVA on self-related relative absolute mean updates. However, the triple interaction did not reach significance ($p > 0.3$), suggesting that the updating difference between BPD patients and healthy controls was not particularly altered for BPD-relevant traits.

## Ratings

To test for differences in participants' initial ratings, we performed 2 (target: self/other) × 3 (group: BPD/ConA/ ConB) ANOVA on mean first ratings (Table 3). The main effect of group was highly significant ($F_{2,100} = 9.88$, $p < 0.001$, $\eta_P^2 = 0.17$), and indicated that overall BPD patients gave lower ratings than healthy controls. The main effect of target was also highly significant ($F_{1,100} = 13.27$, $p < 0.001$, $\eta_P^2 = 0.12$), showing that overall participants made higher self- than other-ratings. There was a trend for the group×target interaction ($F_{2,100} = 2.59$, $p = 0.08$, $\eta_P^2 = 0.05$). Although the interaction effect was only at trend level, we still deconstructed the ANOVA to test two specific *a priori* hypotheses: First, we expected BPD patients to have more negative first self- and other-ratings than healthy controls. This was supported by significant *t* tests between BPD patients and ConA (self: $t_{62} = -4.73$, $p < 0.001$; other: $t_{62} = -2.36$, $p < 0.05$). Second, we hypothesized that healthy participants perceive BPD patients as more negative than they perceive other healthy participants, which was supported by lower first other-ratings of ConB compared to ConA ($t_{79} = 2.09$, $p < 0.05$).

## Memory, positivity ratings of trait adjectives

To exclude that memory effects could explain differences in updating between BPD patients and healthy controls, we compared participants' mean memory errors for the social feedback in a 2 (target: self/other) × 2 (valence: desirable/undesirable) × 3 (group: BPD/ConA/ConB) ANOVA (Table 3). Since none of the interactions reached significance (all $p$'s > 0.5), memory effects could not explain group differences in biased updating.

The three groups perceived the general social positivity of the trait adjectives in the same way since there were no significant effects of group on participants' mean positivity ratings for positive words or negative words ($p$'s > 0.3) (Table 3).

## Discussion

The aim of the present study was to test whether a self- and other-related positivity bias in social feedback processing, which we have previously described in

healthy participants, would be altered in BPD patients. We employed an ecologically valid set-up that combined a real-life social interaction with tight experimental control. Compared to healthy controls, BPD patients showed enhanced updating for undesirable self-relevant feedback about character traits. We found no evidence that BPD patients showed reduced updating for desirable self-relevant feedback. We also found no evidence for an altered positivity bias when BPD patients received social feedback about other persons. As expected, BPD patients rated themselves and others less favourably than controls. Conversely, healthy controls perceived BPD patients less favourably than they perceived other healthy controls but feedback about BPD patients led to similar positively updating as feedback about healthy controls. A control analysis made it unlikely that BPD patients and healthy controls interpreted the character traits differently.

### Self-evaluations and self-relevant feedback processing in BPD

Our data on initial self-ratings are in line with a considerable amount of research showing negative self-evaluations in BPD patients in the domains of self-esteem and self-concept (Klein *et al.* 2001; Sieswerda *et al.* 2005; Rüsch *et al.* 2007; Roepke *et al.* 2011; Vater *et al.* 2015).

Healthy individuals show positivity biases in many cognitive domains (Taylor & Brown, 1988; Leary, 2007). We replicated our previous results by demonstrating positively biased updating in healthy controls (Korn *et al.* 2012). In contrast, BPD patients showed enhanced updating after undesirable feedback. For updating of desirable feedback, no significant difference between BPD patients and healthy controls emerged. A previous report, which used vignettes, could not establish a relationship between BPD features in nonclinical participants and changes in self-esteem after supportive or critical feedback (Bowles *et al.* 2013). Similarly, two further studies, which investigated the instability of self-evaluations in nonclinical participants with BPD features, found conflicting results (Tolpin *et al.* 2004; Zeigler–Hill & Abraham, 2006). Our findings suggest that relating BPD to feedback processing may require a clinical population and an ecological social interaction design. In contrast to previous studies, we defined feedback as desirable or undesirable and thus took participants' initial evaluations into account.

A cognitive model of BPD (Pretzer, 1990) argues that specific, relatively stable knowledge structures organize the perception of the self in BPD patients. Self-relevant cognitive schemata in BPD include 'I am powerless and vulnerable' and 'I am inherently bad and unacceptable' (Pretzer, 1990). Unfavourable self-concepts seem to be less prone to positive modification compared to aspects that are unspecific for BPD. Negative self-evaluations in BPD have been related to maladaptive behavioural strategies such as outbursts of anger and suicidality (Levy *et al.* 2007). In consequence, the enhanced integration of undesirable feedback might explain such self-destructive behaviours in BPD. It might therefore be an interesting question for future research to assess therapeutic contexts that try to promote desirable feedback processing or to discourage undesirable feedback processing. Another interesting future direction with potential relevance to psychotherapy could be studying feedback processing in high stress situations since stress can alter information processing in general (McEwen *et al.* 2015) and since BPD patients report higher levels of stress (Stiglmayr *et al.* 2005).

### Other evaluations and other relevant feedback processing in BPD

Our data indicate that BPD patients judge the traits of other co-players in the game as more unfavourable compared to judgements of healthy individuals on healthy co-players – both for positive and negative traits. These data confirm the described negative other-evaluation bias in BPD (Westen, 1990; Arntz *et al.* 2004; Barnow *et al.* 2009; Roepke *et al.* 2013). We found no evidence for differences in other-relevant feedback processing between BPD patients and healthy participants. Thus, the current study indicates that altered integration of social feedback in BPD may be specific for self-relevant social feedback. We acknowledge, however, that other-related feedback processing is likely to depend on the specific situation and the specific role of the other person. Possibly, the social interaction in the current study was more positive than many social interactions that BPD commonly experience. Future studies should investigate whether other-relevant feedback processing differs for close others (e.g. a family member, romantic partner, or a good friend).

### Evaluations of BPD patients by healthy controls

Healthy controls evaluated the traits of BPD patients less favourably than those of other healthy participants although participants did not talk about their personal life. Prior studies have shown similar results. For example, clinicians rated the self-concept of BPD patients as more incoherent and inconsistent than the self-concept of healthy individuals or of patients with other personality disorders (Wilkinson-Ryan & Westen, 2000).

Probably, altered non-verbal communication such as facial expressions of BPD patients was one reason for why healthy controls judged BPD patients negatively. BPD patients show reduced facial expressions in responding to positive social signals and increased facial expressions in responding to negative social signals (Matzke *et al.* 2014). They further show a general reduction of facial activity in response to positive and negative stimuli (Renneberg *et al.* 2005), and more mixed and fewer positive facial emotions (Staebler *et al.* 2011).

We found no evidence that healthy controls showed a more negative updating pattern when receiving feedback about BPD patients, which suggests that at least in some settings healthy individuals can develop a more positive view of BPD patients when receiving feedback about them.

### Specificity and stability of altered self-relevant feedback processing in BPD

In this study, we observed that the impact of undesirable social feedback on self-evaluations was stronger in BPD patients than in healthy controls. But we did not find differences for other-related feedback, which suggests that the observed effect may be specific for self-relevant processes. This entails that self-evaluations might also decrease more in BPD patients than in healthy controls due to other types of self-relevant information such as information about the personal future (e.g. the probability of adverse life events) or feedback on competence (e.g. the performance in a given task). Establishing whether altered feedback processing extends to these other subclasses of self-relevant feedback is an important question for future research. A notable previous study (Schuermann *et al.* 2011) shows an impairment of BPD patients in feedback processing in a solitary task (i.e. the Iowa gambling task), which involves probabilistic feedback on decision-making competence in form of gains and losses. Thus, the current pattern may extend to competence feedback.

In a similar vein, testing the specificity of the observed findings for BPD was beyond the scope of the current study and thus we did not include a clinical control group. Here, our aim was to establish that BPD patients do show an alteration in social feedback processing. We expect that patients suffering from affective disorders may also show a different pattern than healthy controls but we expect this to be caused by a rather general negative processing pattern (Disner *et al.* 2011). We have recently reported that patients suffering from major depressive disorder (MDD) updated less from desirable information about future life events in comparison to healthy controls (Korn *et al.* 2014*b*).

However, in contrast to the effect reported here for BPD patients, MDD patients did not differ from healthy controls in their updating of undesirable information. Thus, although different psychiatric disorders may be linked to alterations in updating the exact patterns may very well differ.

Another point concerns the stability over time of the impact of social feedback. It remains an open question how long participants would be influenced by the feedback from relative strangers whom they are likely not going to meet again. Previous research is scarce and two studies on the time course of preference changes currently provide mixed results. On the one hand, research on social conformity has indicated that the opinion of others only influences preferences for only a few days (Huang *et al.* 2014). On the other hand, preference changes induced by making a choice between the relevant options lasted for over a year (Sharot *et al.* 2013). In any case, an interesting hypothesis would be that undesirable feedback persists longer than desirable feedback in BPD.

### Strengths and limitations

To our knowledge, this is the first study to assess the processing of self- and other-relevant social feedback on character traits in BPD patients. To increase the ecological validity of our procedure, participants engaged in a real-life social interaction prior to receiving social feedback. Recently, Schillbach and colleagues have convincingly outlined the merits of using such real-life social encounters to investigate social cognition in healthy and psychiatric populations (Schilbach *et al.* 2013).

A crucial feature of our design was that participants' own ratings and the feedback ratings they received were numerical. This allowed a straightforward quantification of updating behaviour. To enhance power and generalizability across character traits, we employed a rather large set of character traits. Nevertheless, we would expect that the obtained results generalize to settings with verbal feedback on a smaller number of traits (e.g. videos from the interaction partners making statements such as 'You were very polite').

Our study provides a multifaceted account of social information processing in BPD because we assessed how BPD patients process social feedback about the traits of themselves and of others. Our design also allowed us to assess how healthy controls process information about BPD patients.

Nevertheless, we would like to acknowledge the following limitations. First, as many studies on BPD, we only tested women and a number of BPD patients had diagnosed co-morbidities and took psychotropic medication. Second, social feedback was manipulated. Although no participant reported doubts about the feedback,

realistic feedback may be perceived as more pertinent. Third, trials for the initial self- and other-evaluations were intermixed and additionally interrupted by social feedback. Separating self- and other conditions into different blocks might result in purer measures of self- and other-relevant feedback processing.

## Conclusions

Our study provides an important first step towards a characterization of social feedback processing in BPD patients. Our results indicate that BPD patients show an altered positivity bias with respect to healthy controls: BPD patients integrate undesirable feedback more than healthy controls when receiving character trait information about the self but not about others. This relates the previously described unstable and negative self-evaluations of BPD patients to aberrant social feedback processing. In the long term, research based on the current findings may thus contribute to providing better treatment options via delineating ways to nudge self-relevant feedback processing in BPD toward the positive.

## Supplementary material

For supplementary material accompanying this paper visit http://dx.doi.org/10.1017/S003329171500207X.

## Acknowledgements

## Declaration of Interest

None.

## References

**Alicke MD, Sedikides C** (2009). Self-enhancement and self-protection: what they are and what they do. *European Review of Social Psychology* **20**, 1–48.

**APA** (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.

**APA** (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Association: Washington, DC.

**Arntz A, Dreessen L, Schouten E, Weertman A** (2004). Beliefs in personality disorders: a test with the personality disorder belief questionnaire. *Behaviour Research and Therapy* **42**, 1215–1225.

**Arntz A, ten Haaf J** (2012). Social cognition in borderline personality disorder: evidence for dichotomous thinking but no evidence for less complex attributions. *Behaviour Research and Therapy* **50**, 707–18.

**Barnow S, Stopsack M, Grabe HJ, Meinke C, Spitzer C, Kronmüller K, Sieswerda S** (2009). Interpersonal evaluation bias in borderline personality disorder. *Behaviour Research and Therapy* **47**, 359–365.

**Beck A, Steer R, Garbin M** (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review* **8**, 77–100.

**Bohus M, Limberger MF, Frank U, Chapman AL, Kühler T, Stieglitz R-D** (2007). Psychometric properties of the Borderline Symptom List (BSL). *Psychopathology* **40**, 126–132.

**Bowles DP, Armitage CJ, Drabble J, Meyer B** (2013). Self-esteem and other-esteem in college students with borderline and avoidant personality disorder features: an experimental vignette study. *Personality and Mental Health* **319**, 307–319.

**Disner SG, Beevers CG, Haigh EaP, Beck AT** (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience* **12**, 467–477.

**First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS** (1997). *Structured Clinical Interview for DSM-IV axis II Personality Disorders (SCID-II)*. American Psychiatric Press, Inc.: Washington, DC.

**Franke GH** (1995). *SCL-90-R the Symptom Checklist of Derogatis*. German Version. Testzentrale: Göttingen, Germany.

**Franzen N, Hagenhoff M, Baer N, Schmidt A, Mier D, Sammer G, Gallhofer B, Kirsch P, Lis S** (2011). Superior 'theory of mind' in borderline personality disorder: an analysis of interaction behavior in a virtual trust game. *Psychiatry Research* **187**, 224–233.

**Hautzinger M, Bailer M, Worall H** (1994). *The Beck Depression Inventory (BDI)* [in German]. Huber: Bern.

**Huang Y, Kendrick KM, Yu R** (2014). Conformity to the opinions of other people lasts for no more than 3 days. *Psychological Science* **25**, 1388–1393.

**King-Casas B, Sharp C, Lomax-Bream L, Lohrenz T, Fonagy P, & Montague PR** (2008). The rupture and repair of cooperation in borderline personality disorder. *Science* **321**, 806–810.

**Klein MH, Wonderlich SA, Crosby R** (2001). Self-concept correlates of the personality disorders. *Journal of Personality Disorders* **15**, 150–156.

**Korn CW, Fan Y, Zhang K, Wang C, Han S, Heekeren HR** (2014*a*). Cultural influences on social feedback processing of character traits. *Frontiers in Human Neuroscience* **8**, 1–18.

**Korn CW, Prehn K, Park SQ, Walter H, Heekeren HR** (2012). Positively biased processing of self-relevant social feedback. *Journal of Neuroscience* **32**, 16832–16844.

**Korn CW, Sharot T, Walter H, Heekeren HR, Dolan RJ** (2014*b*). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine* **44**, 579–592.

**Leary MR** (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology* **58**, 317–344.

Levy KN, Edell WS, McGlashan TH (2007). Depressive experiences in inpatients with borderline personality disorder. *Psychiatric Quarterly* **78**, 129–143.

Matzke B, Herpertz SC, Berger C, Fleischer M, Domes G (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology* **47**, 101–110.

McEwen BS, Gray JD, Nasca C (2015). 60 years of neuroendocrinology: redefining neuroendocrinology: stress, sex and cognitive and emotional regulation. *Journal of Endocrinology* **226**, T67–T83.

Pretzer J (1990). Borderline personality disorder. In *Cognitive Therapy of Personality Disorder* (ed. T. A. Beck and A. Freeman), pp. 176–207. Guilford Press: New York.

Renneberg B, Herm K, Hahn A, Staebler K, Lammers C-H, Roepke S (2012). Perception of social participation in borderline personality disorder. *Clinical Psychology & Psychotherapy* **19**, 473–480.

Renneberg B, Heyn K, Gebhard R, Bachmann S (2005). Facial expression of emotions in borderline personality disorder and depression. *Journal of Behavior Therapy and Experimental Psychiatry* **36**, 183–196.

Roepke S, Schröder-Abé M, Schütz A, Jacob G, Dams A, Vater A, Rüter A, Merkl A, Heuser I, Lammers C, Franklin B (2011). Dialectic behavioural therapy has an impact on self-concept clarity and facets of self-esteem in women with borderline personality disorder. *Clinical Psychology and Psychotherapy* **18**, 148–158.

Roepke S, Vater A, Preißler S, Heekeren HR, Dziobek I (2013). Social cognition in borderline personality disorder. *Frontiers in Neuroscience* **6**, 195.

Rosenberg M (1965). *Society and the Adolescent Self-Image*. Princeton University Press: Princeton, New Jersey.

Rüsch N, Lieb K, Göttler I, Hermann C, Schramm E, Richter H, Jacob GA, Corrigan PW, Bohus M (2007). Shame and implicit self-concept in women with borderline personality disorder. *American Journal of Psychiatry* **164**, 500–508.

Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K (2013). Toward a second-person neuroscience. *The Behavioral and Brain Sciences* **36**, 393–414.

Schmidt KH, Metzler P (1992). *Wortschatztest* (*WST*). Beltz Test GmbH: Weinheim, Germany.

Schuermann B, Kathmann N, Stiglmayr C, Renneberg B, Endrass T (2011). Impaired decision making and feedback evaluation in borderline personality disorder. *Psychological Medicine* **41**, 1917–1927.

Sharot T, Fleming SM, Yu X, Koster R, Raymond J (2013). Is choice-induced preference change long lasting? *Psychological Science* **23**, 1123–1129.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59** (Suppl. 20:22), quiz 34–57.

Sieswerda S, Arntz A, Wolfis M (2005). Evaluations of emotional noninterpersonal situations by patients with borderline personality disorder. *Journal of Behavior Therapy and Experimental Psychiatry* **36**, 209–225.

Staebler K, Renneberg B, Stopsack M, Fiedler P, Weiler M, Roepke S (2011). Facial emotional expression in reaction to social exclusion in borderline personality disorder. *Psychological Medicine* **41**, 1929–1938.

Steyer R, Schwenkmezger P, Notz P, Eid M (1997). *Multidimensional Mood State Questionnaire* (*MDBF*) [*in German*]. Hogrefe: Göttingen, Germany.

Stiglmayr CE, Grathwol T, Linehan MM, Ihorst G, Fahrenberg J, Bohus M (2005). Aversive tension in patients with borderline personality disorder: a computer-based controlled field study. *Acta Psychiatrica Scandinavica* **111**, 372–379.

Taylor SE, Brown JD (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin* **103**, 193–210.

Taylor SE, Collins RL, Skokan LA, Aspinwall LG (1989). Maintaining positive illusions in the face of negative information: getting the facts without letting them get you. *Journal of Social and Clinical Psychology* **8**, 114–129.

Tolpin HL, Gunthert KC, Cohen LH, O'Neill S (2004). Borderline personality features and instability of daily negative affect and self-esteem. *Journal of Personality* **72**, 111–138.

Unoka Z, Seres I, Aspan N, Bodi N, Keri S (2009). Tust game reveals restricted interpersonal transactions in patients with borderline personality disorder. *Journal of Personality Disorders* **23**, 399–409.

Vater A, Schröder-Abé M, Weißgerber S, Roepke S, Schütz A (2015). Self-concept structure and borderline personality disorder: evidence for negative compartmentalization. *Journal of Behavior Therapy and Experimental Psychiatry* **46**, 50–58.

Veen G, Arntz A (2000). Multidimensional dichotomous thinking characterizes borderline personality disorder. *Cognitive Therapy and Research* **24**, 23–45.

Westen D (1990). Towards a revised theory of borderline object relations: contributions of empirical research. *The International Journal of Psychoanalysis* **71**, 661–693.

Wilkinson-Ryan T, Westen D (2000). Identity sisturbance in borderline personality disorder: an empirical investigation. *American Journal of Psychiatry* **157**, 528–541.

Zeigler–Hill V, Abraham J (2006). Borderline personality features: instability of self-esteem and affect. *Journal of Social and Clinical Psychology* **25**, 668–687.