

# **Women with borderline personality disorder show reduced identification of emotional facial expressions and a heightened negativity bias**

Martin Vestergaard<sup>1\*</sup>, Mickey T. Kongerslev<sup>1,2,3</sup>, Marianne S. Thomsen<sup>1,3,4</sup>, Birgit Bork Mathiesen<sup>4</sup>, Catherine J. Harmer<sup>5</sup>, Erik Simonsen<sup>1,6</sup>, Kamilla W. Miskowiak<sup>4,7</sup>,

## **Affiliations**

<sup>1</sup> Psychiatric Research Unit, Psychiatry Region Zealand, Denmark

<sup>2</sup> Department of Psychology, University of Southern Denmark, Denmark

<sup>3</sup> Psychiatric Clinic, Psychiatry Roskilde, Region Zealand, Denmark

<sup>4</sup> Department of Psychology, University of Copenhagen, Denmark

<sup>5</sup> Department of Psychiatry, University of Oxford, United Kingdom

<sup>6</sup> Institute of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

<sup>7</sup> Copenhagen Affective Disorder Research Centre, Psychiatric Center Copenhagen, Rigshospitalet, Copenhagen, Denmark

## **\* Corresponding author**

Martin Vestergaard

Psychiatric Research Unit

Psychiatry Region Zealand, Denmark

Faelledvej 6, 4200 Slagelse

Phone: +45 5150 7031, E-mail: mgoe@regionsjaelland.dk

## **Keywords**

Borderline personality disorder, dot probe task, negativity bias, faces, anger, fear

## **Abstract**

Individuals with borderline personality disorder (BPD) frequently display impairments in the identification of emotional facial expressions paralleled by a negativity bias. However, it remains unclear whether misperception of facial expressions is a key psychopathological marker of BPD. To address this question, we examined 43 women diagnosed with BPD and 56 healthy female controls using an emotion face identification task and a face dot-probe task together with measures on psychopathology. We observed that compared to controls, women with BPD showed impairments in the identification of emotional facial expressions, most notably for disgusted and angry faces, concurrent with a bias to misclassify faces as angry and a faster preconscious vigilance for fearful relative to happy facial expressions. Increased severity of borderline symptoms and global psychopathology in BPD patients were associated with reduced ability to identify angry facial expressions and a stronger negativity bias to anger. The findings indicate that BPD patients who misperceive face emotions have the greatest mental health issues.

## **Introduction**

Borderline personality disorder (BPD) is characterized by maladaptive interpersonal relationships, impulsivity and persistent feelings of intense negative affect (APA, 2013). Affective instability is a core symptom in BPD and converging evidence suggest that emotional dysregulation is linked to a negativity bias in the way subtle situational cues are experienced by individuals who suffer from BPD (Baer, Peters, Eisenlohr-Moul, Geiger, & Sauer, 2012). One critical situational trigger associated with affective instability in BPD may relate to the misperception of social cues signaled by other people, including emotional facial expressions (Herpertz & Bertsch, 2014; Lazarus, Cheavens, Festa, & Rosenthal, 2014). A substantial number of studies have examined if individuals diagnosed with BPD show underlying deficits in the ability to detect and identify the emotional valence of facial expressions (Mitchell, Dickens, & Picchioni, 2014). However, the findings are somewhat mixed and it is currently unclear whether impairments in the identification of emotional facial expressions in BPD are global or emotion-specific. In particular, studies have reported that borderline patients show decreased ability to identify expressions of disgust (Jovev et al., 2011; Nicol, Pope, & Hall, 2014), fear (Jovev et al., 2011), surprise (Domes et al., 2008), anger and happiness (Robin et al., 2012) or global deficits in the identification of emotional faces (Unoka, Fogd, Fuzy, & Csukly, 2011). It was concluded in one meta-analysis that impaired emotional face identification in BPD patients was most pronounced for anger and disgust (Daros, Zakzanis, & Ruocco, 2013) while a more recent meta-analysis established that the most reliable deficits, if any, were restricted to disgusted expressions (Mitchell et al., 2014).

While it remains unclear whether BPD is characterized by emotion-specific impairments in the identification of facial expressions, studies have shown that borderline patients demonstrate a negativity bias in the encoding of facial expressions

(Mitchell et al., 2014). Specifically, patients with BPD tend to perceive ambiguous or neutral facial expressions as aversive (Catalan et al., 2016; Domes et al., 2008; Dyck et al., 2009), and evaluate neutral faces as less trustworthy compared to healthy controls (Fertuck, Grinband, & Stanley, 2013). More borderline-related symptoms in female BPD patients have also been associated with increased misidentification of anger in male faces (Veague & Hooley, 2014). Furthermore, it appears that a negativity bias to perceive facial expressions as aversive is paralleled by a lowered positivity bias for identification of happy facial expressions. Studies have shown that relative to controls, individuals with BPD experienced happy facial expressions with less intensity compared to feelings of anger and loneliness (Thome et al., 2016), misclassified happy facial expressions as aversive (Catalan et al., 2016), and rated happy faces as less approachable (Kobeleva et al., 2014). Borderline patients even appear to display attenuated activity of the levator labii superioris eye muscle, involved in smiling, to happy facial expressions but increased activity in the corrugator supercilii frowning muscle in response to angry, disgusted and sad faces (Matzke, Herpertz, Berger, Fleischer, & Domes, 2014).

It is unresolved whether the observed negativity bias in borderline patients happens at an automatic preconscious stage and/or is modulated at later stages by higher-order selective attention (Mitchell et al., 2014). Adolescent and young adult borderline patients were shown to display a greater vigilance bias for fearful facial expressions on a dot-probe task, when faces were presented preconsciously but not consciously (Jovev et al., 2012). Another study reported that only adolescent female BPD patients in a lowered mood displayed a heightened vigilance for consciously perceived aversive facial expressions compared to healthy and psychiatric controls (von Ceumern-Lindenstjerna et al., 2010). Lastly, a recent study did not observe BPD patients to differ from controls in the evaluation of facial expressions following

preconscious priming with angry or happy facial expressions (Donges, Dukalski, Kersting, & Suslow, 2015).

Emerging evidence indicates that the relationship between BPD and the tendency to misperceive facial expressions as aversive is mediated by the severity of underlying borderline traits. The core symptoms in BPD including trait impulsivity, emotion dysregulation, aggression and self-harm display strong inter-correlations (Terzi et al., 2017), and findings from a longitudinal study in individuals high on borderline traits suggest that affective instability and impulsivity are closely inter-related over time (Tragesser, Solhan, Schwartz-Mette, & Trull, 2007). Increased levels of BPD traits in otherwise healthy young adults have been shown to be associated with reduced identification of angry faces for those individuals who also had high but not low levels of impulsivity (Preti, Richetin, Suttora, & Pisani, 2016), and healthy adults with high levels of trait impulsivity was reported to display increased detection sensitivity to fearful facial expressions in aversive scenarios (Lee, Choi, & Cho, 2012). The bias to misperceive facial expressions as aversive in BPD may originate early in life. Studies have shown that in patients with BPD, a high incidence of self-reported childhood trauma experiences was associated with impaired identification of facial expressions, particularly disgusted faces (Nicol et al., 2014), and with faces being subjectively evaluated as more unapproachable (Nicol, Pope, Sprengelmeyer, Young, & Hall, 2013).

In the present study, we hypothesized that women diagnosed with BPD would show impaired identification of facial expressions paralleled by a negativity bias on a standard emotional face identification task, and a negative vigilance bias on a dot-probe task. If deficits in the identification of facial expressions and a negativity bias for aversive faces are indeed phenotypic traits in BPD, we would expect the degree of borderline-related measures such as trait impulsivity, childhood trauma and/or psychopathological symptoms to show a linear association with the ability to identify

facial expressions and negativity bias. Therefore, planned exploratory analyses were performed to assess whether observed group differences in facial expression identification and negativity bias were coupled to impulsivity, experienced childhood adversity or symptom severity in patients. Finally, we explored whether the recently reported impairments in higher-order cognitive functions in our cohort (Thomsen, Ruocco, Carcone, Mathiesen, & Simonsen, 2016), were associated with inter-individual variations in emotional face identification ability and negativity bias.

## **Material and Methods**

### **Subjects**

Forty-three women aged 18-45 years, who met diagnostic criteria for BPD, for at least two years, according to the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) were included in the study. Fifty-six controls were matched on age, gender and parental education (missing data on parental education for 4 patients) to control for differences in socioeconomic status for the family of origin. Each subject's educational level (missing data for 2 patients) was also recorded to assess whether inter-individual variability in educational level mediated group differences in the present study. Parent education was coded in one of 5 categories (1 = no education; 2 = trained worker; 3 = skilled worker; 4 = theoretical education; 5 = academic education). When data on parental education was available for both parents, the mean of the parental education categories were reported. Similarly, the educational level for each subject was coded in one of 6 categories (1=no education or primary school; 2 = skilled worker; 3 = university-preparatory school; 4 = short academic education; 5=academic education corresponding to a bachelor; 6 = academic education corresponding to a master degree).

Exclusion criteria for all participants included a score  $\leq 70$  on the Verbal

Comprehension Index (VCI), estimated based on the Similarities, Vocabulary and Information sub-tests from the Wechsler Adult Intelligence Scale, Fourth edition (WAIS-IV), lifetime psychotic disorder, lifetime bipolar I disorder or substance use disorder within the past three months according to the DSM-IV, history of significant head trauma, severe chronic physical illness or a neurological condition such as seizure disorder, encephalitis, or stroke. The matching variables are displayed in Table 1.

Patients were recruited from the Psychiatric Clinic in Roskilde, Denmark, and controls were recruited through local advertisement. All participants received thorough written and verbal information regarding the project and were required to provide written informed consent before inclusion to the study. Recruitment procedures are described in more detail elsewhere (Thomsen et al., 2016)

### **Diagnostic assessment**

Trained clinicians used the Mini International Neuropsychiatric Interview (M.I.N.I.) to screen each participant for psychiatric disorders (Sheehan et al., 1998). The Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II) was used to assess personality disorders including BPD and each subject in the patient cohort fulfilled the criteria of BPD (First, 1997).

### **Symptom rating scales**

The Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) was used to assess the severity of current BPD symptoms (Zanarini et al., 2003). The ZAN-BPD is a clinician-administered scale for the assessment of BPD psychopathology. The scale assesses nine items, which reflect the core areas of BPD psychopathology.

Severity of current depressive symptoms was evaluated using the Hamilton Depression Rating Scale (Hamilton, 1960), which shows good reliability (missing data

for 3 patients and one control).

Psychopathology in patients and controls (missing data for 5 patients and one control) was assessed with the Symptom Checklist 90-Revised (SCL-90-R), which is a 90-item self-report questionnaire. The SCL-90-R includes a global severity index (GSI) as well as nine subscales that estimate different domains of psychopathology including somatization, depression, anxiety, phobic anxiety, hostility, obsessive and compulsive behaviour, paranoid ideation, interpersonal insensitivity and psychoticism (Derogatis, 1983).

### **Composite executive function score**

We have previously shown that compared to controls, patients had lower working memory capacity on the Spatial Span task, reduced sustained attention on the Rapid Visual Processing task and a trend for worse inhibition on the Stop Signal task assessed with the Cambridge Neuropsychological Test Automated Battery (CANTAB; [www.cambridgecognition.com/cantab](http://www.cambridgecognition.com/cantab)) reported elsewhere (Thomsen et al., 2016). In the present study we also included outcome measures from the Attention Switching task included in the CANTAB. We estimated performance on the Rapid Visual Processing task and Attention Switching task, respectively, by calculation of the  $d'$  based on signal detection theory (Stanislaw & Todorov, 1999). The  $d'$  provides an index of the ability to discriminate signal from noise and are estimated by subtraction of the Z-transformed alarm rate from the Z-transformed hit rate:  $d' = Z(\text{hits}/(\text{hits} + \text{misses})) - Z(\text{false alarms}/(\text{false alarms} + \text{correct rejections}))$ . Because it is not possible to estimate the  $d'$  if the hit rate or alarm rate is equal to one or zero, respectively, hit rate scores of one were corrected by the calculation of  $1 - 1/(2N)$  whereas alarm rate scores of zero were corrected by calculating  $1/(2N)$ . Finally, the Z-score was calculated for each of the four outcome measures and summed further adding 100 to ease interpretation: composite



executive function score = (Z(Spatial Span task) + Z(Rapid Visual Processing task) + Z(Stop Signal task) + Z(Attention Switching task) + 100). The composite executive function score is reported in Table 1.

### **Childhood trauma and trait impulsivity**

We assessed childhood trauma experiences with the Childhood Trauma Questionnaire (CTQ). The CTQ is a standardized, retrospective 28-item self-report inventory that measures the severity of different types of childhood trauma, which is summed to a total score (Bernstein & Fink, 1998). Five patients and one control had missing data for the CTQ.

The Barratt Impulsiveness Scale 11 (BIS-11) is a 30 item self-report questionnaire designed to assess global impulsiveness accounting for the multi-factorial construct of the concept (Patton et al., 1995). The BIS questionnaire measures attention, motor impulsivity, self-control, cognitive complexity, perseverance and cognitive instability, which are summed to obtain a total impulsivity score. Six patients and one control had missing data for the BIS 11. The CTQ and BIS 11 scores are reported in Table 1.

### **Emotional face identification task**

The emotional face identification task (EFIT) assesses the subject's ability to identify emotional facial expressions at different intensities based on morphed faces. The validity of the emotional face identification paradigm is well established (Young et al., 1997), and the specific EFIT design used in the present study has previously been shown to result in adequate between-group effect sizes (Bhagwagar, Cowen, Goodwin, & Harmer, 2004). The participants viewed images of faces from Ekman and Friesen (Ekman & Friesen, 1976) that displayed either a happy, surprised, sad, fearful, angry,

disgusted or a neutral facial expression. The images were presented on a stationary computer screen in randomized order for 500 ms, after which a black screen immediately replaced the image. Participants were instructed to identify the emotional expressions as quickly and accurately as possible by pressing a corresponding button on the keyboard. For each face identity used in the task, the corresponding facial expression from each emotional category was morphed with the neutral facial expression to obtain different morphing resolutions of 10% increases from a neutral face (0%) to the emotional face (100%). There were a total of 250 stimuli presentations consisting of four examples of every emotion at each intensity level plus a neutral face for every emotion.

To assess the ability to identify anger, disgust, fear, sadness, happiness and surprise from each other, the  $d'$  was estimated for each emotional category by subtracting the Z-transformed alarm rate (false alarms/(false alarms + correct rejections)) from the Z-transformed hit rate (hits/(hits + misses)). Higher  $d'$  values indicate better detectability of emotional facial expressions. To estimate the emotional bias score we used the alarm rate adjusted for hit rate (alarm rate / hit rate  $\times$  100) to ensure that more false alarms was not dependent on a proportional increase in number of hits. Five controls had missing data on the EFIT.

### **Emotional Faces Dot-probe Task**

Vigilance to happy and fearful faces was examined with an emotional face dot-probe task (DPT) design (Murphy, Downham, Cowen, & Harmer, 2008). In each trial, a pair of faces was presented simultaneously, above and below the center, on a computer screen. Each face pair included either an emotional and a neutral expression or two neutral expressions of the same facial identity. Half of the emotional faces expressed fear and the other half happiness. The different facial expressions appeared in the top

and bottom of the screen with equal frequency. Faces were displayed either unmasked or masked. In the unmasked trial condition, face pairs were displayed for 100 milliseconds, and in the masked condition the emotional faces were shown for 17 milliseconds immediately replaced by a neutral mask for 83 milliseconds. On each trial, one of the two faces was immediately replaced by two dots displayed either vertically or horizontally. Participants were instructed to indicate the orientation of the dots as quickly and accurately as possible by pressing the corresponding button on the keyboard. The dots remained on the screen until the subject had responded. The task included a total of 192 trials comprised of 32 masked happy-neutral pairs, 32 masked fear-neutral pairs, 32 masked neutral-neutral pairs, 32 unmasked happy-neutral pairs, 32 unmasked fear-neutral pairs and 32 unmasked neutral-neutral pairs. Eight blocks of unmasked trials and eight blocks of masked trials were presented in an alternating order. Each block included 12 trials. Trials, in which the dot appeared in the location of the previously displayed neutral face was defined as incongruent, while dots that followed the emotional face was defined as congruent. The mean reaction time on hit trials for incongruent and congruent trials were obtained and used for further analyses. Thus, faster reaction times to dots that replace an emotional facial expression on congruent trials may indicate heightened vigilance to the expressed emotion while a slowed reaction time to dots on incongruent trials could indicate reduced attentional disengagement from the emotionally salient face.

To calculate the bias score for the masked and unmasked fear and happy faces, respectively, we subtracted the mean reaction time for the congruent trial condition from the mean reaction time for the incongruent trial condition and normalized the difference score to adjust for individual variations in reaction times:  $((\text{incongruent} - \text{congruent}) / (\text{incongruent} + \text{congruent}) \times 100)$ .

## Statistical analyses

Statistical analyses were conducted using SPSS 23. Group differences for the matching variables and control variables were tested with two-tailed t-tests unless the variables were impossible to normalize in which case group differences were tested with Mann-Whitney U test. We used repeated measures ANCOVA to test our hypotheses and multiple linear regression models were used to further explore the directionality of data. The models were inspected to ensure that the covariates fulfilled criteria of non-collinearity and normal distribution of the residuals. A p-value below 0.05 was considered significant.

Continuous variables were determined to have a significantly non-normal distribution if either of the respective standard errors for the skewness or kurtosis were above or below  $Z \geq \pm 1.96$  (two-sided  $p < 0.05$ ) in the patients and/or controls. Non-normally distributed variables were normalized using the Rankit transformation, and the normalized values were used for the linear analyses. Age, subject education, the EFIT bias scores and the CTQ scores were significantly non-normally distributed ( $p < 0.05$ ) and were successfully normalized with Rankit transformation. The VCI, parent education and BIS 11 in patients and controls, as well as the ZAN-BPD score in patients appeared normally distributed ( $p > 0.05$ ). The EFIT  $d'$  scores appeared normally distributed ( $p > 0.05$ ) except for the EFIT happy  $d'$  ( $p < 0.05$ ). Visual inspection of the data showed that one patient was an extreme outlier in the lower range, and the EFIT happy  $d'$  for that particular subject (one patient) was therefore excluded, after which the EFIT happy  $d'$  appeared normally distributed ( $p > 0.05$ ). Because the DPT masked happy bias and unmasked fear bias scores were non-normally distributed ( $p < 0.05$ ), the DPT bias measures for masked and unmasked faces were normalized to keep the variables scaled to each other. The SCL-90 and HAM-D scores appeared normally distributed in patients ( $p > 0.05$ ) but with non-normal kurtosis and skewness in controls ( $p < 0.05$ ) and

neither SCL-90 nor HAM-D could be normalized. From Table 1 it is apparent that the variability for the SCL-90 and HAM-D in controls is extremely limited and that floor effects are pronounced. The original values of variables that were considered non-normally distributed are displayed with median and lower and upper quartiles in Table 1.

Group differences in the EFIT response bias and EFIT  $d'$  for each emotional face category and group by factor interactions were tested with repeated measures ANOVA. If patients and controls differed in the response bias or the  $d'$  for faces or an interaction for group by factor was observed, two-tailed independent sample t-tests were performed to explore whether group differences were associated with specific faces. Group differences in masked and unmasked attentional bias for the happy versus fearful faces, respectively, on the DPT were likewise tested with repeated measures ANOVA. Cohen's  $d$ , estimated with pooled standard deviations, was reported for significant group differences.

Additional planned exploratory analyses were restricted to the exploration of significant group differences using multiple linear regression models. To ensure that observed group differences were not mediated by the matching variables, multiple linear regression models that tested group differences were controlled for subject age and parent education. Subsequently, each control variable was added to the multiple linear regression models including subject educational level, VCI, composite executive function, BIS or CTQ, respectively, to explore whether the controls variables mediated the findings. Each control variable was included in separate linear regression models, first correcting for the variable, then we included an interaction term for group by control variable, and lastly we explored whether the control variable was associated with the outcome measure across patients and controls excluding group as covariate. Lastly, we examined whether symptoms scale for ZAN-BPD, SCL or HAM-D was

associated with each outcome measure that showed significant group differences.

## **Results**

Patients and controls were matched on age ( $t(97)=0.175$ ;  $p=0.86$ ) and parental education ( $t(93)=-1.350$ ;  $p=0.18$ ). In agreement with our previous findings (Thomsen et al., 2016), compared to controls patients had lower educational level ( $t(95)=-4.664$ ;  $p<0.0001$ ), poorer VCI score ( $t(97)=-5.052$ ;  $p<0.0001$ ), higher CTQ score ( $t(91)=9.854$ ;  $p<0.0001$ ) and lower executive functions ( $t(97)=-5.055$ ;  $p<0.0001$ ), which were assessed by estimating a composite executive function score in the present study. Moreover, patients had more impulsive behavior on the BIS relative to controls ( $t(90)=8.274$ ;  $p<0.0001$ ). As mentioned in the section on statistical analyses, because controls had extremely low scores on the SCL-90 and HAM-D group differences could not be assessed with linear analyses. However, non-parametric analysis with Mann-Whitney U showed that patients and controls differed significantly on SCL-90 ( $p<0.001$ ) and HAM-D ( $p<0.001$ ). From inspection of Table 1, it is apparent that patients scored much higher than controls on the SCL-90 and HAM-D, which was expected since the SCL-90 and HAM-D are clinical symptom scales.

### **Group differences on the emotional face identification task**

Repeated measures ANOVA showed significant group difference in average EFIT  $d'$  ( $F(1,91)=4.265$ ,  $p=0.042$ ,  $\eta^2=0.045$ ) with no significant interaction for group by facial expression category ( $F(1,91)=0.774$ ,  $p=0.38$ ,  $\eta^2=0.008$ ). T-test analysis confirmed that patients relative to controls had lower average EFIT  $d'$  ( $t=-2.065$ ,  $p=0.042$ , Cohen's  $d=0.43$ ), and exploratory analyses for each emotional category showed that patients had lower  $d'$  for angry faces ( $t=-2.004$ ,  $p=0.048$ , Cohen's  $d=0.41$ ), disgusted faces ( $t=-2.530$ ,

$p=0.013$ , Cohen's  $d=0.52$ ) with trend-significance for lower  $d'$  for surprised faces ( $t=-1.927$ ,  $p=0.057$ , Cohen's  $d=0.4$ ) and happy faces ( $t=-1.863$ ,  $p=0.066$ , Cohen's  $d=0.38$ ) but not fearful ( $t=-1.099$ ,  $p=0.28$ , Cohen's  $d=0.23$ ) or sad faces ( $t=-0.962$ ,  $p=0.49$ , Cohen's  $d=0.14$ ). Group differences in  $d'$  for each of the respective face emotions on the EFIT are displayed in Figure 1A.

Repeated measures ANOVA showed a significant group difference in the average of the bias scores ( $F(1,92)=4.587$ ,  $p=0.035$ ,  $\eta^2=0.047$ ) with no apparent interaction for group by emotion bias score ( $F(1,92)=1.502$ ,  $p=0.22$ ,  $\eta^2=0.016$ ). T-test analysis confirmed that patients relative to controls had more false alarms relative to hits on average ( $t=-2.142$ ,  $p=0.035$ , Cohen's  $d=0.44$ ). Post hoc analyses showed that patients had a significantly heightened bias for angry faces ( $t=-2.406$ ,  $p=0.018$ , Cohen's  $d=0.5$ ) while similar group differences were absent for the fearful, disgusted, happy, sad or surprised facial expressions ( $p\geq 0.14$ , Cohen's  $d\leq 0.31$ ). Group differences for each of the face emotion bias scores are displayed in Figure 1B.

### **Group differences on the face dot-probe Task**

The mean performance accuracy on the DPT for the masked and unmasked, happy and fearful, congruent and incongruent trials, respectively, were in the acceptable range for patients (92-95%) and controls (92-94%). The repeated measures ANOVA for the DPT showed a significant interaction for group by masked facial expression category ( $F(1,97)=5.315$ ,  $p=0.023$ ,  $\eta^2=0.052$ ) with no apparent group difference in the average preconscious bias for happy and fearful facial expressions ( $F(1,97)=0.352$ ,  $p=0.55$ ,  $\eta^2=0.004$ ). Visual inspection of the results displayed in Figure 2 and Table 2 suggests that compared to controls, patients had a stronger vigilance bias for masked fearful faces but a lowered bias for masked happy faces, which was confirmed by the T-test

( $t=2.305$ ,  $p=0.023$ , Cohen's  $d=0.46$ ). Note that higher scores indicate that the group on average had faster reaction times on congruent relative to incongruent trials. Relative to controls, patients had significantly lower vigilance bias scores for happy faces ( $t=-2.027$ ,  $p=0.046$ , Cohen's  $d=0.42$ ) with no significant difference for the fear bias score ( $t=1.253$ ,  $p=0.21$ , Cohen's  $d=0.25$ ).

The repeated measures ANOVA did not show significant differences for group by unmasked face emotions ( $F(1,97)=0.794$ ,  $p=0.375$ ,  $\eta^2=0.008$ ) or group difference in the average reaction time bias for unmasked face emotions ( $F(1,97)=1.086$ ,  $p=0.3$ ,  $\eta^2=0.01$ ). Therefore, we did not perform additional analyses to explore group differences for unmasked consciously perceived facial expressions.

### **Exploratory analyses**

We proceeded to explore whether inter-individual variations in the control variables and symptoms scales were associated with the significant group differences for the EFIT anger  $d'$ , disgust  $d'$ , the EFIT bias measure for anger, and the DPT difference score for masked happy relative to fearful faces, respectively.

When we corrected for parent educational level and subject age, group differences remained significant for the EFIT angry  $d'$  ( $t=2.263$ ;  $p=0.026$ ), EFIT disgust  $d'$  ( $t=3.078$ ,  $p=0.003$ ) and the EFIT anger bias score ( $t=-2.542$ ;  $p=0.013$ ) while the DPT negativity bias score approached significance ( $t=1.965$ ;  $p=0.052$ ). After additionally controlling for each of the control variables in separate multiple linear regression models, the group differences for the EFIT angry  $d'$  or EFIT disgust  $d'$  did not survive correction for neither subject educational level, VCI score, BIS 11 or CTQ ( $p>0.05$ ), while group differences persisted after correcting for the composite executive function score for the EFIT disgust  $d'$  ( $t=2.536$ ,  $p=0.013$ ) but not EFIT angry  $d'$  ( $t=1.629$ ,



$p=0.11$ ). The group differences for the EFIT anger bias score persisted when correcting for subject educational level, VCI score or the composite executive function score ( $p<0.05$ ) but not the BIS 11 score ( $p=0.10$ ) or the CTQ score ( $p=0.38$ ). The group differences in the DPT negativity bias score stayed significant when controlling for the VCI score ( $t=2.088$ ;  $p=0.04$ ) and reached trend-significance when controlling for the composite executive function score ( $t=1.916$ ,  $p=0.059$ ) but did not survive correction for subject educational level, BIS 11 or CTQ ( $p>0.13$ ).

We did not observe significant interactions ( $p>0.05$ ) for group by subject educational level, group by VCI score, group by composite executive function score, group by BIS 11 or group by CTQ with the EFIT angry  $d'$ , EFIT disgust  $d'$ , the EFIT anger bias score or the DPT negativity bias score, except a significant group by CTQ score interaction with the EFIT anger bias score ( $t=-2.059$ ,  $\beta=-0.210$ ,  $p=0.043$ , Figure 3A). Exploratory within-group analyses showed that the CTQ score was positively related to the EFIT anger bias score in patients ( $t=2.345$ ,  $\beta=0.345$ ,  $p=0.025$ ) with no apparent relationship in controls ( $p=0.6$ ).

Whole-group analyses showed that the EFIT angry  $d'$  was negatively associated with the CTQ score ( $t=-2.746$ ,  $\beta=-0.261$ ,  $p=0.007$ , Figure 3C) and the BIS 11 score ( $t=-2.333$ ,  $\beta=-0.224$ ,  $p=0.022$ , Figure 3D) but neither with subject educational level ( $t=1.613$ ,  $\beta=0.180$ ,  $p=0.11$ ), the VCI score ( $t=1.903$ ,  $\beta=0.184$ ,  $p=0.06$ ) or the composite executive function score ( $t=1.771$ ,  $\beta=0.169$ ,  $p=0.080$ ). EFIT disgust  $d'$  showed a positive relationship with subject educational level ( $t=2.954$ ,  $\beta=0.328$ ,  $p=0.004$ ), the VCI score ( $t=3.731$ ,  $\beta=0.352$ ,  $p=0.0003$ ), and a negative association with the CTQ score ( $t=-4.595$ ,  $\beta=-0.419$ ,  $p<0.0001$ , Figure 3E) and the BIS 11 score ( $t=-4.908$ ,  $\beta=-0.442$ ,  $p<0.0001$ , Figure 3F) but not the composite executive function score ( $t=1.720$ ,  $\beta=0.169$ ,  $p=0.089$ ). Whole-group analyses showed that across patients and controls, the EFIT anger bias score displayed a positive association with the BIS 11 score ( $t=2.054$ ,

$\beta=0.208$ ,  $p=0.043$ , Figure 3B) but not with subject educational level, VCI or the composite executive function score ( $p>0.11$ ). Across patients and controls, the difference between the fearful and happy reaction time bias scores for masked facial expressions on the DPT was associated with subject educational level ( $t=2.474$ ,  $\beta=0.286$ ,  $p=0.015$ ) but not the VCI score, the composite executive function score, the BIS 11 score or the CTQ score ( $p>0.12$ ).

In patients, the ZAN-BPD score was associated with EFIT angry  $d'$  ( $t=-2.155$ ,  $\beta=-0.316$ ,  $p=0.038$ , Figure 4A), the EFIT anger bias score ( $t=2.027$ ,  $\beta=0.327$ ,  $p=0.050$ , Figure 4C) but not the EFIT disgust  $d'$  ( $p=0.087$ ) or the DPT masked negativity bias score ( $p=0.99$ ). The SCL-90 GSI was negatively associated with the EFIT angry  $d'$  ( $t=-2.485$ ,  $\beta=-0.332$ ,  $p=0.018$ , Figure 4B), and positively with the EFIT anger bias score ( $t=2.158$ ,  $\beta=0.322$ ,  $p=0.038$ ; Figure 4D) but not with EFIT disgust  $d'$  ( $t=-1.202$ ,  $\beta=-0.177$ ,  $p=0.24$ ) the DPT masked negativity bias score ( $p=0.59$ ). The HAM-D showed no relationship with neither the EFIT angry  $d'$ , EFIT disgust  $d'$ , EFIT anger bias score or the DPT masked negativity bias score ( $p>0.3$ ).

## Discussion

As hypothesized, adult women with BPD displayed a reduced ability to identify emotional facial expressions, most notably for angry and disgusted emotions, paralleled by a heightened negativity bias for anger, and a preconscious negativity bias for fearful relative to happy facial expressions compared to controls. Notably, a higher frequency of psychopathological and borderline symptoms in patients was coupled with impaired identification of angry facial expressions and a stronger anger bias. Reported memories of experienced childhood adversity were positively associated with a heightened anger bias in women with BPD, while lowered ability to identify angry and disgusted face emotions were linked to more childhood adversity across patients and controls. Finally,

higher anger bias and reduced identification of angry and disgusted facial expressions was related to more trait impulsivity across the entire cohort.

The results of the present study suggest that BPD is associated with deficits in the identification of facial expressions. Although patients and controls did not appear to differ in their ability to identify specific face emotions, post hoc analyses did indicate that the impaired identification of facial expressions in patients was most pronounced for angry and disgusted faces in agreement with recent meta-analyses (Daros et al., 2013; Mitchell et al., 2014). However, BPD patients also showed a tendency to identify surprised and happy faces with less accuracy. Consistent with earlier studies, the female borderline patients displayed a heightened negativity bias for face emotions (Catalan et al., 2016; Domes et al., 2008; Dyck et al., 2009; Veague & Hooley, 2014). Because the EFIT  $d'$  and bias variables were based on estimates from the same task, the apparent deficits in BPD patients to identify expressed face emotions of disgust and anger appear to go hand in hand with a heightened anger bias. However, the anger bias score in the female BPD patients did not significantly differ from their overall tendency to label facial expressions with the wrong emotional category.

The heightened vigilance to fearful facial expressions and reduced vigilance to happy facial expressions in patients seemed to occur at a preconscious stage. Specifically, the borderline patients were in average faster on trials when fearful facial expressions had to be approached rather than ignored while the opposite response style was apparent on trials that included happy facial expressions, in which patients were slower to approach rather than disengage their attention from happy faces. The heightened negative vigilance bias in BPD patients is consistent with some previous observations (Jovev et al., 2012; von Ceumern-Lindenstjerna et al., 2010) but not (Donges et al., 2015). We note however, that the group difference in fear relative to happy bias scores on the dot-probe task did not survive correction for any of the

matching variables or control measures reported in Table 1, except the verbal comprehension index. Furthermore, the fearful relative to the happy reaction time bias showed no apparent linear relationship with the control measures or the symptom scales and we therefore refrained from making further inferences about the apparent group differences for the dot-probe task.

Neither the lowered ability to identify angry and disgusted facial expressions nor the heightened anger bias is specific to BPD. For instance, it has been shown that patients who suffer from panic disorder had lowered identification of faces that expressed disgust, sadness or anger concurrent with a negativity bias for anger (Kessler, Roth, von Wietersheim, Deighton, & Traue, 2007), and that subjects diagnosed with social anxiety disorder displayed a bias for identification of anger (Mohlman, Carmin, & Price, 2007). Obsessive-compulsive disorder has previously been associated with a decreased ability to identify disgusted facial expressions (Corcoran, Woody, & Tolin, 2008), antisocial violent offenders were shown to display a tendency to interpret ambiguous facial expressions as hostile (Schonenberg & Jusyte, 2014) while adults diagnosed with first episode psychosis had reduced identification of angry facial expressions (Catalan et al., 2016). Finally, a meta-analysis reported that anxiety and major depression are associated with lower emotional face identification ability (Demenescu, Kortekaas, den Boer, & Aleman, 2010). In the present study, the effect sizes for the group differences in the ability to identify angry and disgusted face emotions and anger bias were modest with group means differing about a half standard deviation from each other. However, multiple linear regression analyses showed that patients who scored higher on the ZAN-BPD scale and SCL-90 global severity scale had a stronger anger bias and lower ability in the identification of angry and disgusted facial expressions. Therefore, BPD patients who misperceived aversive face emotions the most appeared to display more borderline symptoms and global psychopathology.

Thus, the emotional face identification task might provide critical clues to the specific emotional and cognitive vulnerabilities that underlie the symptoms experienced by the individual patient.

After correcting for subject educational level or the verbal comprehension index score, group differences persisted for the anger bias score but not for the angry or disgusted  $d'$  on the EFIT. While we matched patients and controls on parental educational level in an effort to control for socioeconomic background, the patients themselves had significantly lower education levels compared to controls. It is commonly observed that patients with BPD have shorter educations than healthy controls (Jovev et al., 2011; Matzke et al., 2014) and more borderline traits in adolescence are predictive of lower academic attainment later in life (Bagge et al., 2004; Winograd, Cohen, & Chen, 2008). This is consistent with our finding that better identification of disgusted face emotions was associated with higher educational levels across the entire cohort. Moreover, because the verbal comprehension index score in the entire cohort showed a positive linear relationship with disgusted EFIT  $d'$ , the ability to identify emotional facial expressions is perhaps coupled to better intellectual functioning in healthy individuals as well as borderline patients.

Consistent with previous studies, patients scored significantly higher on traits of impulsivity assessed with the BIS 11 scale relative to controls (Domes et al., 2008; Jacob et al., 2010; Kobeleva et al., 2014). Impulsivity is considered a core symptom in BPD and is associated with poor emotion regulation (Sebastian, Jacob, Lieb, & Tuschler, 2013), more aggressive behaviour and tendencies to self-mutilation (Terzi et al., 2017). Group differences in the ability to identify face emotions and anger bias did not persist when correcting for individual variations in trait impulsivity. Interestingly however, a higher BIS 11 score was associated with lower identification ability of angry and disgusted facial expressions and a heightened anger bias across patients and controls.

Our results extend the findings from previous studies, which reported that more trait impulsivity in healthy adults was associated with increased sensitivity to fearful faces in aversive scenarios (Lee et al., 2012), and that more borderline traits in otherwise healthy individuals were coupled with reduced identification of angry faces, only for those subjects who also displayed high levels of impulsivity (Preti et al., 2016). As such, the results indicate that anger bias and ability to identify aversive facial expressions are linked to traits of impulsivity not just in patients but also in healthy individuals.

As previously reported (Thomsen et al., 2016), and in agreement with earlier studies (Carvalho Fernando et al., 2014; Nicol et al., 2013), patients diagnosed with BPD reported more adverse experiences during childhood assessed with the CTQ compared to controls. Recently it was shown that borderline patients who subjectively evaluated faces as being more unapproachable had more childhood trauma experiences relative to controls (Nicol et al., 2013), and that lowered ability to identify facial expressions, particularly disgusted faces, was associated with higher levels of childhood trauma experiences (Nicol et al., 2014). Likewise, we observed that a higher CTQ score was related to heightened anger bias in BPD patients but not controls. However, neither the anger bias score nor the identification ability of angry and disgusted facial expressions differed significantly between patients and controls when correcting for inter-individual variations in self-reported childhood adversity, which therefore seems to mediate the group differences. Additionally, more childhood adversity reported by patients and controls was associated with lowered identification of angry and disgusted faces. Experienced adversity in childhood may thus predispose individuals to misperceive facial expressions as aversive later in life. This notion is substantiated by evidence that the ability to identify the emotional valence in facial expressions undergoes prolonged development from early childhood through adolescence (Mancini, Agnoli, Baldaro, Bitti, & Surcinelli, 2013; Thomas, De Bellis, Graham, & LaBar,

2007).

A major unresolved issue concerns the apparent inconsistencies among studies on emotional face identification in BPD. One reason for the seemingly large discrepancies between results is probably related to the use of different task designs including the emotional categories examined, how responses are recorded, the sampling size of the task stimuli, and how outcome measures are estimated across studies. Moreover, ceiling effects, particularly for happy facial expressions, often tamper firm interpretations of the results (Catalan et al., 2016; Domes et al., 2008; Dyck et al., 2009; Unoka et al., 2011). Similarly, the trend for lower identification of happy facial expressions in the borderline patients of the present study may partly be influenced by lack of variance. Although psychopathology were assessed with the SCL-90-R and HAM-D in the healthy controls, we refrained from doing linear regression analyses that included SCL-90-R or HAM-D outcome measures obtained from controls because of substantial floor effects. This is not surprising since the SCL-90-R and HAM-D were developed to assess the degree of psychopathology in psychiatric populations, and the lack of variability on the HAM-D in healthy controls has previously been noted (Fertuck et al., 2013). Although patients did differ significantly from controls on the dot-probe task, it is possible that effect sizes would have been larger if we had included angry facial expressions in the task because a negativity bias for anger appears to be a more consistent finding in BPD than fear (Daros et al., 2013; Mitchell et al., 2014).

In conclusion, relative to healthy controls, adult women diagnosed with BPD had impaired identification of angry and disgusted facial expressions paralleled by a heightened negativity bias. Increased severity in global psychopathology and borderline symptoms in patients showed a linear relationship with impaired identification of angry facial expressions and a stronger bias to misperceive faces as angry. While atypical reading of emotional facial expressions is not a specific to BPD, our findings indicate

that the BPD patients who do misperceive the social cues signaled by facial expressions may have the greatest mental health issues.

## References

- APA. (2013). American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th edition). *Arlington, VA: American Psychiatric Publishing.*
- Baer, R. A., Peters, J. R., Eisenlohr-Moul, T. A., Geiger, P. J., & Sauer, S. E. (2012). Emotion-related cognitive processes in borderline personality disorder: a review of the empirical literature. *Clin Psychol Rev*, 32(5), 359-369. doi:10.1016/j.cpr.2012.03.002
- Bagge, C., Nickell, A., Stepp, S., Durrett, C., Jackson, K., & Trull, T. J. (2004). Borderline personality disorder features predict negative outcomes 2 years later. *J Abnorm Psychol*, 113(2), 279-288. doi:10.1037/0021-843X.113.2.279
- Bernstein, D. P., & Fink, L. (1998). Childhood Trauma Questionnaire: A retrospective self-report manual. *San Antonio, TX: The Psychological Corporation.*
- Bhagwagar, Z., Cowen, P. J., Goodwin, G. M., & Harmer, C. J. (2004). Normalization of enhanced fear recognition by acute SSRI treatment in subjects with a previous history of depression. *Am J Psychiatry*, 161(1), 166-168. doi:10.1176/appi.ajp.161.1.166
- Carvalho Fernando, S., Beblo, T., Schlosser, N., Terfehr, K., Otte, C., Lowe, B., . . . Wingenfeld, K. (2014). The impact of self-reported childhood trauma on emotion regulation in borderline personality disorder and major depression. *J Trauma Dissociation*, 15(4), 384-401. doi:10.1080/15299732.2013.863262
- Catalan, A., Gonzalez de Artaza, M., Bustamante, S., Orgaz, P., Osa, L., Angosto, V., . . . Gonzalez-Torres, M. A. (2016). Differences in Facial Emotion Recognition between First Episode Psychosis, Borderline Personality Disorder and Healthy Controls. *Plos One*, 11(7), e0160056. doi:10.1371/journal.pone.0160056
- Corcoran, K. M., Woody, S. R., & Tolin, D. F. (2008). Recognition of facial expressions in obsessive-compulsive disorder. *J Anxiety Disord*, 22(1), 56-66. doi:10.1016/j.janxdis.2007.01.003
- Daros, A. R., Zakzanis, K. K., & Ruocco, A. C. (2013). Facial emotion recognition in borderline personality disorder. *Psychol Med*, 43(9), 1953-1963. doi:10.1017/S0033291712002607
- Demenescu, L. R., Kortekaas, R., den Boer, J. A., & Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *Plos One*, 5(12), e15058. doi:10.1371/journal.pone.0015058
- Derogatis, L. R. (1983). SCL-90-R: Administration, scoring, and procedures manual—II. *Towson, MD: Clinical Psychometric Research.*
- Domes, G., Czeschnek, D., Weidler, F., Berger, C., Fast, K., & Herpertz, S. C. (2008). Recognition of facial affect in Borderline Personality Disorder. *J Pers Disord*, 22(2), 135-147. doi:10.1521/pedi.2008.22.2.135



- Donges, U. S., Dukalski, B., Kersting, A., & Suslow, T. (2015). Automatic processing of facial affects in patients with borderline personality disorder: associations with symptomatology and comorbid disorders. *Ann Gen Psychiatry, 14*, 20. doi:10.1186/s12991-015-0058-y
- Dyck, M., Habel, U., Slodczyk, J., Schlummer, J., Backes, V., Schneider, F., & Reske, M. (2009). Negative bias in fast emotion discrimination in borderline personality disorder. *Psychol Med, 39*(5), 855-864. doi:10.1017/S0033291708004273
- Ekman, P., & Friesen, W. V. (1976). Pictures of Facial Affect. *Consulting Psychologists Press, Palo Alto, CA*.
- Fertuck, E. A., Grinband, J., & Stanley, B. (2013). Facial trust appraisal negatively biased in borderline personality disorder. *Psychiatry Res, 207*(3), 195-202. doi:10.1016/j.psychres.2013.01.004
- First, M. B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S. (1997). Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) *American Psychiatric Press, Washington D.C*.
- Hamilton, M. (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry, 23*, 56-62. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14399272>
- Herpertz, S. C., & Bertsch, K. (2014). The social-cognitive basis of personality disorders. *Curr Opin Psychiatry, 27*(1), 73-77. doi:10.1097/YCO.0000000000000026
- Jacob, G. A., Gutz, L., Bader, K., Lieb, K., Tuscher, O., & Stahl, C. (2010). Impulsivity in borderline personality disorder: impairment in self-report measures, but not behavioral inhibition. *Psychopathology, 43*(3), 180-188. doi:10.1159/000304174
- Jovev, M., Chanen, A., Green, M., Cotton, S., Proffitt, T., Coltheart, M., & Jackson, H. (2011). Emotional sensitivity in youth with borderline personality pathology. *Psychiatry Res, 187*(1-2), 234-240. doi:10.1016/j.psychres.2010.12.019
- Jovev, M., Green, M., Chanen, A., Cotton, S., Coltheart, M., & Jackson, H. (2012). Attentional processes and responding to affective faces in youth with borderline personality features. *Psychiatry Res, 199*(1), 44-50. doi:10.1016/j.psychres.2012.03.027
- Kessler, H., Roth, J., von Wietersheim, J., Deighton, R. M., & Traue, H. C. (2007). Emotion recognition patterns in patients with panic disorder. *Depress Anxiety, 24*(3), 223-226. doi:10.1002/da.20223
- Kobeleva, X., Seidel, E. M., Kohler, C., Schneider, F., Habel, U., & Derntl, B. (2014). Dissociation of explicit and implicit measures of the behavioral inhibition and activation system in borderline personality disorder. *Psychiatry Res, 218*(1-2), 134-142. doi:10.1016/j.psychres.2014.04.027
- Lazarus, S. A., Cheavens, J. S., Festa, F., & Rosenthal, Z. M. (2014). Interpersonal functioning in borderline personality disorder: a systematic review of behavioral and laboratory-based assessments. *Clin Psychol Rev, 34*(3), 193-205. doi:10.1016/j.cpr.2014.01.007
- Lee, T. H., Choi, J. S., & Cho, Y. S. (2012). Context modulation of facial emotion perception differed by individual difference. *Plos One, 7*(3), e32987. doi:10.1371/journal.pone.0032987
- Mancini, G., Agnoli, S., Baldaro, B., Bitti, P. E., & Surcinelli, P. (2013). Facial expressions of emotions: recognition accuracy and affective reactions during late childhood. *J Psychol, 147*(6), 599-617. doi:10.1080/00223980.2012.727891

- Matzke, B., Herpertz, S. C., Berger, C., Fleischer, M., & Domes, G. (2014). Facial reactions during emotion recognition in borderline personality disorder: a facial electromyography study. *Psychopathology*, 47(2), 101-110. doi:10.1159/000351122
- Mitchell, A. E., Dickens, G. L., & Picchioni, M. M. (2014). Facial emotion processing in borderline personality disorder: a systematic review and meta-analysis. *Neuropsychol Rev*, 24(2), 166-184. doi:10.1007/s11065-014-9254-9
- Mohlman, J., Carmin, C. N., & Price, R. B. (2007). Jumping to interpretations: social anxiety disorder and the identification of emotional facial expressions. *Behav Res Ther*, 45(3), 591-599. doi:10.1016/j.brat.2006.03.007
- Murphy, S. E., Downham, C., Cowen, P. J., & Harmer, C. J. (2008). Direct effects of diazepam on emotional processing in healthy volunteers. *Psychopharmacology (Berl)*, 199(4), 503-513. doi:10.1007/s00213-008-1082-2
- Nicol, K., Pope, M., & Hall, J. (2014). Facial emotion recognition in borderline personality: an association, with childhood experience. *Psychiatry Res*, 218(1-2), 256-258. doi:10.1016/j.psychres.2014.04.017
- Nicol, K., Pope, M., Sprengelmeyer, R., Young, A. W., & Hall, J. (2013). Social judgement in borderline personality disorder. *Plos One*, 8(11), e73440. doi:10.1371/journal.pone.0073440
- Preti, E., Richetin, J., Suttora, C., & Pisani, A. (2016). Individual differences in components of impulsivity and effortful control moderate the relation between borderline personality disorder traits and emotion recognition in a sample of university students. *Psychiatry Res*, 238, 109-115. doi:10.1016/j.psychres.2016.02.037
- Robin, M., Pham-Scottez, A., Curt, F., Dugre-Le Bigre, C., Speranza, M., Sapinho, D., . . . Kedia, G. (2012). Decreased sensitivity to facial emotions in adolescents with Borderline Personality Disorder. *Psychiatry Res*, 200(2-3), 417-421. doi:10.1016/j.psychres.2012.03.032
- Schonenberg, M., & Jusyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *Eur Arch Psychiatry Clin Neurosci*, 264(1), 61-69. doi:10.1007/s00406-013-0440-1
- Sebastian, A., Jacob, G., Lieb, K., & Tuscher, O. (2013). Impulsivity in borderline personality disorder: a matter of disturbed impulse control or a facet of emotional dysregulation? *Curr Psychiatry Rep*, 15(2), 339. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/23424747>
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (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. *J Clin Psychiatry*, 59 Suppl 20, 22-33;quiz 34-57. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9881538>
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behav Res Methods Instrum Comput*, 31(1), 137-149. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10495845>
- Terzi, L., Martino, F., Berardi, D., Bortolotti, B., Sasdelli, A., & Menchetti, M. (2017). Aggressive behavior and self-harm in Borderline Personality Disorder: The role of impulsivity and emotion dysregulation in a sample of outpatients. *Psychiatry Res*, 249, 321-326. doi:10.1016/j.psychres.2017.01.011

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, 10(5), 547-558. doi:10.1111/J.1467-7687.2007.00614.X

Thome, J., Liebke, L., Bungert, M., Schmahl, C., Domes, G., Bohus, M., & Lis, S. (2016). Confidence in facial emotion recognition in borderline personality disorder. *Personal Disord*, 7(2), 159-168. doi:10.1037/per0000142

Thomsen, M. S., Ruocco, A. C., Carcone, D., Mathiesen, B. B., & Simonsen, E. (2016). Neurocognitive Deficits in Borderline Personality Disorder: Associations With Childhood Trauma and Dimensions of Personality Psychopathology. *J Pers Disord*, 1-19. doi:10.1521/pedi\_2016\_30\_265

Tragesser, S. L., Solhan, M., Schwartz-Mette, R., & Trull, T. J. (2007). The role of affective instability and impulsivity in predicting future BPD features. *J Pers Disord*, 21(6), 603-614. doi:10.1521/pedi.2007.21.6.603

Unoka, Z., Fogd, D., Fuzy, M., & Csukly, G. (2011). Misreading the facial signs: specific impairments and error patterns in recognition of facial emotions with negative valence in borderline personality disorder. *Psychiatry Res*, 189(3), 419-425. doi:10.1016/j.psychres.2011.02.010

Veague, H. B., & Hooley, J. M. (2014). Enhanced sensitivity and response bias for male anger in women with borderline personality disorder. *Psychiatry Res*, 215(3), 687-693. doi:10.1016/j.psychres.2013.12.045

von Ceumern-Lindenstjerna, I. A., Brunner, R., Parzer, P., Mundt, C., Fiedler, P., & Resch, F. (2010). Attentional bias in later stages of emotional information processing in female adolescents with borderline personality disorder. *Psychopathology*, 43(1), 25-32. doi:10.1159/000255960

Winograd, G., Cohen, P., & Chen, H. (2008). Adolescent borderline symptoms in the community: prognosis for functioning over 20 years. *J Child Psychol Psychiatry*, 49(9), 933-941. doi:10.1111/j.1469-7610.2008.01930.x

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, 63(3), 271-313. doi:10.1016/S0010-0277(97)00003-6

Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD): a continuous measure of DSM-IV borderline psychopathology. *J Pers Disord*, 17(3), 233-242. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12839102>

**Figure 1.** The plots in (A) display the outcome measures for the emotional face identification task (EFIT)  $d'$  while the bar plots in (B) show the outcome measures for the EFIT bias scores. The mean and standard deviation is shown for each outcome variable coded in black for the patients and in grey for the controls. Note that the displayed values for the bias scores in (B) were normalized using Rankit transformation.

**Figure 2.** The plot shows the vigilance bias scores on the dot-probe task (DPT) for masked fearful and masked happy faces. The mean and standard deviation is shown for each bias score coded in black for the patients and in grey for the controls. Note that the displayed values for the bias scores were normalized using Rankit transformation.

**Figure 3.** The partial regression plots display whole-group analyses for the emotional face identification task (EFIT) angry bias score predicted the interaction term for group by the CTQ score (A), whole-group analyses for the EFIT angry bias score predicted by the BIS 11 score (B), whole-group analyses for the EFIT angry  $d'$  predicted by by CTQ score (C), whole-group analyses for the EFIT angry  $d'$  predicted by the BIS 11 score (D), whole-group analyses for the EFIT disgust  $d'$  predicted by by CTQ score (E), and whole-group analyses for the EFIT disgust  $d'$  predicted by the BIS 11 score (F).

. Note that different regression lines are shown for the controls and patients in (A) because the plot displays the significant interaction effect for the group by CTQ score. The patients are coded in black and controls are coded in grey. The values displayed on the Y-axes and X-axes are the regression model residuals. The plots are corrected for age and parent educational level. The dotted lines correspond to the 95% confidence interval of the group mean. Note that in (A) the 95% confidence interval of the group mean is shown for the patients but not controls for the ease of visual interpretation.

**Figure 4.** The partial regression plots display within-group analyses of the patients coded in black for the emotional face identification task (EFIT) angry  $d'$  predicted by the ZAN-BPD score (A), EFIT angry  $d'$  predicted by the SCL-90 GSI score (B), EFIT anger bias score predicted by the ZAN-BPD score (C) and EFIT anger bias score predicted by the SCL-90 GSI score (D). The values displayed on the Y-axes and X-axes are the regression model residuals. The plots are corrected for age and parent educational level. The dotted lines correspond to the 95% confidence interval of the group mean.

**Table 1.** The table displays the matching variables, control measures and psychopathology scales for the patient group and control group, as well as the symptom scales for the patients.

	Patients (N=43)	Controls (N=56)
	<b>Matching variables</b>	
Age <sup>b</sup>	26.0 (23.0   33.1)	26.0 (22.1   33.1)
Parental education <sup>a</sup>	2.47 ± 0.92 <sup>(-4)</sup>	2.70 ± 0.73
	<b>Control measures</b>	
Subject education <sup>b</sup>	2.0 (1.0   4.0) <sup>(-2)</sup>	3.5 (3.0   5.0)
WAIS VCI <sup>a</sup>	93.4 ± 10.0	102.6 ± 8.1
Composite executive function score <sup>b</sup>	87.2 (73.8   98.7)	107.2 (95.7   122.4)
BIS 11 <sup>a</sup>	73.4 ± 11.6 <sup>(-3)</sup>	55.1 ± 8.4 <sup>(-1)</sup>
CTQ <sup>b</sup>	60.5 (46.3   68.5) <sup>(-5)</sup>	29.0 (25.0   35.0) <sup>(-1)</sup>
	<b>Symptom scales</b>	
SCL-90 GSI <sup>b</sup>	1.6 (1.3   2.2) <sup>(-5)</sup>	0.2 (0.1   0.3) <sup>(-1)</sup>
HAM-D <sup>b</sup>	14.0 (10.3   17.8) <sup>(-3)</sup>	0.0 (0.0   2.0) <sup>(-1)</sup>
ZAN-BPD <sup>a</sup>	13.7 ± 7.1	-

	<b>Psychiatric Diagnostic Comorbidity</b>	
	N (%)	
Major Depressive Disorder	4 (9.3)	0 (0.0)
Bipolar II Disorder	0 (0.0)	0 (0.0)
Agoraphobia	9 (20.9)	0 (0.0)
Panic Disorder without Agoraphobia	8 (18.6)	0 (0.0)
Social Phobia	13 (30.2)	0 (0.0)
General Anxiety Disorder	7 (16.3)	0 (0.0)
Obsessive Compulsive Disorder	2 (4.7)	0 (0.0)
Posttraumatic Stress Disorder	7 (16.3)	0 (0.0)
	<b>Personality Disorders</b>	
	N (%)	
Avoidant	19 (44.2)	0 (0.0)
Dependent	6 (14.0)	0 (0.0)
Obsessive Compulsive	8 (18.6)	0 (0.0)
Paranoid	8 (18.6)	0 (0.0)
Schizotypal	0 (0.0)	0 (0.0)
Schizoid	0 (0.0)	0 (0.0)
Histrionic	2 (4.7)	0 (0.0)
Narcissistic	2 (4.7)	0 (0.0)
Antisocial	4 (9.3)	0 (0.0)

Data <sup>(a)</sup> shown with mean  $\pm$  standard deviations appeared normally distributed while data <sup>(b)</sup> reported with medians and lower and upper quartiles significantly deviated from the normal distribution. Note that the non-normally distributed variables were normalized with Rankit transformation for the linear analyses. Note that missing data for a variable is displayed with a superscript encapsulated in hyphens <sup>(-n)</sup> in the descriptive statistic cells for the patients and controls, respectively. Missing data is also described in the method section.

**Table 2.** The table displays the descriptive statistics for the emotional face identification task (EFIT) and dot-probe task (DPT) outcome measures for the patient group and control group.

	Patients (N=43)	Controls (N=56)
	<b>Outcome measures</b>	
EFIT angry $d'$	1.53 $\pm$ 0.38	1.68 $\pm$ 0.37
EFIT disgust $d'$	1.39 $\pm$ 0.67	1.71 $\pm$ 0.59
EFIT fear $d'$	1.75 $\pm$ 0.56	1.89 $\pm$ 0.61
EFIT happy $d'$	2.63 $\pm$ 0.37	2.76 $\pm$ 0.30
EFIT sad $d'$	1.80 $\pm$ 0.54	1.87 $\pm$ 0.51
EFIT surprise $d'$	1.83 $\pm$ 0.40	1.98 $\pm$ 0.36
EFIT angry bias	14.3 (9.5   21.1)	10.9 (6.3   16.2)
EFIT disgust bias	9.5 (5.8   23.8)	9.5 (4.0   13.6)
EFIT fear bias	6.3 (4.5   10.7)	4.8 (2.4   7.9)

EFIT happy bias	1.0 (0.7   2.0)	0.8 (0.4   2.2)
EFIT sad bias	7.9 (3.0   12.7)	6.3 (3.0   10.5)
EFIT surprise bias	8.6 (5.2   13.0)	6.6 (4.6   11.4)
DPT masked fear bias (ms)	2.2 (-1.0   4.0)	0.9 (-1.6   3.1)
DPT masked happy bias (ms)	-0.9 (-4.4   1.7)	-0.3 (-1.7   2.2)
DPT unmasked fear bias (ms)	0.8 (-2.9   3.4)	-0.6 (-2.5   4.1)
DPT unmasked happy bias (ms)	0.6 (-2.2   3.2)	-0.9 (-3.9   1.0)

Data <sup>(a)</sup> shown with mean  $\pm$  standard deviations appeared normally distributed while data <sup>(b)</sup> reported with medians and lower and upper quartiles significantly deviated from the normal distribution. Note that the non-normally distributed variables were normalized with Rankit transformation for the linear analyses. Note that five controls had missing data for the EFIT.