Research Paper

Altered amygdalar emotion space in borderline personality disorder normalizes following dialectical behaviour therapy

Seth M. Levine*, PhD; Katharina Merz*, MD; Daniel Keeser, PhD; Julia I. Kunz, MSc; Barbara B. Barton, PhD; Matthias A. Reinhard, MD, PhD; Andrea Jobst, MD; Frank Padberg, MD; Corinne Neukel, PhD; Sabine C. Herpertz, MD; Katja Bertsch†, PhD; Richard Musil†, MD

Background: Borderline personality disorder (BPD) is a mental health condition characterized by an inability to regulate emotions or accurately process the emotional states of others. Previous neuroimaging studies using classical univariate analyses have tied such emotion dysregulation to aberrant activity levels in the amygdala of patients with BPD. However, multivariate analyses have not yet been used to investigate how representational spaces of emotion information may be systematically altered in patients with BPD. **Methods:** Patients with BPD performed an emotional face matching task while undergoing MRI before and after a 10-week inpatient program of dialectical behavioural therapy. Representational similarity analysis (RSA) was applied to activity patterns (evoked by angry, fearful, neutral and surprised faces) in the amygdala and temporo-occipital fusiform gyrus of patients with BPD and in the amygdala of healthy controls. **Results:** We recruited 15 patients with BPD (8 females, 6 males, 1 transgender male) to participate in the study, and we obtained a neuroimaging data set for 25 healthy controls for a comparative analysis. The RSA of the amygdala revealed a negative bias in the underlying affective space (in that activity patterns evoked by angry, fearful and neutral faces were more similar to each other than to patterns evoked by surprised faces), which normalized after therapy. This bias-to-normalization effect was present neither in activity patterns of the temporo-occipital fusiform gyrus of patients nor in amygdalar activity patterns of healthy controls. **Limitations:** Larger samples and additional questionnaires would help to better characterize the association between specific aspects of therapy and changes in the neural representational space. **Conclusion:** Our findings suggest a more refined role for the amygdala in the pathological processing of perceived emotions and may provide new diagnostic and prognostic imaging-based markers of emotion dysregulation and personality disorders. **Clinical trial registration:** DRKS00019821, German Clinical Trials Register (Deutsches Register Klinischer Studien).

Introduction

Borderline personality disorder (BPD) is a severe mental disorder affecting about 1.7% of the population and 15%–28% of inpatients in psychiatric care.[1] It is characterized by a pattern of instability in affect, self-image and interpersonal relations as well as impulsivity, risk-taking behaviour and hostility.[2] According to a prominent theory,[3] emotion dysregulation is conceptualized as the core feature of BPD, rendering it a primary target for evidence-based interventions, such as dialectical behaviour therapy (DBT). Emotion dysregulation involves faster and elevated responses to stimuli, slower return to baseline, and fewer adaptive and more maladaptive regulation strategies.[4] In DBT, patients are trained to, for example, better differentiate their own emotions and decide which emotions are adaptive (and which are overbearing).[3] This form of therapy has been shown to have a mitigating effect on emotion dysregulation.[5]

Previous investigations into emotion processing of patients with BPD have shown heightened emotional sensitivity,[6] negativity biases[7] and altered processing of facial expressions compared with healthy individuals.[8] At the neurobiological level, univariate analyses of functional neuroimaging data from patients with BPD have implicated aberrant activity levels in the amygdala in altered emotion processing[9–11] — the consistency of which has been reported in meta-analyses[12,13] — whereas a normalization of such amygdalar activity has been reported following psychotherapy.[14,15] Such

Correspondence to: S. Levine, Department of Psychology, LMU Munich, Leopoldstraße 13, 80802 Munich, Germany; seth.levine@psy.lmu.de

*Share first authorship; †share senior authorship.

Submitted May 29, 2023; Revised July 24, 2023; Revised Aug. 28, 2023; Accepted Sept. 5, 2023

Cite as: *J Psychiatry Neurosci* 2023 November 7;48(6). doi: 10.1503/jpn.230085

neuroimaging findings have supported theories designating the amygdala as a key brain region in emotion regulation.[16]

However, univariate analyses of functional neuroimaging data have found limited success in generating reliable biomarkers of mental disorders.[17] To that end, the adoption of multivariate methods from cognitive neuroscience has attempted to close the explanatory gap between biological psychiatry and neuroscience. Representational similarity analysis (RSA),[18] a form of multivariate pattern analysis,[19] has only recently been used to examine how the cognitive structure of information is altered in different patient groups, such as individuals with posttraumatic stress disorder,[20] autism[21] and schizophrenia.[22] This method of analysis allows researchers to understand the relative informational content represented by multivariate activity patterns by using similarity metrics (e.g., Euclidean distance, correlation distance) to examine the association between such activity patterns evoked by distinct concepts.[23] This approach effectively places such concepts in a high-dimensional abstract space based on their association with each other, thereby allowing for a richer interpretation of mental representations,[24] thus rendering RSA more sensitive than traditional univariate analyses to meaningful variability of activity patterns in a given brain region.[25,26] To our knowledge, so far no study has used RSA to investigate these abstract spaces as they pertain to the organization of emotion information using neural activity patterns[27–29] specifically in individuals with BPD.

As such, in the present study, we sought to extend prior neuroimaging findings by using RSA to explore whether such neural "emotion spaces," measured using a classic perceptual matching task of emotional facial expressions,[30] show systematic alterations,[24] such as a higher degree of similarity among negative stimuli, before and after patients with BPD underwent a 10-week DBT program.

Methods

Participants

We recruited inpatients with BPD to this study as part of an ongoing trial for patients with BPD and patients with persistent depressive disorder at the Department of Psychiatry and Psychotherapy of the University Hospital LMU Munich.[31–33] All experimental procedures were approved by the ethics committee of the Faculty of Medicine of the LMU and complied with the Declaration of Helsinki following its most recent amendments. Participants provided written informed consent before participating in the study.

Clinical scales

At the beginning of treatment, patients were assessed for common comorbidities of BPD using the Structured Clinical Interview for DSM-5, Clinician Version (SCID-5-CV).[34,35] For clinical assessment, we administered the Beck Depression Inventory (BDI-II),[36,37] the 24-item Hamilton Depression Rating Scale (HAMD-24),[38,39] the Borderline Symptom List — Short Version (BSL-23),[40,41] the Borderline Personality Disorder Severity Index Version IV (BPDSI-IV)[42,43] and the short form of the Childhood Trauma Questionnaire (CTQ-SF)[44,45] at admission. After the 10-week treatment of DBT, patients completed a second administration of 1 or more of these assessments.

Experimental paradigm

While lying in the MRI scanner, participants performed a classic perceptual matching task[30] in which they were visually presented with alternating blocks of emotional faces (i.e., angry, fearful, neutral, surprised) or shapes (which served as the control condition). On a given trial, participants saw 3 stimuli (following the emotional theme of the current block) simultaneously — 1 at the top of the screen (the target stimulus) and 2 at the bottom of the screen — with the goal of determining which of the 2 stimuli at the bottom was identical to the target stimulus above. Each emotion block lasted 48 seconds and included 6 stimuli of a given emotional expression appearing in each block for 4 seconds with a variable interstimulus interval of 2–6 seconds. Each shape block lasted 36 seconds and included 6 stimuli of different shapes appearing in each block for 4 seconds with a fixed interstimulus interval of 2 seconds. Inter-block intervals were 12 seconds in duration. The run started and ended with a shape block. The task was administered using Presentation software, version 18.0 (Neurobehavioural Systems, Inc.), and stimuli were projected onto a screen that participants viewed using a mirror in the MRI scanner. Responses were provided via an MRI-compatible keypad (Current Designs).

Neuroimaging acquisition parameters

Neuroimaging data acquisition was carried out at the Neuroimaging Core Unit Munich (NICUM) of the LMU using a 3T Siemens Magnetom Prisma scanner with a 32-channel head coil (Siemens AG). Functional sequences consisted of 650 volumes acquired with a T_2*-weighted echo-planar imaging (EPI) sequence with the following parameters: 72 slices per volume in ascending interleaved order with multiband factor 8, voxel size 2 mm^3 isotropic, repetition time (TR) 800 ms, echo time (TE) 37 ms, flip angle 52°, field of view (FOV) 208 mm. The first 5 volumes of functional scans were dummy volumes to account for T_1-saturation and were discarded before image preprocessing. To coregister the functional images with the high-resolution anatomic images, 208 slices of T_1-weighted scans were acquired using a magnetization prepared rapid gradient-echo (MP-RAGE) sequence with the following parameters: voxel size 0.8 mm^3 isotropic, TR 2500 ms, TE 2.22 ms, flip angle 8°, FOV 256 mm).

Neuroimaging data analysis

Neuroimaging data were analyzed using SPM12, MATLAB R2020a (The Mathworks) and CoSMoMVPA.[46] Preprocessing of the neuroimaging data made use of default settings of the SPM12 preprocessing pipeline (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) and included spatially realigning the functional images to the mean image in the time series using a 6-parameter rigid body transformation and fourth degree b-spline interpolation, coregistering the functional images to a

given participant's T_1-weighted structural scan, normalizing the coregistered images to a standard 2-mm Montreal Neurological Institute (MNI) template using fourth degree b-spline interpolation, and spatially smoothing the images with an 8-mm full-width at half-maximum (FWHM) Gaussian kernel. Slice-timing correction was not performed, as the task was a block design.

The preprocessed functional images were analyzed using a general linear model (GLM) containing 1 regressor per condition. Regressors corresponding to the task blocks were modelled as box-car functions and convolved with a canonical hemodynamic response function. Motion-correction parameters were modelled as regressors of no interest in addition to a constant term.

Region of interest definition

As this experimental paradigm is known to activate the amygdala when contrasting face blocks with shape blocks,[30] we sought to determine whether the amygdala also systematically represents patterns of multivariate activity pertaining to emotion information; as such, we obtained a bilateral amygdala region of interest (ROI) from the probabilistic Harvard–Oxford atlas,[47] which was masked at a probability threshold of 0.8, yielding a region size of 209 voxels. Each participant's whole-brain *t*-scores (from the parameter estimates generated via the GLM) for the contrasts of interest (i.e., Anger > Shapes, Fear > Shapes, Neutral > Shapes, and Surprise > Shapes) were masked using this amygdala ROI for the RSA.

Representational similarity analysis

Using the Pearson correlation coefficient, the *t*-scores within the extracted voxels were correlated for each participant across conditions from the first time point and again, separately, from the second time point. This procedure yielded 6 correlation values per time point, which were visualized as correlation matrices. The pattern in the difference between the surprise stimuli and the other stimuli (in the first session) led us to investigate between-condition correlations within sessions, which in turn allowed us to compare these relative differences between sessions. To this end, for a given participant and a given time point, we separated correlation values that involved the surprise condition from correlations that did not involve the surprise condition and averaged these 2 sets independently (i.e., correlations between anger–fear, anger–neutral and fear–neutral were averaged together, and correlations between anger–surprise, fear–surprise and neutral–surprise were averaged together). This procedure yielded an "other v. surprise" analysis that we investigated before and after therapy (Fisher transforming all participants' averaged correlation values) via a 2-factor repeated-measures analysis of variance (ANOVA) using SPSS, version 28.0 (IBM Corp.). Statistical thresholds were set at an α level of 5%.

Regional control

To determine whether the multivariate findings were specific to the amygdala, we also obtained a bilateral ventrotemporal cortex ROI from the probabilistic Harvard–Oxford atlas, which is known to encode object categories.[48,49] The expectation was that this ROI would encode face information (and potentially the corresponding emotional expressions) but that there would be no systematic changes in the representational space following DBT. Specifically, we obtained a bilateral temporo-occipital fusiform gyrus ROI, which was masked at a probability threshold of 0.63 (yielding a region size of 217 voxels), so that the control ROI contained roughly the same number of voxels as the amygdala ROI. Here, we used a 3-factor repeated-measures ANOVA (region × emotion × time) to compare the results from the amygdala with those from the fusiform gyrus.

Healthy control group

In an additional follow-up analysis, we sought to determine whether the observed amygdala-specific effect was also specific to patients. To this end, we incorporated a neuroimaging data set from a sample of healthy controls who underwent the same emotion task in 2 separate sessions separated by a period of approximately 7 weeks. These neuroimaging data were acquired at Heidelberg University between 2016 and 2018 using a 3 T Siemens Tim Trio scanner (Siemens AG) with a 32-channel head coil. Functional sequences consisted of 40 transverse slices per volume acquired with a T_2*-weighted gradient EPI sequence (voxel size 2.3 mm^3 isotropic, TR 2340 ms, TE 26 ms, flip angle 80°, FOV 220 mm). To coregister the functional images to high-resolution anatomic images, structural scans were acquired using a T_1-weighted MP-RAGE sequence (voxel size 1 mm^3 isotropic, TR 1900 ms, TE 2.52 ms, flip angle 9°, FOV 256 mm). Further details about this data set have been published previously.[15] Preprocessing of the neuroimaging data from the healthy controls also made use of the default settings in SPM12 and likewise included realignment and coregistration to the mean functional image, spatial normalization to the standard 2-mm MNI template, and smoothing with an 8-mm FWHM Gaussian filter.

The same ROI analysis on the amygdala was carried out using the data set of the healthy controls, and a 3-factor mixed ANOVA (group × emotion × time) was used to compare the results from the patients with BPD with those from the healthy controls.

Follow-up correlation analyses

Given the findings from the similarity analysis, we wanted to explore whether there was any correlation between the changes in the amygdalar emotion space and other participant characteristics. To this end, the difference scores for each clinical scale (with the exception of the CTQ-SF, which was administered only once) were rank-correlated (using Kendall τ_b) with the interaction values from the activity patterns in the amygdala. In addition to the overall score of the BPDSI-IV, we also used the scores from the subscales for impulsivity (symptom area 4) and affective instability (symptom area 6), as these aspects of BPD have been associated with functionality of the amygdala.[50] Participants' ages were Pearson correlated with the interaction values. Corresponding *p* values were generated following 10 000 iterations of permutation testing.

Results

Participants

We initially recruited 21 inpatients between the ages of 19 and 54.5 years (12 females, 6 males, 2 transgender males, 1 unspecified; mean age 27 ± standard deviation [SD] 10 yr). Six patients between the ages of 19 and 33.6 years (4 females, 1 transgender male, 1 unspecified; mean age 22.7 ± SD 5.6 yr) did not participate in the second neuroimaging session, as they were either discharged early from the clinic or refused to participate in the second session. As such, full data sets for the remaining 15 participants between the ages of 19.8 and 54.5 years (8 females, 6 males, 1 transgender male; mean age 28.6 ± SD 11 yr) were included in the present analysis. The neuroimaging data set for the healthy control group came from a sample of 25 volunteers (18 females, 9 males; mean age 30.2 ± SD 7.8 yr).

Among the patients with BPD, 12 had a current major depressive episode, 8 patients were diagnosed with life-time PTSD (current symptomatology, $n = 5$), 5 patients had a life-time binge eating disorder (current symptomatology, $n = 1$) and 1 had a life-time (and current) bulimic eating disorder. After the 10-week treatment of DBT, 15 patients completed a second administration of the HAMD-24, 14 completed a second administration of the BPDSI-IV and 12 completed a second administration of the BDI-II and BSL-23 (Table 1 and Appendix 1, Tables 1–3, available at https://www.jpn.ca/lookup/doi/10.1503/jpn.230085/tab-related-content).

Similarity of amygdalar activity patterns reflects negative shift in emotion space that normalizes after therapy

Representational geometry of patients' emotion spaces within the amygdala (Figure 1A) showed a negative bias in the first session that was not detected in the second session (time × emotion: $F_{1,14} = 5.027$, $p = 0.042$; Figure 1D). Specifically, before DBT, activity patterns evoked by angry, fearful and neutral facial expressions showed a greater average similarity (i.e., higher correlation) to each other (i.e., "other pairs") than to facial expressions depicting surprise ($t_{14} = 2.805$, $p = 0.014$). Following DBT, this imbalance in the emotion space was no longer evident, as the representational geometry revealed a more uniform degree of similarity among the activity patterns ($t_{14} = 0.005$, $p > 0.99$).

Table 1: Descriptive statistics for the clinical scores

Clinical scale	Median (IQR)	*p* value*
BDI-II	–3.5 (–9.00 to 2.00)	0.054
BSL-23	0.025 (–0.52 to 0.57)	0.079
HAMD-24	–2.0 (–8.5 to 4.5)	0.11
BPDSI-IV	–5.57 (–12.39 to 1.25)	0.013
Impulsivity†	–0.64 (–1.275 to –0.005)	0.021
Affective instability‡	–0.1 (–1.7 to 1.5)	0.69
CTQ-SF§	48.0 (34.5 to 61.5)	–

BDI-II = Beck Depression Inventory; BPDSI-IV = Borderline Personality Disorder Severity Index Version IV; BSL-23 = Borderline Symptom List — Short Version; CTQ-SF = short form of the Childhood Trauma Questionnaire; HAMD-24 = 24-item Hamilton Depression Rating Scale; IQR = interquartile range.
*Wilcoxon signed-rank tests contrasting the median score of the 2 sessions.
†Fourth symptom area of the BPDSI-IV.
‡Sixth symptom area of the BPDSI-IV.
§CTQ-SF scores were acquired only during the first session and therefore do not reflect a difference score.

Emotion space in object-selective cortex remains relatively stable

To determine whether the systematic change in the emotion space was specific to the amygdala, we ran the same analysis in the temporo-occipital fusiform gyrus, knowing that face information is reportedly encoded by the ventrotemporal cortex. Here the representational geometry showed a dramatically higher overall degree of similarity between all facial expressions compared with that of the amygdala (region: $F_{1,14} = 29.995$, $p = 8.2 \times 10^{-5}$; Figure 1B). The interaction between emotions and time, as observed in the amygdala, also differed between regions (region × emotion × time: $F_{1,14} = 5.866$, $p = 0.03$), with no detectable evidence for such an interaction effect in the fusiform gyrus (emotion × time: $F_{1,14} = 0.174$, $p = 0.68$; Figure 1E).

Emotion spaces differ between patients and healthy controls

The last follow-up control analysis sought to determine whether the dynamic aspect of the emotion space underlying the amygdala was specific to patients with BPD, or whether time alone could explain this effect, in that a similar systematicity would be observable in healthy controls at 2 different points in time. To this end, we applied the same analysis to amygdalar voxels of healthy controls. There was an overall difference between the groups, in that pattern correlations of the patients with BPD tended to be higher than those of healthy controls (group: $F_{1,38} = 7.054$, $p = 0.011$; Figure 1C). More importantly, the previously reported interaction effect differed between the groups (group × emotion × time: $F_{1,38} = 5.184$, $p = 0.029$) and was not observed in the healthy controls (emotion × time: $F_{1,24} = 0.63$, $p = 0.80$; Figure 1F). Additionally, the emotion spaces of healthy controls did not show any systematic changes in terms of emotions ($F_{1,24} = 0.622$, $p = 0.44$) or time ($F_{1,24} = 3.685$, $p = 0.07$).

Correlation analyses

Rank-correlating the interaction values from the amygdalar activity patterns with the difference scores in the clinical scales showed a slight positive correlation between the fourth symptom area of the BPDSI-IV (i.e., impulsivity). Namely, decreasing pattern similarity of the facial expressions other than surprise (with respect to the changing pattern similarity of the surprised facial expressions; i.e., the interaction effect) corresponded to decreasing impulsivity scores ($\tau = 0.35$, $p = 0.03$; Figure 2). The remaining correlations for the BDI-II ($\tau = –0.11$, $p = 0.66$), BSL-23 ($\tau = –0.09$, $p = 0.64$), HAMD-24 ($\tau = 0.12$, $p = 0.26$), BPDSI-IV total ($\tau = 0.05$, $p = 0.38$), BPDSI-IV affective instability (sixth symptom area; $\tau = 0.01$, $p = 0.46$) and CTQ-SF total score ($\tau = –0.03$, $p = 0.56$) did not surpass the statistical threshold.

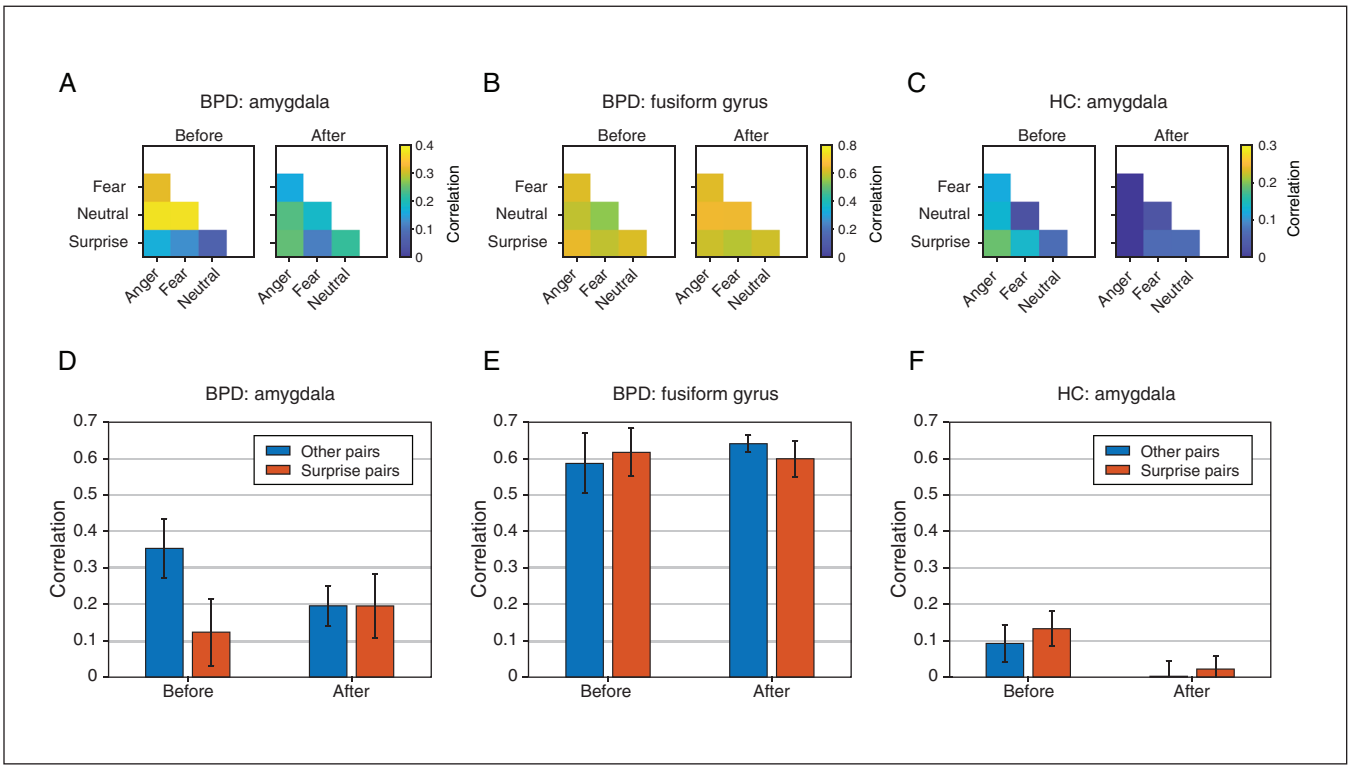

Figure 1: Results from the representational similarity analysis depicted as correlation matrices of the multivariate patterns evoked by the emotional facial expressions for both sessions in (A) the amygdala of patients with borderline personality disorder (BPD), (B) the fusiform gyrus of patients with BPD and (C) the amygdala of the healthy controls (HC). The amygdalar emotion space of the patients with BPD showed (D) a higher degree of similarity between angry, fearful and neutral expressions (blue bars) compared with the similarity of surprised expressions with the other facial expressions (red bars), which normalized in the second session after dialectical behaviour training (DBT; $F_{1,14} = 5.027$, $p = 0.042$). This interaction effect from (D) was observed neither in (E) the fusiform gyrus of patients with BPD ($F_{1,14} = 0.174$, $p = 0.68$) nor in (F) the amygdala of the healthy controls ($F_{1,24} = 0.63$, $p = 0.80$). Error bars represent standard errors of the mean.

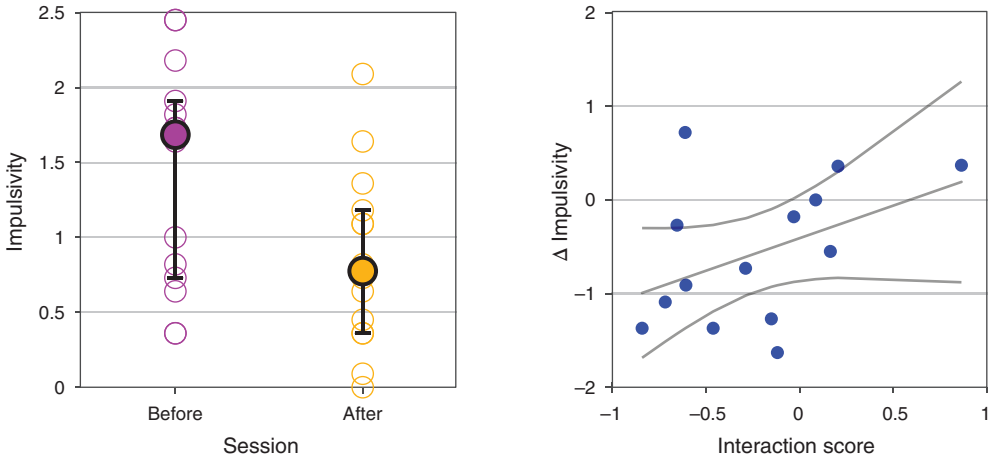

Figure 2: (Left) Changes in the impulsivity scores from the Borderline Personality Disorder Severity Index Version IV interview (median and interquartile range overlaid), and (right) their respective correlation (Kendall $\tau = 0.35$, $p = 0.03$) with the interaction in the multivariate pattern changes (reflecting alterations within the emotion space) in the amygdala. The line of best fit and its corresponding 95% confidence intervals are shown in grey.

Likewise, Pearson correlation of participants' ages with their interaction values from the amygdalar activity patterns yielded a negative correlation that also did not pass the statistical threshold ($r = –0.29$, $p = 0.15$). All p values presented here were generated through permutation testing and reflect 1-sided statistical tests. However, none of these correlations survived a correction for multiple comparisons; as such, these results should be considered exploratory.

Discussion

Emotion dysregulation is a core symptom of BPD,[51] and DBT focuses on this dysregulation by training patients to differentiate their emotions.[3] Functional neuroimaging studies of such emotion dysregulation in patients with BPD have used univariate analyses to consistently show altered activation levels of the amygdala,[13,12] also with respect to treatment programs incorporating DBT.[52] However, the use of multivariate pattern analysis opens up new avenues for interpreting the role of the amygdala in individuals with BPD, as RSA allows one to hypothesize not only about the involvement of a brain region, but also more specifically about the representational content underlying its activity patterns.[23]

As such, this study sought to provide a first look into high-dimensional neural representations of perceived emotions in patients with BPD. We combined RSA with functional MRI to investigate how the representational geometry of emotion information in the amygdala differs in patients with BPD before and after DBT. We found that, before therapy, there was an unusual negative shift in the representational space of perceived emotions in patients with BPD, in that neutral faces were represented more similarly to angry and fearful faces (leading to surprised faces being represented less similarly to the faces expressing other emotions). After therapy, this systematicity normalized, such that all representations of emotional expressions maintained a comparable degree of similarity to each other (i.e., the emotions were more evenly distributed across the representational space). This negatively shifted structure in the affective representational space of the amygdala is consistent with negativity biases observed in patients with BPD.[53] The fact that this negatively shifted structure was detected neither in the object-selective (i.e., ventrotemporal) cortex of patients with BPD nor in the amygdala of healthy controls allows for the interpretation that the amygdala may be playing a more fine-grained role in patients with BPD (i.e., beyond that of a hyperactive node in an emotion circuit), in that it processes neutral social cues more similarly to negative social cues.

Our findings are supported by those of prior studies showing that multivariate patterns in the amygdala reflect aversive learning,[54] subjective valence[55] and facial expressions.[56] Here we extend such work by showing that a diagnosis of BPD can also contribute to alterations in amygdalar affective spaces. The specificity of this finding in the amygdala, with respect to the fusiform gyrus, is also corroborated by previous work showing that changes in representational spaces following fear conditioning occurred in downstream regions involved in affective processing rather than in the object-selective cortex.[57–59]

Additionally, Puccetti and colleagues recently used RSA to show that a decreased persistence of the amygdala to represent negative information corresponded to higher psychological well-being.[60] This discovery is in line with our result that the negative bias in the amygdalar affective space normalized in patients with BPD following DBT, which raises the question of whether systematic variations in this space might be indicative of meaningful individual differences and have prognostic value. As such, the findings we present here offer a new perspective on the involvement of the amygdala in (pathologically) representing emotion information and may reflect a neural mechanism of emotion dysregulation that classically characterizes BPD.

Limitations

One of the primary methodological limitations of our study derives from the sample of healthy volunteers having not been specifically matched demographically to the patients with BPD in the current study and having been acquired on a different MRI scanner with different scanning parameters.[15] However, as we did not directly compare the groups, but rather within-group longitudinal changes between the groups, concerns regarding potential methodological biases are diminished. The presence of such biases would be expected if 1 set of scanning parameters was intrinsically more sensitive to detecting multivariate activity patterns of interest.[61] For this reason, we additionally carried out the same analysis in the fusiform gyrus of the healthy controls in order to show that, like in the patients with BPD, the category of faces was represented among the corresponding multivariate activity patterns, as indicated by a high degree of pattern similarity (Appendix 1, Figure 1). Nevertheless, we acknowledge that a more rigorous control sample with matching acquisition protocols would ultimately be favourable; as such, this control analysis represents only a first step in determining the specificity of the effects reported here.

Another limitation of our study involves the extent to which we can associate the representational geometry in the amygdala to specific pathological aspects of BPD. Although the correlation analysis revealed a possible link between the altered affective space and impulsivity scores, the association was not particularly robust, as evidenced by the failure of the correlations to survive a correction for multiple comparisons; however, this null effect could simply have been due to our small sample size.

Conclusion

As this study is, to our knowledge, the first to apply RSA to functional MRI data of patients with BPD, follow-up work incorporating similar methodology, larger samples and additional questionnaires is warranted in order to better characterize the association between neural representational spaces, emotion dysregulation and BPD. One idea would involve carrying out several neuroimaging scans throughout the course of a DBT program in conjunction with a dismantling design.[52] This approach could help to constrain our understanding of the association between specific aspects of therapy and

changes in the neural representational geometry, potentially revealing how such altered representational spaces map onto pathological behaviour in patients with BPD, thereby increasing the prognostic value of functional MRI in the clinic. Another idea would involve applying the same analyses to neuroimaging data from a different patient population — also characterized by issues with interpersonal interactions (e.g., depression) — to determine the diagnostic specificity of altered emotion spaces in the amygdala. Finally, given the complexity of mental disorders and the brain networks contributing to them, one could additionally carry out follow-up connectivity analyses using the amygdala as a seed region, in conjunction with representational connectivity analysis,[18,62,63] to determine how such multidimensional representations are altered within and between large-scale networks.

Many studies over the past decades have reported abnormal activation levels of the amygdala as a potential mechanism underlying the behaviour of patients with BPD. In this brief report, we provide a first glimpse into the combination of multivariate pattern analysis with functional MRI data acquired from patients with BPD. Before and after patients underwent a 10-week inpatient program of DBT, we used RSA to explore the informational content of activity patterns in the amygdala evoked from a task involving identification of facial expressions. Our approach showed a negative shift in the representational space before therapy, in which angry, fearful and neutral faces were represented unexpectedly similarly to one another, while surprised faces were unexpectedly dissimilar to the other expressions. This bias normalized following therapy. Such findings indicate that RSA can reveal novel insights into the neurobiological underpinnings of information processing in personality disorders, which has the potential to increase the diagnostic and prognostic value of functional neuroimaging for clinical psychology and psychiatry.

An earlier version of this manuscript was published as a preprint: Levine SM, Merz K, Keeser D, et al. Altered amygdalar emotion space in borderline personality disorder normalizes following dialectical behavioral therapy. *medRxiv* 2023 Jan. 14. doi: 10.1101/2023.01.14.23284531.

Affiliations: From the Department of Psychology, LMU Munich, Munich, Germany (Levine, Bertsch); the NeuroImaging Core Unit Munich (NICUM), University Hospital, LMU Munich, Munich, Germany (Levine, Merz, Keeser, Bertsch); the Department of Psychiatry and Psychotherapy, University Hospital, LMU Munich, Munich, Germany (Merz, Keeser, Kunz, Barton, Reinhard, Jobst, Padberg, Musil); the Department of General Psychiatry, Center for Psychosocial Medicine, Heidelberg University, Heidelberg, Germany (Neukel, Herpertz, Bertsch); and the German Center for Mental Health (DZPG), Munich, Germany (Padberg, Bertsch).

Competing interests: F. Padberg is a member of the European Scientific Advisory Board of BrainsWay Inc., Jerusalem, Israel; and the International Scientific Advisory Board of Sooma, Helsinki, Finland. He has received speaker's honoraria from Mag&More GmbH; the neuroCare Group, Munich, Germany; and Brainsway Inc. His lab has received support with equipment from neuroConn GmbH, Ilmenau, Germany; Mag&More GmbH; and BrainsWay Inc. R. Musil has received financial research support from the EU (H2020 No. 754740) and the Tourette Gesellschaft Deutschland e.V., and served as principal investigator (PI) in clinical trials from Abide Therapeutics, Böhringer Ingelheim, Emalex Biosciences, Lundbeck GmbH, Nuvelution TS Pharma Inc., Oryzon, Otsuka Pharmaceuticals and Therapix Biosciences. He is a member of the advisory board of the Tourette Gesellschaft Deutschland e.V. He has received speakers' honoraria from Otsuka Pharmaceuticals and Lundbeck. No other competing interests were declared.

Contributors: D. Keeser, B. Barton, M. Reinhard, A. Jobst, F. Padberg, S. Herpertz, K. Bertsch and R. Musil designed the study. K. Merz, J. Kunz and C. Neukel acquired the data, which S. Levine and K. Merz analyzed. S. Levine, K. Merz and K. Bertsch wrote the article. All of the authors critically revised it for important intellectual content and gave final approval of the version to be published.

Data sharing: All data produced in the present study are available upon reasonable request to the authors.

Funding: This project was supported by the FöFoLe program (PI: M. Reinhard, grant no. 996) and the FöFoLe^{PLUS} program (PI: M. Reinhard, grant no. 003, Munich Clinician Scientist Program) from the Faculty of Medicine of the LMU Munich, as well as from the Friedrich Baur Stiftung (PI: M. Reinhard, grant no. 74/18). The Heidelberg study was part of the Clinical Research Group (Klinische Forschungsgruppe [KFO] 256, spokesperson: Christian Schmahl, deputy spokesperson: Sabine C. Herpertz) on the "Mechanisms of Disturbed Emotion Processing in Borderline Personality Disorder" and was funded by grants from the German Research Foundation (Deutsche Forschungsgemeinschaft [DFG], grant nos.: HE 2660/12-2 and BE 5292/3-2). The procurement of the Prisma MRI 3T Magnetom Scanner was supported by the DFG grant for major research instrumentation (DFG, INST 86/1739-1 FuGG).

References

1. Gunderson JG, Herpertz SC, Skodol AE, et al. Borderline personality disorder. *Nat Rev Dis Primers* 2018;4:18029.
2. Levine D, Marziali E, Hood J. Emotion processing in borderline personality disorders. *J Nerv Ment Dis* 1997;185:240-6.
3. Linehan M. *Cognitive-behavioral treatment of borderline personality disorder*. New York (NY): Guilford Press; 1993.
4. Carpenter RW, Trull TJ. Components of emotion dysregulation in borderline personality disorder: a review. *Curr Psychiatry Rep* 2013;15:335.
5. Neacsiu AD, Eberle JW, Kramer R, et al. Dialectical behavior therapy skills for transdiagnostic emotion dysregulation: a pilot randomized controlled trial. *Behav Res Ther* 2014;59:40-51.
6. Lynch TR, Rosenthal MZ, Kosson DS, et al. Heightened sensitivity to facial expressions of emotion in borderline personality disorder. *Emotion* 2006;6:647-55.
7. Kleindienst N, Hauschild S, Liebke L, et al. A negative bias in decoding positive social cues characterizes emotion processing in patients with symptom-remitted borderline personality disorder. *Borderline Personal Disord Emot Dysregul* 2019;6:17.
8. Daros AR, Zakzanis KK, Ruocco AC. Facial emotion recognition in borderline personality disorder. *Psychol Med* 2013;43:1953-63.
9. Bilek E, Itz ML, Stößel G, et al. Deficient amygdala habituation to threatening stimuli in borderline personality disorder relates to adverse childhood experiences. *Biol Psychiatry* 2019;86:930-8.
10. Donegan NH, Sanislow CA, Blumberg HP, et al. Amygdala hyperreactivity in borderline personality disorder: implications for emotional dysregulation. *Biol Psychiatry* 2003;54:1284-93.
11. Herpertz SC, Dietrich TM, Wenning B, et al. Evidence of abnormal amygdala functioning in borderline personality disorder: a functional MRI study. *Biol Psychiatry* 2001;50:292-8.
12. Mitchell AE, Dickens GL, Picchioni MM. Facial emotion processing in borderline personality disorder: a systematic review and meta-analysis. *Neuropsychol Rev* 2014;24:166-84.

13. Schulze L, Schmahl C, Niedtfeld I. Neural correlates of disturbed emotion processing in borderline personality disorder: a multimodal meta-analysis. *Biol Psychiatry* 2016;79:97-106.
14. Goodman M, Carpenter D, Tang CY, et al. Dialectical behavior therapy alters emotion regulation and amygdala activity in patients with borderline personality disorder. *J Psychiatr Res* 2014;57:108-16.
15. Neukel C, Bertsch K, Wenigmann M, et al. A mechanism-based approach to anti-aggression psychotherapy in borderline personality disorder: group treatment affects amygdala activation and connectivity. *Brain Sci* 2021;11:1627.
16. LeDoux J. The amygdala. *Curr Biol* 2007;17:R868-74.
17. Woo C-W, Chang LJ, Lindquist MA, et al. Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* 2017;20:365-77.
18. Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis — connecting the branches of systems neuroscience. *Front Syst Neurosci* 2008;2:4.
19. Haxby JV, Gobbini MI, Furey ML, et al. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 2001;293:2425-30.
20. Morey RA, Haswell CC, Stjepanovic D, et al. Neural correlates of conceptual-level fear generalization in posttraumatic stress disorder. *Neuropsychopharmacology* 2020;45:1380-9.
21. O'Hearn K, Larsen B, Fedor J, et al. Representational similarity analysis reveals atypical age-related changes in brain regions supporting face and car recognition in autism. *Neuroimage* 2020;209:116322.
22. Williams AB, Liu X, Hsieh F, et al. Memory-based prediction deficits and dorsolateral prefrontal dysfunction in schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2023;8:71-8.
23. Mur M, Bandettini PA, Kriegeskorte N. Revealing representational content with pattern-information fMRI–an introductory guide. *Soc Cogn Affect Neurosci* 2009;4:101-9.
24. Levine SM, Schwarzbach JV. Individualizing representational similarity analysis. *Front Psychiatry* 2021;12:729457.
25. Popal H, Wang Y, Olson IR. A guide to representational similarity analysis for social neuroscience. *Soc Cogn Affect Neurosci* 2019;14:1243-53.
26. Dimsdale-Zucker HR, Ranganath C. Representational similarity analyses. Handbook of in vivo neural plasticity techniques. *Elsevier* 2018;509–525.
27. Levine SM, Wackerle A, Rupprecht R, et al. The neural representation of an individualized relational affective space. *Neuropsychologia* 2018;120:35-42.
28. Riberto M, Paz R, Pobric G, et al. The neural representations of emotional experiences are more similar than those of neutral experiences. *J Neurosci* 2022;42:2772-85.
29. Koide-Majima N, Nakai T, Nishimoto S. Distinct dimensions of emotion in the human brain and their representation on the cortical surface. *Neuroimage* 2020;222:117258.
30. Hariri AR, Tessitore A, Mattay VS, et al. The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage* 2002;17:317-23.
31. Konvalin F, Grosse-Wentrup F, Nenov-Matt T, et al. Borderline personality features in patients with persistent depressive disorder and their effect on CBASP outcome. *Front Psychiatry* 2021;12:608271.
32. Reinhard MA, Rek SV, Nenov-Matt T, et al. Association of loneliness and social network size in adulthood with childhood maltreatment: analyses of a population-based and a clinical sample. *Eur Psychiatry* 2022;65:e55.
33. Wolf J, Padberg F, Nenov-Matt T, et al. Suicidal behaviors are associated with loneliness and decrease during inpatient CBASP treatment for persistent depressive disorder. *J Psychiatr Res* 2022;154:139-44.
34. First MB, Williams JB, Karg RS, et al. *Structured clinical interview for DSM-5 disorders: SCID-5-CV clinician version*. American Psychiatric Association; 2015.
35. Beesdo-Baum K, Zaudig M, Wittchen H-U. *Strukturiertes Klinisches Interview für DSM-5-Störungen — Klinische Version.* Göttingen: Hogrefe; 2019.
36. Beck AT, Steer RA, Brown G. *Manual for the Beck Depression Inventory-II*. San Antonio, TX: The Psychological Corporation; 1996.
37. Hautzinger M, Keller F, Kühner C. *Beck depressions inventar revision*. Frankfurt, Germany: Harcourt Test Services; 2006.
38. Zimmerman M, Martinez JH, Young D, et al. Severity classification on the Hamilton Depression Rating Scale. *J Affect Disord* 2013;150:384-8.
39. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967;6:278-96.
40. Bohus M, Kleindienst N, Limberger MF, et al. The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. *Psychopathology* 2009;42:32-9.
41. Kleindienst N, Jungkunz M, Bohus M. A proposed severity classification of borderline symptoms using the borderline symptom list (BSL-23). *Borderline Personal Disord Emot Dysregul* 2020;7:11.
42. Giesen-Bloo JH, Wachters LM, Schouten E, et al. The Borderline Personality Disorder Severity Index-IV: psychometric evaluation and dimensional structure. *Pers Individ Dif* 2010;49:136-41.
43. Schuppert HM, Bloo J, Minderaa RB, et al. Psychometric evaluation of the Borderline Personality Disorder Severity Index-IV — adolescent version and parent version. *J Pers Disord* 2012;26:628-40.
44. Bernstein DP, Stein JA, Newcomb MD, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse Negl* 2003;27:169-90.
45. Wingenfeld K, Spitzer C, Mensebach C, et al. Die deutsche Version des Childhood Trauma Questionnaire (CTQ): Erste Befunde zu den psychometrischen Kennwerten. *Psychother Psychosom Med Psychol* 2010;60:442-50.
46. Oosterhof NN, Connolly AC, Haxby JV. CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Front Neuroinform* 2016;10:27.
47. Frazier JA, Chiu S, Breeze JL, et al. Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *Am J Psychiatry* 2005;162:1256-65.
48. Ishai A, Ungerleider LG, Martin A, et al. The representation of objects in the human occipital and temporal cortex. *J Cogn Neurosci* 2000;12:35-51.
49. Heekeren HR, Marrett S, Bandettini PA, et al. A general mechanism for perceptual decision-making in the human brain. *Nature* 2004;431:859-62.
50. Silvers JA, Hubbard AD, Biggs E, et al. Affective lability and difficulties with regulation are differentially associated with amygdala and prefrontal response in women with borderline personality disorder. *Psychiatry Res Neuroimaging* 2016;254:74-82.
51. Glenn CR, Klonsky ED. Emotion dysregulation as a core feature of borderline personality disorder. *J Pers Disord* 2009;23:20-8.
52. Iskric A, Barkley-Levenson E. Neural changes in borderline personality disorder after dialectical behavior therapy-a review. *Front Psychiatry* 2021;12:772081.
53. Bertsch K, Hillmann K, Herpertz SC. Behavioral and neurobiological correlates of disturbed emotion processing in borderline personality disorder. *Psychopathology* 2018;51:76-82.
54. Dunsmoor JE, Kragel PA, Martin A, et al. Aversive learning modulates cortical representations of object categories. *Cereb Cortex* 2014;24:2859-72.
55. Jin J, Zelano C, Gottfried JA, et al. Human amygdala represents the complete spectrum of subjective valence. *J Neurosci* 2015;35:15145-56.
56. Zhang H, Japee S, Nolan R, et al. Face-selective regions differ in their ability to classify facial expressions. *Neuroimage* 2016;130:77-90.
57. Levine SM, Pfaller M, Reichenberger J, et al. Relating experimentally-induced fear to pre-existing phobic fear in the human brain. *Soc Cogn Affect Neurosci* 2018;13:164-72.
58. Levine SM, Kumpf M, Rupprecht R, et al. Supracategorical fear information revealed by aversively conditioning multiple categories. *Cogn Neurosci* 2021;12:28-39.
59. Undeger I, Visser RM, Becker N, et al. Model-based representational similarity analysis of blood-oxygen-level-dependent fMRI captures threat learning in social interactions. *R Soc Open Sci* 2021;8:202116.
60. Puccetti NA, Schaefer SM, van Reekum CM, et al. Linking amygdala persistence to real-world emotional experience and psychological well-being. *J Neurosci* 2021;41:3721-30.
61. Tahedl M, Levine SM, Weissert R, et al. Early remission in multiple sclerosis is linked to altered coherence of the cerebellar network. *J Transl Med* 2022;20:488.
62. Coutanche MN, Akpan E, Buckser RR. Representational connectivity analysis: identifying networks of shared changes in representational strength through jackknife resampling; 2020.
63. Karimi-Rouzbahani H, Woolgar A, Henson R, et al. Caveats and nuances of model-based and model-free representational connectivity analysis. *Front Neurosci* 2022;16:755988.