Multiscale classification reveals a multivariate functional connectivity marker for borderline personality disorder

Juha M. Lahnakoski[1,2,3], Tobias Nolte[4,5], Alec Solway[6], Iris Vilares[7], Andreas Hula[8], Janet Feigenbaum[9], Terry Lohrenz[10], Peter Fonagy[5,9], P. Read Montague[4,10,11,12], Leonhard Schilbach[1,13]

1. Independent Max Planck Research Group for Social Neuroscience, Max Planck Institute of Psychiatry, Munich, Germany
2. Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Center Jülich, Wilhelm-Johnen-Straße, 52428, Jülich, Germany
3. Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Moorenstr. 5, 40225, Düsseldorf, Germany
4. Wellcome Trust Centre for Neuroimaging, University College London, London, United Kingdom
5. Anna Freud National Centre for Children and Families, London, United Kingdom
6. Department of Psychology, University of Maryland, College Park, MD, USA
7. Department of Psychology, University of Minnesota, Minneapolis, MN, USA
8. Austrian Institute of Technology, Vienna, Austria
9. Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom
10. Virginia Tech Carilion Research Institute, Virginia Tech, Roanoke, VA, USA
11. Department of Physics, Virginia Tech, Blacksburg, VA
12. Department of Psychiatry and Behavioral Medicine, Virginia Tech Carilion School of Medicine, Virginia Tech, Roanoke, VA
13. Department of Psychiatry, Ludwig-Maximilians-Universität, Munich, Germany

Corresponding author: **Juha M. Lahnakoski**

Email: j.lahnakoski@fz-juelich.de
Address Forschungszentrum Jülich,
Wilhelm-Johnen-Straße,
52428 Jülich, Germany

Keywords: Borderline personality disorder, BPD, fMRI, multivariate, functional connectivity, classification

Acknowledgements and disclosures

This study was supported by Wellcome Trust Principal Research Fellowship (PRM), The Kane Family Foundation (PRM), National Science Foundation (PRM, TN, IV, AH, AS and TL), Virginia Tech (PRM), Max Planck Society via an Independent Max Planck Research Group (LS) and Finnish Cultural Foundation (grant #150496 to JL). For the patient recruitment, and thus enabling this study, we wish to acknowledge The Personality and Mood Disorder Research Consortium, which is a recruitment network consisting of the following health services: West London Mental Health Trust, Central and North West London NHS Trust, North East London NHS Foundation Trust, Barnet, Enfield and Haringey NHS Trust, Camden & Islington NHS Trust, South West London and St Georges NHS Foundation Trust, London Probation Trust. All authors report no conflicts of interest.

Abstract

Background

Functional connectivity measures have garnered interest as possible biomarkers of psychiatric disorders including borderline personality disorder (BPD). However, small sample sizes and lack of within-study replications have led to divergent findings with no clear spatial foci. Therefore, we adopted an exploratory full-brain approach in the current study to evaluate which combinations of regions are most consistently predictive of BPD diagnosis.

Methods

We studied fMRI resting state functional connectivity in matched subsamples of 116 BPD and 72 control individuals defined by three grouping strategies: 1) referral diagnosis, 2) clinical diagnostic interview excluding patients no longer filling diagnostic criteria or controls scoring above threshold in a screening questionnaire and 3) self-reported symptom severity. We predicted BPD status using classifiers with repeated cross-validation based on multiscale functional connectivity within and between regions of interest (ROIs) covering the whole brain— global ROI-based network, seed-based ROI-connectivity, functional consistency and voxel-to-voxel connectivity within and between ROIs. Finally, we evaluated the generalizability of the classification in the left-out portion of non-matched data.

Results

Full-brain connectivity allowed successful classification (~70%) of BPD patients vs. control individuals in matched inner cross-validation. The classification remained significant when applied to unmatched out-of-sample data, but accuracies were lower (~61–70%) than in fully matched samples. The over-estimation of inner cross-validation accuracy was exacerbated by univariate regression of nuisance variables, particularly in smaller samples. Highest seed-based accuracies were in a similar range to global accuracies (~70–75%), but spatially more specific. In the seed-based classification, the regions implicated most often included midline, temporal and somatomotor regions. Highest accuracies were achieved with the clinical interview followed by referral diagnosis group definition. Self-report results remained at chance level. The accuracies were affected by an interaction of medications and global signal and univariate nuisance regression. Pairwise correlations, local consistencies and fine-scale connectivity matrices were not significantly predictive of BPD after multiple comparison corrections, but weak local effects coincided with the most discriminative ROIs in the classification.

Conclusions

Our multivariate results indicate that complex global functional connectivity differences are moderately predictive of BPD despite heterogeneity of the patient population. However, univariate nuisance regression applied to full cross-validation dataset can cause inflation of accuracies compared with left-out test data.

Introduction

Impaired social interactions are a central characteristic of personality disorders (1). Borderline personality disorder (BPD), in particular, has been empirically characterized by profound social deficits, with patients notably suffering from dysfunctional relationships (2). Core clinical features of BPD include affective dysregulation (3), impulsivity (4), heightened risk for self-harm and suicidality (5), relational instability and hypervigilance to motives of others in close relationships (6) as well as a bias in attributing hostility to others (7). It has been suggested that at the core of patients' difficulties lies a problem with mentalizing, i.e. an inability to imagine and thus perceive and interpret human behavior in terms of underlying mental states (8,9). These social cognitive impairments may reflect changes in the capacity to understand the internal states of self and others and respond to implicit trust gestures (10). This may be particularly relevant for BPD as a developmental psychopathology, for which early adversity and disorganized attachment relationships have been established as critical contributing factors. A cascade of deteriorating disease states in patients is often provoked by heightened interpersonal stress/threat (e.g. rejection, abandonment, or isolation) accompanied by the experience of a reduced sense of agency and a resulting propensity to react (2,11). While there is accumulating evidence that symptoms such as emotional dysregulation are associated with functional and structural differences in frontal and limbic regions (e.g. 12, 13), relatively few studies have examined resting state functional connectivity differences in BPD. Some studies have reported connectivity increases while others have reported seed region-specific decreases, with the overall effect sizes being small and spatially scattered (e.g. 14–19). Other approaches have been proposed to evaluate brain connectivity differences, for example, estimates of anatomical connectivity based on diffusion tensor imaging. These have suggested mainly reduced fractional anisotropy (FA) values in BPD in varying white matter tracts, although the locations appear to differ from study to study. One study found reduced FA values in the inferior longitudinal, uncinate and occipitofrontal fasciculi (20), but these changes were only observed in adolescent and not adult participants. Another study tested the fractional anisotropy (FA) only in the uncinate and cingulum, finding decreased FA values in BPD sample compared to controls in the uncinate and not in the cingulum. By contrast, two further studies did find effects in the cingulum (21,22) and, in the case of one of the studies, not in the uncinate fasciculus (21). A recent meta-analysis of data from four studies also showed support for decreased FA values in BPD localized in the corpus callosum and the fornix (23). Some further studies have produced high accuracies and effect sizes based on other measures, such as BOLD power at specific frequency bands or network analysis techniques (24), but like most other studies to date, they have relied on small samples and the findings have not yet been replicated. Therefore, there is a concerning lack of consensus on whether, or in which ways, BPD may be characterized by aberrations in brain connectivity.

Generally, an important limitation of prior functional connectivity studies of BPD has been the small sample size, usually approximately 20 patients in the studies cited above, although some studies with double the number of subjects exist (e.g. 24–26). Moreover, the lack of within-study replication through independent and repeated cross-validation has limited the insight into the reliability of the findings. To address these issues, we recruited a larger group of individuals with BPD from a number of referral services and a group of healthy controls (HC) in order to assess the reliability of the findings. Due to the inconclusive nature of previous results, we adopted an explorative approach by calculating the functional connectivity of fMRI data over the whole brain at different spatial scales and used these connectivity values as features in a machine learning classification approach. The different scales of analysis are schematically visualized in **Figure 1** using a small set of hypothetical regions. At the global scale, we calculated a global connectivity matrix between activity time courses of 273 regions of interest (ROIs) covering the entire brain using linear correlations. At the seed scale, we extracted the rows of the global connectivity matrix to evaluate how important each region is for classification performance. At

the voxel-scale, we first calculated the accuracies based on voxel–voxel connectivity matrices within ROIs and between all pairs of ROIs. Finally, at the ROI scale, the functional consistency of the ROIs and ROI pairs (mean of the correlation matrix entries within/between ROIs reflecting the (inter)regional homogeneity of the voxels) was compared between groups and univariate contrasts and classification were performed to compare the local results directly to multivariate accuracies at other scales. We also evaluated the effects of spatial smoothing and global signal regression on prediction accuracies. In all classification analyses, we used 5-fold cross-validation of a balanced and matched (according to age, sex, scanning site, education level, and scores on the Raven's standard progressive matrices (RSPM) intelligence test) subsets the participants (inner cross-validation). To gain insight on the stability of the results, we repeated the inner cross-validation multiple times with random splits into folds of training and test data. Finally, we evaluated the generalizability of the findings by applying the classifiers to the participants that were excluded from the matched inner cross-validation.

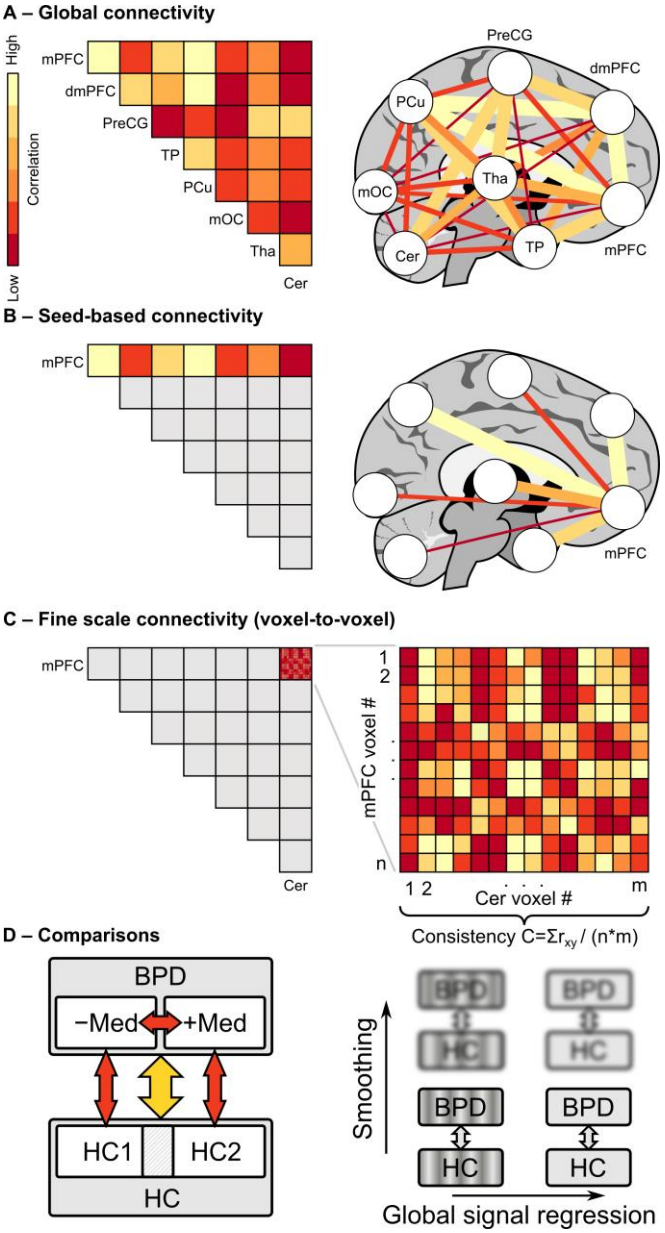


Figure 1: Schematic representation of the scales of analysis for a selection of regions. A *Global connectivity – The full global scale connectivity matrix between the mean time courses of each ROI was used as features for classification.* ***B*** *Seed-based connectivity – The connectivity from one ROI to all other ROIs (rows of the global connectivity matrix) are used for classification to improve spatial specificity of results.* ***C*** *Fine scale connectivity – Voxel-to-voxel connectivity between voxels in all unique pairs of ROIs are used for classification. Additionally, the consistency of ROIs and ROI pairs was calculated as the mean of all unique fine scale connectivity values.* ***D*** *Comparisons – Main comparison was between matched BPD and HC groups (left panel, yellow arrow). Additionally, to evaluate the effect of medications taken by a subset of the patient group, we repeated the global and seed-based analyses between matched groups of HC participants and sub-groups of medicated (BPD–Med) and unmedicated (BPD+Med) BPD patients as well as between the two BPD groups (red arrows). Finally, we evaluated the effects of spatial smoothing and global signal regression to the accuracies in the different analyses (right panel).*

Methods and materials

Participants

One hundred and eighty-seven adult participants were selected from a larger study investigating social exchanges in BPD and antisocial personality disorder. Here, we included only the control participants and the patients with BPD. From those, we excluded participants that had more than 10% of their data affected by excessive motion (defined as >.5 mm framewise displacement; N=17) or whose data was otherwise noisy (extensive signal distortion in the EPI images; N=1), and participants with incomplete data on sex/gender (N=2), leaving 167 participants (63 HC, 104 BPD as reported by referring clinician) for subsequent matching and analysis.

Patients were referred by psychiatrists, care coordinators, and (trainee) clinical psychologists within specialist personality disorder services of seven London NHS Mental Health Trusts across five London boroughs. In the following, the diagnosis reported during this referral process is called the "referral diagnosis". A further clinical interview was performed to evaluate whether the patients currently filled the diagnostic criteria for BPD and a diagnostic screening questionnaire was filled by the control participants to evaluate whether they showed symptoms consistent with BPD (see Participant matching section below). Individuals with recent psychotic episodes, severe learning disabilities, or current or past neurological disorders or traumas were excluded. Additionally, participants who were currently taking substances (other than their prescribed medication) were excluded in prior screenings (self-reported). Patients were interviewed by psychologists with the Structured Clinical Interview for DSM-IV Axis II Diagnoses (28). Since each rater interviewed different patients, rating consistency was assessed with a mixed-effects model. BPD symptom scores were placed at the within-level and clustered by rater at the between-level. The intra-class coefficient (ICC) was then estimated for BPD symptom scores. While higher ICCs typically reflect better reliability when reliability of diagnostic classification is evaluated, low scores are preferred using the current model as they reflect greater within-level vs. between-level variation (i.e., differences between patients within raters vs. between raters). The ICC for BPD symptom scores was 0.05 (95% CI [.01, .2]). Thus, raters were consistent in how they scored patients; the main source of variation occurred between patients within a given rater. Variability around the mean symptom rating was also low (M = 6.8, SD = 0.9), which further supports the consistency in ratings.

The entire dataset reported here consisted of 71 controls and 116 patients (63 and 104 after exclusion, respectively), who varied widely in age from 17 to 62. Of all participants, 138 were females and 47 males (1 transgender, 1 no response) yielding a female to male ratio of 2.9/1. For the patients, female to male ratio was 4.7/1 (94 female, 20 male) and for the control group 1.6/1 (44 female, 27 male). Due to this discrepancy and the size difference between the groups, we excluded participants to make group sizes and sex distributions equal in the main analyses (see Participant matching-section) and applied the classifiers on this matched data to the rest of the (unmatched) sample. Of the patients, 42 were not medicated, 5 patients were on four medications, 7 on three medications, 29 on two medications and 33 on one medication. None of the control participants reported to be currently taking prescription medications. Medication included most often antidepressants (various types, 52% of patients), antipsychotics (various types, 19% of patients), benzodiazepines (9% of patients), anticonvulsants/epilepsy medication (8%), beta blockers (9%), sleeping pills (6%) and antihistamines (3%).

Procedure

Over several visits, all participants underwent a comprehensive study protocol comprising a number of behavioral paradigms, questionnaires, diagnostic and developmental interviews, as well as the RSPM test to estimate their non-verbal fluid intelligence, and structural and functional (neuroeconomics-based

two party social exchange neuroimaging tasks) magnetic resonance imaging. Resting state functional connectivity data were acquired between the last task-based fMRI paradigm and multi-parameter mapping MRI sequences. The order of the task paradigms was counter-balanced to avoid potential mean differences introduced by the preceding task, although we did not expect lasting, consistent connectivity differences to remain after the task had ended. During resting state functional imaging, participants were instructed to lie still inside the scanner with their eyes open while looking at the MS Windows 2000 start screen, with the windows logo at the center. Participants were instructed to "think of whatever comes to mind and to let their mind wander". Eye tracking was used to control whether they stayed awake during the 4:30 minutes of data acquisition and those data sets excluded where alertness was questionable or where participants fell asleep.

Functional magnetic resonance imaging

Functional imaging was performed at three sites in London, United Kingdom with similar Siemens MAGNETOM Trio 3-tesla MRI scannners. The voxel size of data was 3.4375 X 3.4375 X 4 mm^3, FOV 220 mm, 37 slices, TR 2 s, TE 25 ms, flip angle 90°. Total duration of the functional scan was 5 minutes (150 volumes). Additionally, T1 weighted MPRAGE images (matrix size 512x448, in-plane resolution 0.4785x0.4785 mm, 192 slices, slice thickness 1 mm, TR 1200 ms, TE 2.66, inversion time 600 ms, flip angle 12°) were acquired for anatomical registration. Because of the three different locations and scanners, the scanning site was used as an additional hard criterion during pairwise matching of subjects and it was added as an effect of no interest in the analyses.

Preprocessing

The fMRI data was preprocessed using FSL (fMRIB Software Library; 21, 22). The first four volumes were removed from the data to allow the tissue magnetization to stabilize. Motion correction was applied using MCFLIRT (31) and brain extraction was performed using BET (32). Anatomical images were corrected for intensity inhomogeneities during the brain extraction. Motion parameters were regressed out of the data using the general linear model (GLM) in FSL (*fsl_glm*) in subjects' native space. Additionally, volumes containing excessive motion were blacklisted using the *fsl_motion_outliers* tool, and the blacklist matrix was later used for excluding the effects of those volumes when calculating the connectivity matrices. Functional data was registered to the anatomical data of each subject and then to 2-mm MNI standard space (Montreal Neurological Institute) using FLIRT (31) using 9 degrees of freedom (translation, rotation and scaling). Spatial smoothing of the data was omitted during preprocessing, but the effect of spatial smoothing was explored during analysis (preprocessed data was smoothed with Gaussian kernel with 6-mm full width at half maximum). High-pass temporal filter with Gaussian-weighted least squares straight line fitting (σ=200s) was applied to remove low-frequency drifts from the data. When loading the data for analysis, an additional Parks-McClellan finite impulse response, linear-phase band-pass filter was applied in Matlab (0.02–0.08 Hz, minimum attenuation in the stop-band 40 dB, maximum ripple in the pass-band 3 dB).

Participant matching

The subject groups were formed using three strategies: 1 – prior referral diagnosis of BPD 2 – diagnosis based on structured clinical interview according to DSM-IV (SCID II; subsequently referred to as "SCID grouping") excluding patients who no longer filled the diagnostic criteria for BPD and control participants scoring above the cutoff in the Standardised Assessment of Personality: Abbreviated Scale (SAPAS; 25), and 3 – cutoff on the borderline subscale of Personality Assessment Inventory (PAI-BOR; 26) measuring the self-reported symptom severity (here referred to as "PAI-BOR grouping"). In the main analysis, to be conservative and to remove variance of no interest and generate groups of comparable size, each HC participant was individually matched with a BPD participant. The same script was used for participant matching for all grouping strategies according to the following pre-defined algorithm, which

was chosen prior to analysis. First, for each HC participant, BPD participants that were either of different sex or were imaged at a different site were removed. Then the 10 BPD participants that were closest in age to the current HC participant were retained and the person with the most similar RSPM score and education level was selected as the matching BPD participant for the HC participant. The resulting age, education level and RSPM (35) score distributions are depicted in **Figure 2** for the SCID II grouping. The distributions were largely similar for all groupings. Patient (and control) group sizes for different grouping strategies were N=39 for SCID II, N=52 for Referral and PAI-BOR groupings, and N=26 for unmedicated SCID II grouping. Because a subset of the patients showed higher levels of motion based on the total root-mean-square (RMS) motion, but did not reach the exclusion criteria, we additionally performed motion matching to evaluate the effects of motion on classifier performance. To achieve this, we repeated the generalization analyses at 5 different motion matching levels by splitting the participants into 1–5 equally sized groups based on the RMS motion percentiles. This was used as an additional hard matching criterion in the inner cross-validation as described above for sex and imaging site. We additionally calculated correlations between classifier scores and levels of RMS motion across participants to evaluate whether the classifier performance was driven by overall motion.

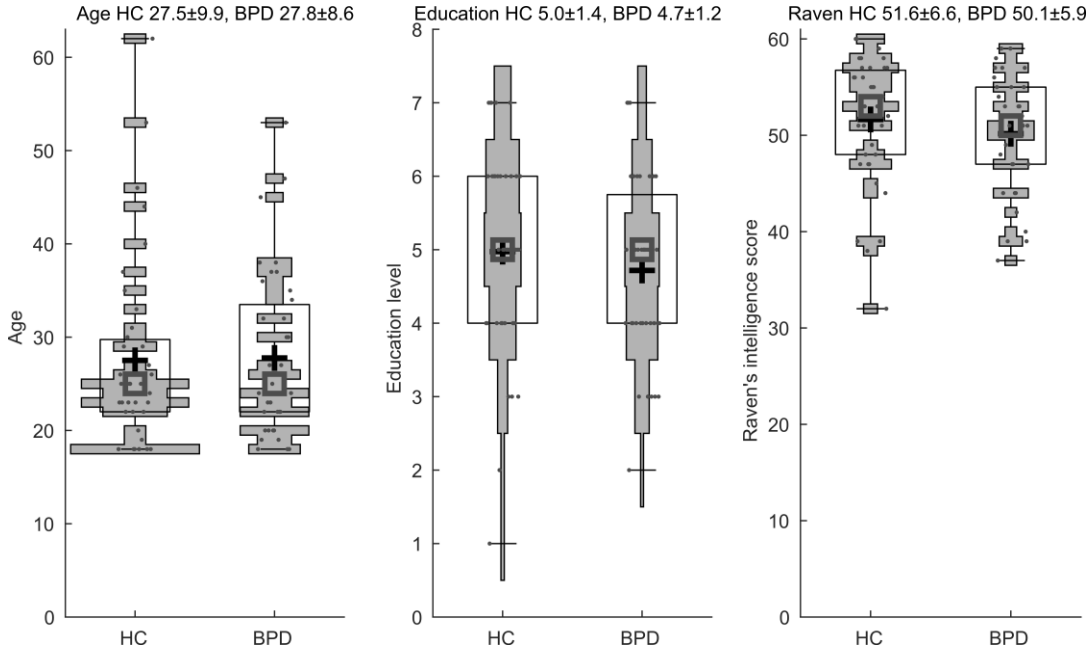


Figure 2: *Violin plot histograms of the distributions of age, education level and RSPM scores after subject matching. These data show distributions for grouping based on the SCID II diagnostic interview including the medicated patients. Distributions were largely similar with the other grouping approaches.*

To further reduce effects of age, sex and scanning site, we regressed these variables out of the raw connectivity values of the whole population at each voxel/ROI pair in the main analysis. To evaluate the generalizability of the classification, we additionally split the sample into a balanced training set where we performed a similar nuisance regression as above, but additionally applied the fitted regression weights to the unbalanced left-out data to perform the out-of-sample analysis described below. Because more than half of the participants in the entire BPD sample were on at least one medication (See Table 1 for summary of medication status and sample size with different strategies in full sample and after exclusion and group matching, respectively), we visualized the confusion matrix of three-group classification (HC, BPD+Med, BPD–Med). Finally, we repeated the two-group (HC vs. BPD) analyses with matched groups of only unmedicated patients and healthy controls using the SCID II diagnostic criteria

to explore the effect of medication on the classification performance. This left 26 people in each group after matching for sex, scanning site, age, education level and RSPM score.

Table 1: *Medication status of patients and sample sizes with different grouping strategies.*

		Number of medications							
	Grouping	0	1	2	3	4	>0	N	N_{BPD}
Full	Referral	36.2%	28.4%	25.0%	6.0%	4.3%	63.8%	187	116
Included	SCID	40.5%	23.0%	21.6%	8.1%	6.8%	59.5%	130	74
	PAIBOR	45.4%	23.1%	21.3%	6.5%	3.7%	54.6%	161	108
	Referral	33.7%	29.8%	25.0%	6.7%	4.8%	66.3%	167	104
Matched	SCID	46.2%	28.2%	17.9%	5.1%	2.6%	53.8%	78	39
	PAIBOR	50.0%	23.1%	21.2%	5.8%	0.0%	50.0%	104	52
	referral	36.5%	32.7%	21.2%	5.8%	3.8%	63.5%	104	52

fMRI data analysis

Data were analyzed in Matlab (R2017a; MathWorks, Inc., Natick, MA, USA). Data were loaded using NIfTI tools for Matlab (Jimmy Shen; http://de.mathworks.com/matlabcentral/fileexchange/8797-tools-for-nifti-and-analyze-image). We used "Brainnetome" (36) and cerebellar connectivity (37) atlases masked based on the EPI data intensity over participants to define 273 ROIs covering the entire gray matter. ROI time courses were calculated as the mean of the voxel time courses belonging to that ROI[1]. Connectivity matrices were then calculated as linear correlations between ROI time courses while controlling, through linear regression, for volumes that were affected by excessive motion (FSL motion outliers output file) estimated during preprocessing (additionally, motion parameters were regressed out during preprocessing as described in the Preprocessing section). Additionally, the analyses were repeated after also regressing out the mean signal over all voxels within the brain to elucidate the effects of global signal regression (GSR) on connectivity structure and classification accuracy.

We used linear support vector machine (SVM) classifiers with 5-fold cross-validation implemented in Statistics and Machine Learning Toolbox in Matlab. Classifiers were trained on data at multiple scales (see **Figure 1**): 1 – the full correlation matrices (global network classification), 2 – the rows of the correlation matrices (seed based classification), 3 – mean correlation between voxels in a ROI (ROI consistency) and between ROI-pairs, 4 – full correlation matrices between voxels within ROIs (within ROI connectivity) and 5– correlation matrices between voxels of pairs of different ROIs (between-ROI connectivity). For the global network classification, we additionally performed a three-group classification using a combination of three one-class SVM classifiers (radial basis function kernel) to visualize the confusion between healthy controls and patients with and without medication.

[1] Additionally, we repeated the analyses using the first principal component time courses of the ROIs. This resulted in a reduction of correlation between ROIs and slight reduction in prediction performance of the classifiers. Seed based accuracies were positively correlated over ROIs, but variable (r~0.38) and the significantly predictive regions overlapped, but the overlap was not perfect (DICE index ~0.61)). Employing the mean time courses rather than principal components also yielded connectivity matrices with a clearer modular structure. Therefore, we focus here only on the results based on the mean time courses of the ROIs.

All the main classification analyses were repeated multiple times to estimate the stability of the results (global classification 1000 times, seed-based classification 50 times/ROI, ROI consistency and finally within/between-ROI classification 10 times due to computational costs). Significance of all classification analyses were estimated by repeating identical analysis pipelines with randomly permuted labels for the training data and comparing the true accuracies to the obtained empirical null distribution. The number of permutations was equal to the number of iterations of the classifier with real labels as we aimed to estimate the distribution of both the real accuracies in the population as well as the associated null distributions. Null distributions for ROIs and ROI-pairs were assumed to be identically distributed and thus the entries of the connectivity matrices were aggregated into a single null distribution for each analysis type to reduce computation time, but different distributions were estimated for different analysis scales (global, seed-based and voxel-scale analyses). To improve the estimates of probabilities at the extreme values of the null distributions, we fit semi-parametric distributions to the empirical null values where the bulk of the distribution was represented directly by the empirical probabilities and the tails ($p<0.1$ and $p>0.9$) were approximated by pareto distributions as implemented by paretotails-function in Matlab.

In addition to the main analysis, an out-of-sample analysis was performed to assess the generalizability of the classifiers to completely unseen data. Here, a classifier was trained on random balanced subsamples, the inner 5-fold cross-validation accuracy was estimated similarly to the main analysis and, additionally, the classifiers trained on inner sample were tested on the left-out data, which was inherently unbalanced both regarding sex and diagnosis (more female than male and more patients than controls). This analysis was performed with and without nuisance regression (weights of which were based only on the inner sample, see above). Additionally, for the global connectivity matrices, the analysis was performed with different training sample sizes (1/3, 1/2 and 2/3 of the full matched sample size of the main analysis) and the analysis was repeated 500 times for each combination. For the seed-based and within/between-ROI analyses, we only estimated the out-of-sample accuracy with ½ of the total sample and reduced number of iterations (50 for seed-based, 10 for within/between ROI matrices) due to computational complexity. Null distribution was estimated similarly but with permuted labels, as in the main analyses.

In addition to the classification analyses, we contrasted the univariate connectivity values (correlations between pairs of ROIs and local consistencies) with two-sample t-tests with unequal variances. To evaluate the replicability of these findings, we additionally repeated the analyses 100 times by splitting the data into a “discovery” and “replication” samples (40% of the data, each, to reduce the overlap between subsequent iterations) and evaluated the t-value distributions and replication rates in only those links that showed a significant effect in the discovery data. This evaluation was performed at different thresholds ($p<.05$, $p<.005$ and $p<.001$, uncorrected). Finally, we performed a simple univariate threshold-based out-of-sample analysis to assess the discriminability of the connectivity values. Here, we selected a threshold that best discriminated a training sample (defined as the midpoint between two participants’ values that maximized the training accuracy[2]) and applied this to the left-out data. This analysis was repeated 50 times and again compared to a null distribution estimated using randomized group labels.

[2] Note that this procedure may not have a unique solution as more than one threshold can lead to the same accuracy. In cases when this happened, we selected the lowest such threshold. Here, we performed the univariate classification as a supplementary analysis to the univariate t-test to allow a simpler comparison with the classification analyses and selected this approach due to its computational efficiency. More robust methods may be preferable when high generalizability is of primary concern. However, the results presented here were largely consistent with a more standard linear discriminant analysis in a smaller validation analysis.

Results

We explored the effects of several factors in the data: 1 – scale of analysis (global, seed-based, ROI/ROI-pair consistency, fine-scale voxel-level connectivity), 2 – diagnostic approach (referral, SCID II or PAI-BOR), 3 – medication status (control vs. BPD in medicated, unmedicated and mixed groups; BPD unmedicated vs. medicated), 4 – global signal regression, 5 – spatial smoothing.

At the level of single links between ROIs (correlation between ROI time courses), there were no significant differences (t-test) with any of the group definition strategies or preprocessing approaches after correcting for multiple comparisons. Applying the global signal regression, the connectivity values were reduced leading to more negative correlations, which has been mathematically shown to occur (38), but the group differences remained non-significant. Because the group differences did not survive corrections for multiple comparisons, we evaluated the reproducibility of local link-level differences at a more liberal threshold by repeatedly (100 iterations) splitting the data to non-overlapping discovery and replication subsamples and evaluating the replication rates (Figure 3). No effects replicated at even moderately conservative uncorrected thresholds ($p<.001$; Figure 3A, replication percentages shown black outline next to the vertical p-value thresholds). Even at liberal thresholds ($p<.05$, uncorrected) replication rates remained low (6.1% for positive and 8.6% for negative effects). However, there was a tendency for the effects to be in the same direction in both discovery and replications subsamples as evidenced by the shift in the distributions and the lower reversal (percentages without outline) than replication rates. While replication rate across all links was relatively low, a subset of links replicated in up to 38% of all iterations (Figure 3B), suggesting that weak, but reproducible effects may exist. However, the pattern of differences is relatively sparsely spread around the brain. To summarize the regions that most often showed group differences, Figure 3C shows the number of significantly replicable links originating at each ROI. These results highlight a mix temporal, temporo-parietal, midline and subcortical regions as the most discriminative "hubs", showing the highest number of discriminative links.

Figure 4 A shows the accuracy distributions of cross-validated support-vector machine classifiers (1000 iterations; 5-fold cross-validation within iterations) at the level of global connectivity structure compared with the null distributions of the classifiers trained with randomly permuted labels. Highest classification accuracies were achieved with grouping based on the SCID II interviews (mean accuracy 70%, $p<0.00001$) followed by referral diagnosis (mean accuracy ~58%, $p<0.05$). This drop in accuracy was driven by chance level prediction of individuals who either no longer fulfilled the diagnostic criteria for BPD in the SCID II interview (patients) or scored above threshold in the SAPAS screening questionnaire for personality disorders (controls). Accuracies using only the scores of the self-reported symptom severity (PAI-BOR) remained at chance level (~56%, $p\sim0.07–0.11$, n.s.). Overall, these results suggest that a thorough diagnosis at the time of intake is critical to uncover brain fingerprints of current psychopathology. Thus, in the following figures, we will focus mainly on the SCID grouping strategy.

Global signal regression reduced accuracies with all grouping strategies (SCID II 65%, $p<0.001$; referral diagnosis 55%; n.s., PAI-BOR ~53%, n.s). Results with spatial smoothing have been omitted in the figure because, at the scale of ROI–ROI connectivity, spatial smoothing had no discernible effect on the accuracies. On average, the predictive performance was largely balanced, i.e. the mean sensitivity and specificity were approximately equal to each other and to the mean accuracy (see also **Supplementary Figure 1**).

The seed-based ROI–ROI connectivity (**Figure 4 B**) revealed widely distributed areas whose connectivity profiles were predictive of BPD diagnosis with the most discriminative SCID II grouping strategy, that are similar to those showing weak but reproducible local effects in Figure 3. The most discriminative regions

largely overlap with social brain networks that are activated during mentalizing (39). However, these regions extend beyond the midline and temporoparietal regions of the "mentalizing network", to the left hemisphere dominant temporo-frontal regions that are important for language as well as parietal regions along the intraparietal sulcus (IPS), supplementary motor area and premotor regions that are important for action understanding as well as pain perception during naturalistic social observation (see e.g. Lahnakoski et al., 2012). To compare the most discriminative regions across all three grouping strategies, **Figure 3 C** shows the overlap of the significantly predictive seed regions. The regions that most consistently discriminate between patients and controls with all grouping strategies include the dorsomedial prefrontal cortex, precuneus and posterior cingulate cortex, anterior temporal lobe, and left-lateralized posterior superior temporal sulcus, inferior frontal gyrus, anterior insula and supplementary motor area/paracentral lobule. The seed ROIs that reached a significant classification accuracy with at least two of the three group definition strategies are listed in **Table 2** including the top rated "Behavioral domains" listed in the Brainnetome atlas. Most common high-level domains are *cognition* (13 regions; most common sub domains were *social cognition*, *memory* and *language*, each mentioned 3 times), *perception* (8 regions, 6 of which were *visual*), *action* (4 regions including *execution*, *inhibition* and *imagination*), followed by *emotion* and *interoception* (2 regions each). The full list of significant classification accuracies with each group definition strategy and the MNI coordinates of ROI centroids and atlas labels are listed in **Supplementary Table 1**).

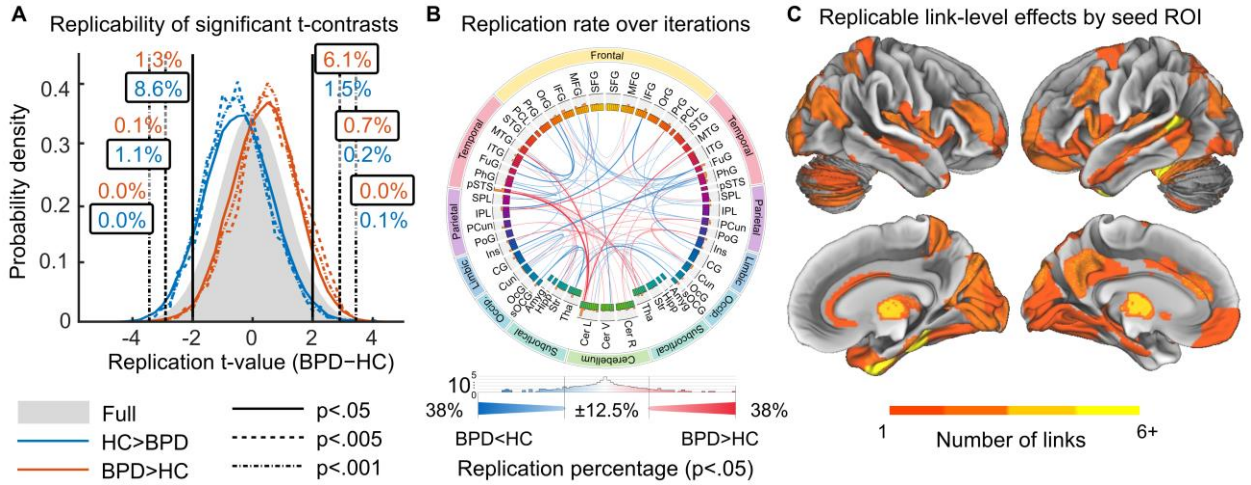


Figure 3: Reproducibility of local differences. A *The distribution of all local difference t-values across connections and iterations is shown in gray. The distributions of t-values in the outer cross-validation of those links that showed significant positive (BPD>HC) or negative (HC>BPD) effects in the inner cross-validation sample are shown in red and blue, respectively. The distributions and the associated (one-tailed) t-value thresholds for three p-levels are plotted in solid (p<.05), dashed (p<.005) and dot-dashed (p<.001) lines. Replication rates are shown surrounded by a black outline next to the lines denoting each p-value threshold (red numbers indicate BPD>HC and blue numbers HC>BPD effects in the discovery sample). Reversal rates of the effects are shown similarly, except without a black outline.* ***B*** *Replication rates of individual links at initial uncorrected threshold of p<.05 between discovery and replication analysis. Thickness and color intensity indicate the percentage of replications over iterations and threshold for visualization is based on the maximum value observed in identical analysis with randomly permuted labels. Colors are similar to panel A. The distribution of replicabilities over all links is shown on a logarithmic scale below the connectivity plot.* ***C*** *Number of links showing replicable effects from each ROI (node degree in panel B). Colormap is capped at 6 to visualize differences between ROIs. However, highest degree is 11 at the left pSTS ROI.*

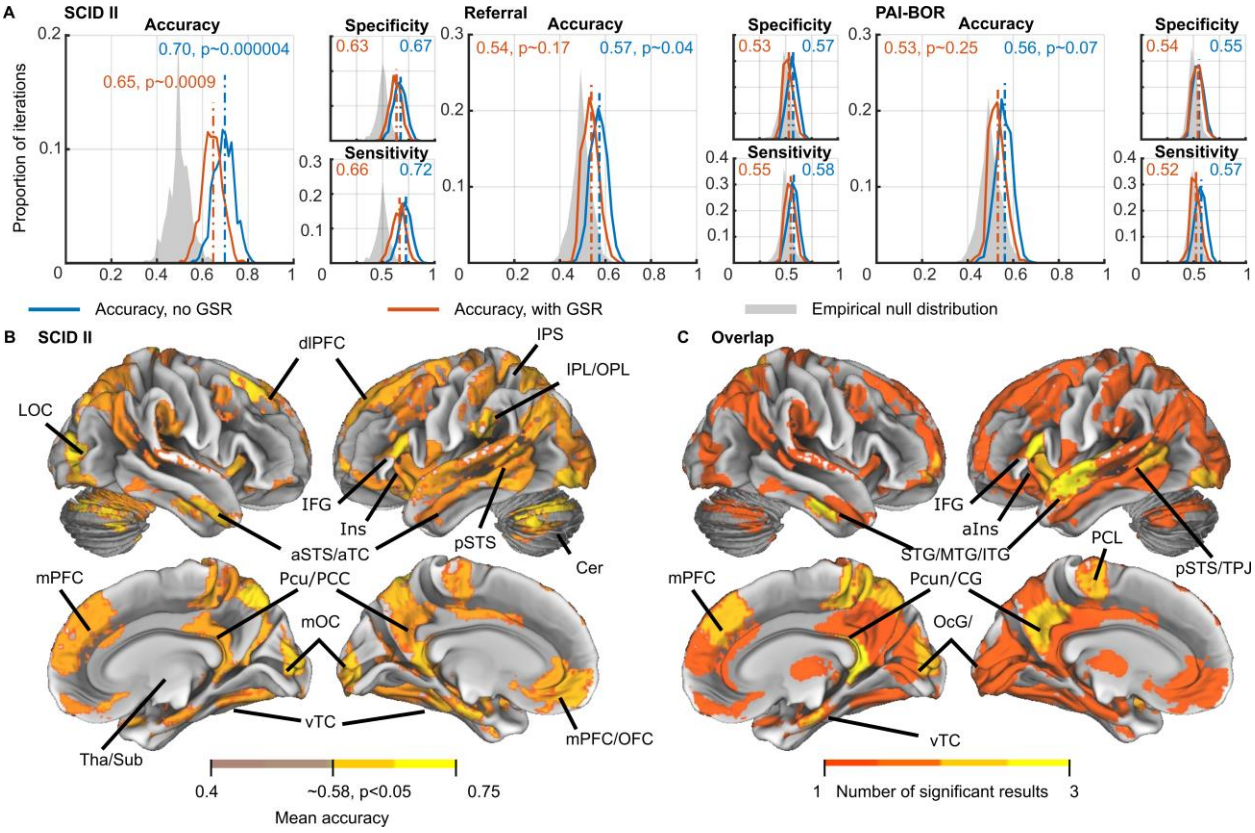


Figure 4: Classification accuracies based on global network and seed-based ROI connectivity. A *Distributions of accuracies based on full global network connectivity between ROIs over 1000 iterations with different group definition strategies (left – SCID II diagnosis, middle – referral diagnosis, right – PAIBOR self-report cutoff) showing highest accuracies for the SCID II diagnosis, reduced accuracies for the referral diagnosis grouping and null accuracies for grouping based on self-reports. Null distribution using randomly permuted labels is shown in gray; accuracies without global signal regression are shown in blue and accuracies with global signal regression in red. Global signal regression reduces accuracies in all cases, but results are unaffected by spatial smoothing at the ROI scale.* ***B*** *Mean seed based accuracies visualized on the brain surface at each seed ROI location for SCID II diagnosis. Threshold is set at p<0.05 in the full distribution without averaging.* ***C*** *Overlap of significantly predicting regions with the three grouping strategies. Color indicates in how many (out of three) analyses the prediction accuracy is significant. Abbreviations: aSTS/ATC – anterior superior temporal sulcus/anterior temporal cortex, Cer – Cerebellum, IPL/OPL – inferior/opercular parietal lobule, IPS – intraparietal sulcus, LOC – lateral occipital cortex/complex, mOC – medial occipital cortex, mPFC – medial prefrontal cortex, OFC – orbitofrontal cortex, PreCG – precentral gyrus, pSTS – posterior superior temporal sulcus, TPJ – temporoparietal junction, vTC – ventral temporal cortex.*

To evaluate the extent to which the findings in the matched subsample generalize to new data, we applied the classifiers trained on balanced subsamples of the data to the unbalanced left out participants. **Figure 5A** shows the inner and outer cross-validation accuracies with the different grouping strategies before (left) and after (right) removing nuisance covariates based on the training data. The colors of the dots indicate the size of the groups in the inner cross-validation sample. Although the classification accuracy remains above chance level in the outer generalization sample, the inner cross-validation tends to over-estimate the accuracy, particularly when the nuisance covariates are regressed out of the connectivity matrices, suggesting either non-generalizable nuisance effects or possible data leakage from the nuisance regression in the inner cross-validation sample.

Although the accuracies were significant, generalizable and stable with multiple training sample sizes, the applicability of these accuracies (0.6-0.7, depending on the specific approach), the real-world applicability depends on the base rate of the condition in the population. To illustrate this, **Figure 5B** shows the accuracy and the positive and negative predictive values sampled at different prevalence levels in the outer cross-validation sample. While the accuracy remains stable at all prevalence levels, the positive predictive value drops (and negative predictive increases) quickly as the prevalence drops to more realistic levels.

Figure 5C and D show the ROIs whose connectivity profile (rows of the connectivity matrix) was significantly predictive of BPD diagnosis using the SCID grouping. The data are visualized as percentage of iterations where both inner and outer cross-validation results were significant **(Figure 5C)** and percentage of iterations where either inner (red: **Figure 5D**) or outer (blue) cross-validation accuracies were significant (p<0.05, FWER controlled). The areas are largely consistent with both approaches.

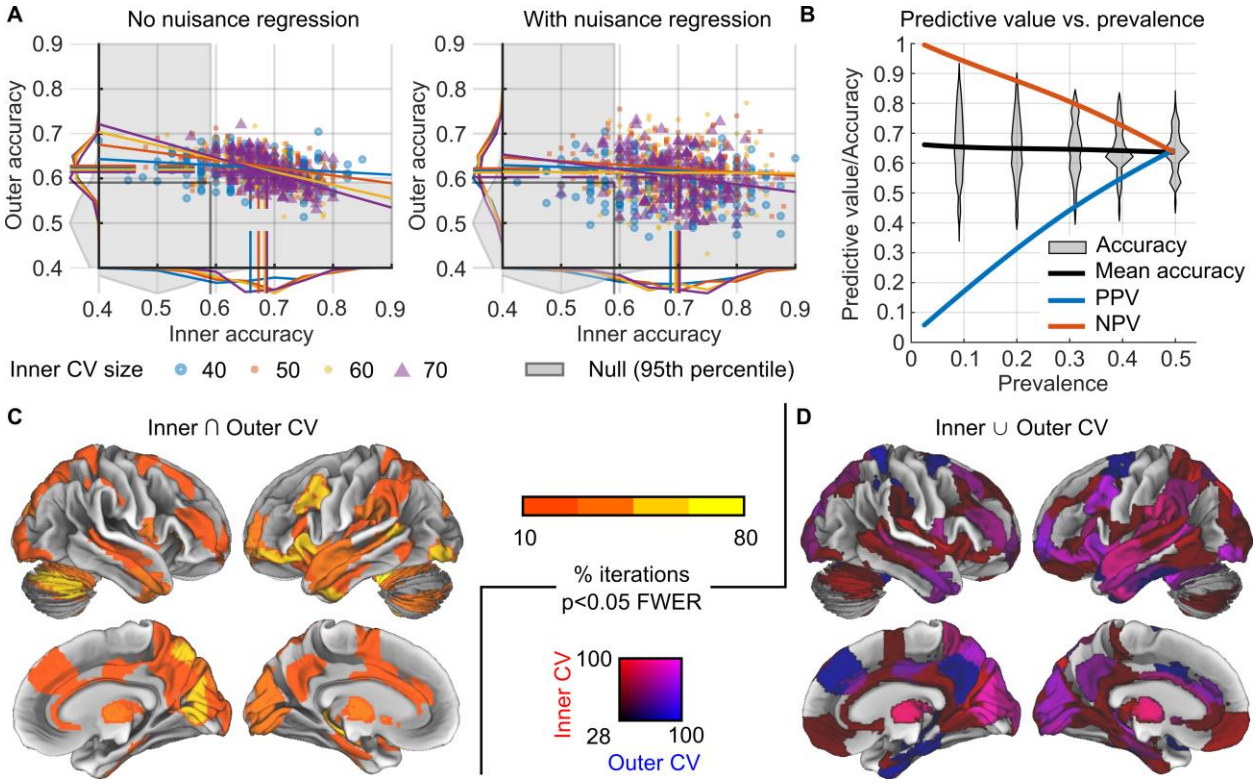


Figure 5: Generalization of classification to unseen data. A *Distributions of classifier performance in inner vs. outer cross-validation samples in the same iterations. Left panel shows results without nuisance regressors and right panel shows results when nuisance effects were estimated in the whole inner cross-validations sample.* ***B*** *Effects of prevalence of BPD diagnosis in the outer cross-validation sample. The mean accuracies, positive predictive values and negative predictive values are plotted against the proportion of patients in randomly selected subsets of the left-out data.* ***C*** *ROIs whose connectivity profiles significantly (p<0.05, FWER controlled in empirical null analyses) predicted BPD diagnosis in both inner and outer cross-validation in the same iteration.* ***D*** *ROIs whose connectivity profiles significantly (p<0.05, FWER) predicted BPD diagnosis either in the inner (red) or outer (blue) cross-validation.*

Table 2: Consistent accuracies with different group selection strategies. *Region labels, accuracies and top Behavioral Domains listed in the Brainnetome atlas (http://atlas.brainnetome.org/bnatlas.html; retrieved Feb. 15th 2019)*

		Classification accuracy			
Region label		**SCID II**	**Referral**	**PAI-BOR**	**Behavioral Domain**
SFG_R_7_3	A9l	0.633	0.588		Cognition.Social_Cognition
SFG_R_7_6	A9m	0.619	0.603		Cognition
IFG_L_6_3	A45c	0.690	0.625	0.590	Cognition.Memory.Explicit
PrG_R_6_2	A6cdl	0.694	0.607		Perception.Vision.Shape
PCL_R_2_1	A1/2/3ll	0.629		0.638	
PCL_L_2_2	A4ll	0.624	0.593		Action.Imagination
STG_L_6_3	TE1.0&1.2	0.627	0.612		Cognition.Music
STG_L_6_5	A38l	0.638	0.598		
STG_L_6_6	A22r	0.624	0.625	0.591	Emotion.Sadness
MTG_L_4_3	A37dl	0.605	0.602		Cognition.Language.Syntax
ITG_L_7_2	A37elv	0.595	0.601		Cognition.Memory.Working
PhG_L_6_3	TL	0.659	0.592		Cognition.Social_Cognition
pSTS_L_2_2	cpSTS	0.674	0.594		Cognition.Language.Speech
SPL_R_5_1	A7ip	0.601	0.596		Interoception.Sexuality
Pcun_R_4_1	A7m	0.696	0.601		Perception.Vision.Motion
Pcun_R_4_2	A5m	0.606		0.602	Cognition.Social_Cognition
Pcun_L_4_4	A31	0.614		0.623	Cognition.Language
INS_L_6_2	vIa	0.631	0.591		Cognition
INS_L_6_3	dIa	0.605	0.620		
INS_R_6_3	dIa	0.644	0.588		Action.Inhibition
CG_R_7_3	A32p	0.597	0.588		Perception.Olfaction
CG_R_7_4	A23v	0.633	0.593	0.593	Emotion
Cun_L_5_1	cLinG	0.588	0.613		Perception.Vision
Cun_R_5_3	cCunG	0.668	0.613		Perception.Vision
OcG_L_4_4	iOccG	0.662	0.590		Perception.Vision
sOcG_L_2_2	lsOccG	0.588	0.596		Cognition.Reasoning
Hipp_L_2_2	cHipp	0.629		0.589	Cognition.Memory.Explicit
Str_R_6_6	dlPu	0.626	0.593		Action.Execution
Tha_L_8_1	mPFtha		0.588	0.594	Interoception.Bladder
Tha_L_8_2	mPMtha	0.606	0.597		Action.Execution
Tha_L_8_5	PPtha	0.604		0.607	Perception.Vision.Color
Tha_R_8_5	PPtha	0.600		0.603	Perception.Somesthesis.Pain
Cer_L_9	Crus I	0.697	0.615		
Cer_R_9	Crus I	0.687	0.598		

Because there were differences in the amount of in-scanner motion between groups, we also performed the classification analysis after matching the training sample for level of motion by splitting the data to 1–5 equally sized motion percentile groups based on their total RMS motion (1=no matching). The mean accuracies remained in a similar range for all motion matching scenarios (**Supplementary Table 2**) and classifier score were not significantly correlated with the amount of motion (**Supplementary Table 3**) with any of the matching strategies. This suggests that the classifier performance was not primarily motion related.

To explore whether symptom dimensions or other individual characteristics were associated higher likelihoods to be classified as a patient, we explored the correlations between classifier scores and the PAI-BOR total and sub-scale scores as well as demographic variables, focusing only on the SCID grouping. We found that there were significant differences in all the diagnostic scales between groups, as expected, but there were no consistent linear effects between the scales and classifier scores after controlling for mean group differences. That is, higher subjective ratings of BPD-like symptoms did not lead to higher likelihood that the classifier predicted within either group, or across the whole sample after controlling for group mean effects.

At the global level, we observed a significant correlation between group likelihood and RSPM scores ($r\sim0.7$) as well as education level ($r\sim0.6$) after controlling for the linear effects of these variables in the SCID II sample before classification. This relation was caused by regressing out the link-level effects of these variables; when the regression was omitted, the correlation was abolished ($r\sim-0.2$), but the classification accuracy at the global level (SCID II, no GSR) increased (up to ~80% with some combinations of sample size and motion matching). This suggests that differences of functional connectivity related to performance on the RSPM test may also modulate the same connectivity patterns that are important for predicting BPD diagnosis and removing these effects from the data through univariate linear regression may introduce unwanted effects into subsequent multivariate classification.

At the fine-scale connectivity between and within regions, we saw no significant prediction accuracies, similarly to the single link results. Although accuracies of 70–80% were observed for some pairs of ROIs and the distribution of accuracies over ROIs was shifted slightly to the right compared with the null distribution, no ROI pairs remained significant after correcting for multiple comparisons because high accuracies also existed in the null distribution. Moreover, at a more liberal threshold, the most predictive ROI were spread sparsely across the brain with no discernible pattern further suggesting that single connections may be unreliable predictors of BPD.

To investigate the role of medication on the predictive performance, we repeated classification between three groups: healthy controls, BPD patients without medication (BPD–Med), and BPD patients that were taking one or more medications for their condition (BPD+Med). These analyses were only performed with the SCID II based group definition strategy to reduce the number of combinations. The mean confusion matrix (over 100 iterations) for the three-class classification based on global connectivity matrices is presented in **Figure 6A**. As is apparent, the controls are successfully classified from both patient groups (chance level 33%). However, the medicated and unmedicated patients are confused with each other suggesting the biggest distinction is between HC and BPD irrespective of medication status.

Finally, to more thoroughly explore the effect of medication and its interaction with parameters such as spatial smoothing and global signal regression, we also performed classification between all pairs of groups (HC vs. BPD–Med, HC vs. BPD+Med, BPD–Med vs. BPD+Med). **Figure 6B** shows the distributions

of pairwise accuracies between all three pairs of groups. Accuracies are significant between HC and both patient groups, although accuracy for BPD+Med group falls just into chance level with global signal regression (from 64–65% to 58–59%). By contrast, in HC vs. BPD–Med analysis GSR slightly improves accuracy (59% to 61–62%). The accuracies between medicated and unmedicated patient groups are at chance level (52–53%). At the level of seed-regions, there was relatively little overlap between the three pairwise classifications. However, regions where controls are successfully classified from patient groups are similar to those showing most consistent accuracy across grouping strategies in **Figure 4**.

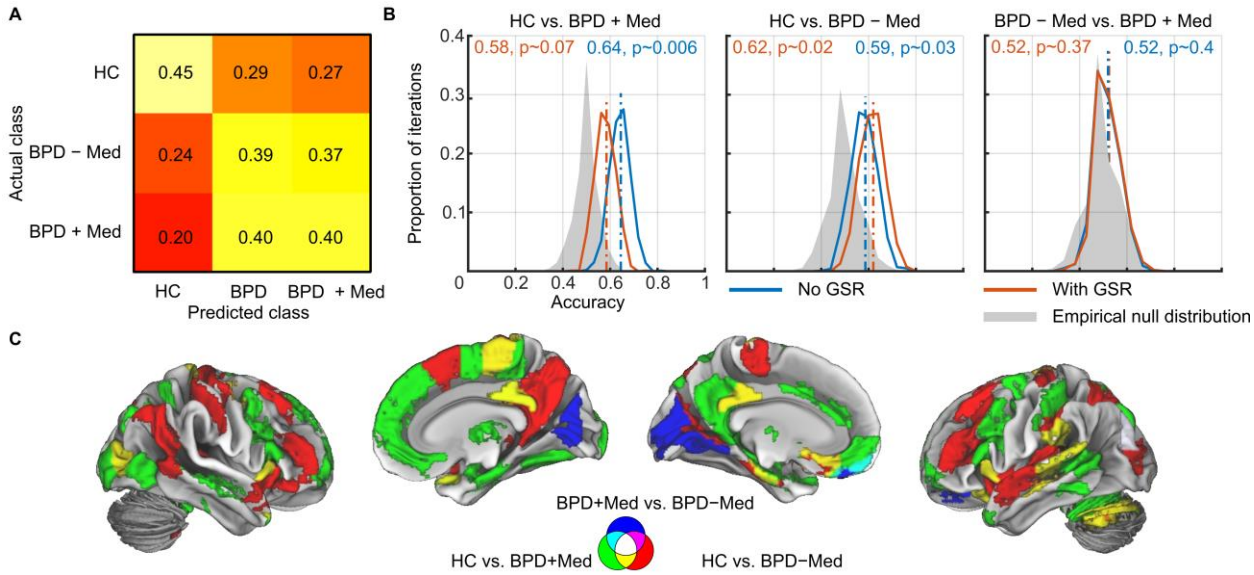


Figure 6: Effects of medication on global classification accuracies with SCID II grouping. ***A*** *Confusion matrix in three-class classifier between healthy controls (HC), patient without medication (BPD–MED) and patients with medication (BPD+Med).* ***B*** *Distributions of accuracies between pairs of matched subgroups. Between HC and BPD+Med accuracies are reduced by GSR, a pattern that is reduced and reversed when classifying between HC and BPD–MED groups. Accuracies between the two BPD groups remain at chance level.* ***C*** *ROI-maps showing the seed regions that were discriminative of the groups in each analysis. Blue indicates classification between BPD+Med and BPD–Med, red classification between HC and BPD+Med and green between HC and BPD–Med. Overlapping regions are shown in mixed colors.*

Discussion

Our results demonstrate that significant and wide-spread differences of functional connectivity exist in patients with borderline personality disorder compared with healthy control participants at multiple scales. By contrast, we observed no strong local univariate changes in connectivity, although some weak effects could be replicated in multiple non-overlapping subsamples. Our findings further underline that a thorough diagnostic procedure based on a structured clinical interview as well as a screening for personality disorders in the control group are associated with better discriminative performance in classification of patients. Clinically, our findings suggest that rsMRI may be sensitive to some aspects of psychopathology of BPD. However, additional work should be undertaken to evaluate whether data-driven classification schemes can predict outcome and treatment assignment, as is already starting to be done for other psychopathologies (e.g. 33–35).

Taken together, our multi-scale analyses indicate that the predictive power of functional connectivity is most consistent in cortical midline structures, posterior and anterior aspects of the superior temporal sulcus, and lateral parietal regions. These regions overlap with multiple different canonical functional networks that are associated with mentalizing and social processing, sensorimotor functions, action observation and bodily sensations/homeostatic monitoring. The former set of regions has recently been shown to increase in grey matter volume as a result of BPD-tailored psychotherapy (44). Prior findings on symptom-related alterations of connectivity in BPD have been rare and reported effects relatively small. This is also consistent with the lack of robust differences of connectivity using univariate contrasts in the current data. A recent study did find differences between BPD and HC individuals in degree centrality and fractional amplitude of low-frequency fluctuations in partly overlapping regions of the left posterior temporal lobe and precuneus as those shown in the current study and found the changes to be correlated with attachment scores in the patient population (15). Here, we used self-reported BPD and depression symptoms and medication status as predictors of classifier group likelihoods to assess their importance for classification. However, we observed no linear effects of the symptom scores beyond a mean group difference with the current symptom measures and sample size. Therefore, in future studies, larger cohorts and additional phenotyping would be important to map specific clinical aspects of the disorder to specific brain regions. For example, the predictive power of the somatosensory regions might be higher for those individuals showing repeated self-harming behaviors, which might manifest in different representation of bodily sensations. Or, for instance, increased symptoms of social paranoia or affect regulation could specifically be reflected in the connectivity of "social brain" regions. Comparing the different group specifications in the current study, self-report questionnaires of symptoms appear to be unreliable predictors of abnormal brain function and should be complemented by thorough clinician-based diagnostic procedures and quantitative interaction-based phenotyping (45).

While we consistently observed significantly higher accuracies than with null data based on relatively wide-spread connectivity patterns, the observed local differences are subtle with no significant univariate group differences in the link strengths after correcting for multiple comparisons. Our replication analysis revealed a number of weakly discriminative links that co-occurred with the most discriminative seed ROIs in the classification analysis. This suggest that the classification performance at larger spatial scales is likely driven by distributed patterns of small, but replicable connectivity changes. However, the lack of strong link-level effects or information on causal direction of functional connections limits the interpretability of the connectivity differences between groups. Moreover, the accuracies achieved with the connectivity-based machine learning approach varied relatively widely (from ~40 to over 90% in extreme cases) between repeated iterations of the same analysis pipeline with different assignments of training and test groups. This underlines the heterogeneity of the population

(there are effectively over 100 permissible symptom combinations, which would qualify for a DSM-based clinical diagnosis; 38) and suggests that multiple classifications with random group splits are crucial for estimating the true expected accuracy and its variability in machine learning approaches employing resting state functional connectivity. To further evaluate the generalizability of our classifiers, we applied the classifiers trained on balanced subsets to the unseen, left out participants. These results were largely consistent with those attained in the smaller matched samples, but the accuracies were inflated by nuisance regression in the inner cross-validation data compared to the left-out sample. Moreover, when the inner cross-validation sizes increased, we observed a negative correlation of inner vs. outer cross-validation accuracies as the outer cross-validation sample was increasingly constrained and imbalances between participant samples were exacerbated by the limited left-out sample size (Figure 5A). While this is not the perfect solution for evaluating generalizability due to the reliance on the same limited and dependent sample, it should help in avoiding focusing on the peculiarities of a single cross-validation sample. Ideally, however, larger samples with entirely independent replication data should be preferred in future studies. Moreover, while the current results are promising, their applicability to cross-sectional data with realistic prevalence of BPD is very limited; when prevalence is low, the positive predictive power of the classification drops drastically as most of the positive findings are false positives merely because of the relative overrepresentation of the healthy (non-BPD) group. Thus, significant advances are still required for clinical applications. Additionally, preprocessing procedures, such as spatial smoothing and global signal regression as well as linear regression of nuisance covariates across participants may affect the classification accuracies depending on the scale of observation. Also, although in the currents study, all scanning sites used the same scanner model, the effects of different scanning sites can cause variance in multi-site studies that may be difficult to control fully. Therefore, these effects should be carefully scrutinized in future studies.

It is notable that with the different types of symptoms in BPD and large variance in age and education level, the assessed patients represented a very heterogeneous group and therefore very high accuracies are unlikely. Indeed, exceedingly high accuracies are a cause for suspicion of dependency problems or that the classifier is fitting noise (47). Importantly, finding robust patterns of connectivity abnormalities in psychiatric disorders has proven difficult. For example, a recent meta-analysis (48) found no reproducible functional connectivity differences over 99 studies of unipolar depression highlighting the importance of careful scrutiny of reliability of individual findings. To improve the situation, within-study replication and careful consideration of stability of findings seem crucial in the future. Indeed, the current findings suggest that functional connectivity effects of intelligence and BPD interact in non-obvious ways as evidenced by the changes in global accuracy and correlations between group likelihood and RSPM scores due to nuisance regression.

Global signal regression has raised considerable controversy, particularly in the field of resting state functional connectivity (38). On one hand, it has been argued to reflect confounding artifacts of motion and physiology. On the other hand, regressing out the global signal has been mathematically shown to induce negative correlations in the data (38). In the current study, global signal regression generally reduced accuracies, however, the effect was small and it was even reversed when only non-medicated patients were classified against healthy controls. It seems, therefore, that spatially specific effects of the global signal may moderately improve the discriminative power of resting state functional connectivity in the current data, at least when medicated patients are included in the analyses. A recent report also showed that spatially distinct patterns of brain regions contributing to the global signal are predictive of schizophrenia (49). In the current study, we excluded participants with excessive motion, regressed out the motion time courses as well as the volumes containing large spikes of motion, and signals from the white matter and cerebrospinal fluid to combat artefactual global signal changes, but admittedly linear

regression strategies do not perfectly remove all traces of these signals and they may still affect multivariate classification results. Therefore, the specific contributions of motion, cardiac and breathing rhythms and brain signals of interest should still be explored further in the future.

The reason for observing higher accuracies for the "gold standard" SCID grouping strategy could be explained by real differences in brain connectivity in the excluded participants or mere technical issues due to different group sizes. To evaluate these options, we explored the classifier scores for the participants who were included in the Referral grouping, but were excluded from the SCID grouping, because the groupings were in other ways equivalent. The participants were excluded due to either not meeting diagnostic criteria (patients) or scoring over exclusion threshold using the SAPAS screening questionnaire for personality disorders (controls), suggesting that they might show intermediate connectivity patterns, in-between the two groups. We observed that the excluded participants were classified at chance level, thus explaining at least part of the lower accuracies in the Referral grouping. This tentatively supports the interpretation that the individuals not matching the criteria for either group exhibited brain connectivity patterns that were particularly difficult to classify as either BPD or control. Thus, it appears particularly important that the current diagnosis and symptoms are carefully confirmed before participants are included in such analyses, rather than relying on old diagnoses or self-reports. However, the low number of excluded participants and other individual differences (e.g. in motion, brain size or education) preclude strong interpretation of the causes that made these individuals particularly difficult to classify.

Importantly, in addition to group differences of interest, there are also other sources of variation in resting state functional connectivity. While the main source of variance seems to be due to individual factors (50), and intrinsic functional connectivity appears largely stable during rest and different audiovisual stimulation conditions (51), the test-retest reliability is not perfect and depends on the specific brain regions and the amount of available data (52), and moderate individual–task interactions may also exist (50). Reliability might be improved in future studies by providing the participants with an attention-inducing naturalistic stimulus, which appears to reduce participants' motion (53), at least in children. Moreover, an engaging stimulus may reduce the intraindividual variability of connectivity between subsequent measurements, which during resting state measurements could arise due to unconstrained, idiosyncratic patterns of thought. Reducing this unwanted variability could potentially even accentuate intergroup variability (for discussion and conceptual demonstration, see (54)).

A significant limitation for the specificity of the current findings is the lack of other clinical groups in addition to BPD in our sample. Specifically, the general level of psychopathology, sometimes referred to as the "p-factor", appears to affect largely similar brain regions that were predictive of BPD in the current study (55). Moreover, the current study did not include a thorough diagnosis of comorbid psychiatric disorders, which may further blur the line between disorders. Thus, the findings presented here could be disorder-general rather than specific to BPD. A recent reanalysis of two large representative surveys of US adults (56) confirmed the general psychopathology factor across psychopathologies. In particular, latent BPD was found to closely reflect the general psychopathology factor and could even be considered a unitary construct rather than two separate entities. Thus, identifying unique mechanism for BPD could enable better understanding of the nature of general psychopathology. However, disentangling signals unique to BPD that are not associated with severity other psychopathologies appears challenging due to the strong association with p factor and BPD. It is, thus, crucial that future analyses extend the current findings in transdiagnostic samples, which is becoming increasingly feasible due to large-scale projects like the UK Biobank. Moreover, categorical comparisons between heterogeneous psychiatric disorder labels, although efficient in clinical practice using traditional diagnostic tools, may not be the most appropriate way of evaluating the brain

mechanisms of psychiatric symptoms, which may be better understood by dimensional approaches to psychopathology. In the current case, we explored only self-reported symptoms, which in this case appeared to be weaker predictors of connectivity than categorical classification based on the SCID interview. High interrater reliabilities (Cronbach's Kappa ≥ 0.89) have generally been reported for the SCID interview, although some studies or study sites have produced lower reliabilities both between raters and in test-retest evaluations (57). However, more objective and reliable dimensional models of psychopathology may be particularly fruitful for elucidating the interindividual variability both within and between categorical diagnostic labels.

Due to the large number of comparisons, we have here limited the machine learning methods to only simple linear SVM classifiers. A different choice of classifier could also have been more successful at detecting subtle differences in regions not highlighted by the current methodology. Here, we chose a simple linear classifier to limit problems with overfitting in high-dimensional data that a non-linear classifier could be more susceptible to, given the limited sample size. However, a non-linear classifier could be more sensitive to effects that are not accessible to linear methods. Potentially, also optimizing the hyperparameters of the current model could also have improved the performance. Future studies should further explore the optimal choices of machine learning methods and associated parameters for clinical prediction of BPD, including both continuous regression and categorical classification methods, preferably in a large transdiagnostic sample.

Finally, medications may confound the interpretation of the results in studies on patients with disorders such as BPD, who are often on at least one medication affecting the central nervous system. In the current study, we repeated the analyses on a subset of patients with no current medication. We observed a slightly reduced accuracy specifically in analyses employing data without global signal regression – an effect that was reversed in the data where global signals were regressed out. This might suggest that the drugs taken by the patient group (mainly antidepressant and antipsychotic medication) may cause global connectivity differences that inflate predictive performance of classifiers based on connectivity. However, it is unclear whether this effect is due to neural signals or is driven by random variability. For example, differences in physiological signals or motion that was not removed by the motion regression and data exclusion procedures. Moreover, additional investigations are required to confirm this observation because we cannot rule out alternative explanations, such as other characteristics of the unmedicated vs. medicated patient samples or the smaller pool of healthy controls that might have contributed to the differences in the accuracies.

Conclusions

Our results demonstrate that widespread functional connectivity patterns reliably discriminate between BPD and control participants. However, only a subset of these regions were consistently able to classify patients irrespective of medication as well as group selection and preprocessing strategies suggesting considerable heterogeneity in the patient population that may be related to the current or past use of medications and symptom differences. Importantly, the classification results generalized to unbalanced left-out participants. However, the specificity of these finding should be further evaluated in relation to other psychiatric disorders. Local link-level differences showed only weak differences between groups, which did not survive corrections for multiple comparisons. Our replication analysis further suggests that only a small minority of these effects were likely to replicate and a subset even showed significant opposite effects, which may explain the lack of consensus in the previous literature. Therefore, it is imperative that the reliability of observations is carefully scrutinized and larger, well-characterized samples including multiple patient groups are favored in future studies to avoid discrepant findings and

over-interpretation of the spatial patterns of functional differences associated with psychiatric disorders.

Bibliography

1. Schilbach L (2016): Towards a second-person neuropsychiatry. *Philos Trans R Soc Lond B Biol Sci* 371: 20150081.

2. Fonagy P, Luyten P (2009): A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Dev Psychopathol* 21: 1355–1381.

3. Reisch T, Ebner-Priemer UW, Tschacher W, Bohus M, Linehan MM (2008): Sequences of emotions in patients with borderline personality disorder. *Acta Psychiatr Scand* 118: 42–48.

4. Grootens KP, van Luijtelaar G, Buitelaar JK, van der Laan A, Hummelen JW, Verkes RJ (2008): Inhibition errors in borderline personality disorder with psychotic-like symptoms. *Prog Neuro-Psychopharmacology Biol Psychiatry* 32: 267–273.

5. Black DW, Blum N, Pfohl B, Hale N (2004): Suicidal Behavior in Borderline Personality Disorder: Prevalence, Risk Factors, Prediction, and Prevention. *J Pers Disord* 18: 226–239.

6. Fonagy P, Allison E (2014): The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy* 51: 372–380.

7. Critchfield KL, Levy KN, Clarkin JF, Kernberg OF (2007): The relational context of aggression in borderline personality disorder: using adult attachment style to predict forms of hostility. *J Clin Psychol* 64: 67–82.

8. Allen JG, Fonagy P, Bateman A (2008): *Mentalizing in Clinical Practice*. American Psychiatric Pub.

9. Euler S, Nolte T, Constantinou M, Griem J, Montague PR, Fonagy P (2019): Interpersonal Problems in Borderline Personality Disorder: Associations with Mentalizing, Emotion Regulation, and Impulsiveness. *J Pers Disord* 1–17.

10. Debbané M, Nolte T (2019): Mentalization-Based Therapy in the Light of Contemporary Neuroscientific Research. In: Bateman A, Fonagy P, editors. *Handbook of Mentalizing in Mental Health*. Washington DC, USA: American Psychiatric Publishing.

11. Nolte T, Campbell C, Fonagy P (2019): A Mentalization-Based and Neuroscience-Informed Model of Severe and Persistent Psychopathology. In: Pereira JG, Gonçalves J, Bizzari V, editors. *The Psychotherapy-Neurobiology-Pharmacology Intervention Triangle*. Malaga, Spain, pp 161–184.

12. Schulze L, Schmahl C, Niedtfeld I (2016): Neural Correlates of Disturbed Emotion Processing in Borderline Personality Disorder: A Multimodal Meta-Analysis. *Biol Psychiatry* 79: 97–106.

13. Herpertz SC, Bertsch K, Jeung H (2018): Neurobiology of Criterion A: self and interpersonal personality functioning. *Curr Opin Psychol* 21: 23–27.

14. Quattrini G, Pini L, Pievani M, Magni LR, Lanfredi M, Ferrari C, *et al.* (2019): Abnormalities in functional connectivity in borderline personality disorder: Correlations with metacognition and

emotion dysregulation. *Psychiatry Res Neuroimaging* 283: 118–124.

15. Lei X, Liao Y, Zhong M, Peng W, Liu Q, Yao S, *et al.* (2018): Functional Connectivity Density, Local Brain Spontaneous Activity, and Their Coupling Strengths in Patients With Borderline Personality Disorder. *Front Psychiatry* 9. https://doi.org/10.3389/fpsyt.2018.00342

16. Salvador R, Vega D, Pascual JC, Marco J, Canales-Rodr\'{\i}guez EJ, Aguilar S, *et al.* (2016): Converging Medial Frontal Resting State and Diffusion-Based Abnormalities in Borderline Personality Disorder. *Biol Psychiatry* 79: 107–116.

17. O'Neill A, D'Souza A, Samson AC, Carballedo A, Kerskens C, Frodl T (2015): Dysregulation between emotion and theory of mind networks in borderline personality disorder. *Psychiatry Res Neuroimaging* 231: 25–32.

18. Wolf R (2011): Aberrant connectivity of resting-state networks in borderline personality disorder. *J Psychiatry Neurosci* 36: 402–411.

19. Krause-Utz A, Schmahl C (2016): A More Global Look at Altered Neural Structure and Resting-State Function in Borderline Personality Disorder. *Biol Psychiatry* 79: 76–77.

20. New AS, Carpenter DM, Perez-Rodriguez MM, Ripoll LH, Avedon J, Patil U, *et al.* (2013): Developmental differences in diffusion tensor imaging parameters in borderline personality disorder. *J Psychiatr Res* 47: 1101–1109.

21. Ninomiya T, Oshita H, Kawano Y, Goto C, Matsuhashi M, Masuda K, *et al.* (2018): Reduced white matter integrity in borderline personality disorder: A diffusion tensor imaging study. *J Affect Disord* 225: 723–732.

22. Goldstein KE, Haznedar MM, Alloy LB, Drabick DAG, McClure MM, New AS, *et al.* (2019): Short communication: Diffusion tensor anisotropy in the cingulate in borderline and schizotypal personality disorder. *Psychiatry Res* 279: 353–357.

23. Xu T, Cullen KR, Mueller B, Schreiner MW, Lim KO, Schulz SC, Parhi KK (2016): Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI. *NeuroImage Clin* 11: 302–315.

24. Baczkowski BM, van Zutphen L, Siep N, Jacob GA, Domes G, Maier S, *et al.* (2017): Deficient amygdala–prefrontal intrinsic connectivity after effortful emotion regulation in borderline personality disorder. *Eur Arch Psychiatry Clin Neurosci* 267: 551–565.

25. Lei X, Zhong M, Liu Y, Jin X, Zhou Q, Xi C, *et al.* (2017): A resting-state fMRI study in borderline personality disorder combining amplitude of low frequency fluctuation, regional homogeneity and seed based functional connectivity. *J Affect Disord* 218: 299–305.

26. Lei X, Zhong M, Zhang B, Yang H, Peng W, Liu Q, *et al.* (2019): Structural and Functional Connectivity of the Anterior Cingulate Cortex in Patients With Borderline Personality Disorder. *Front Neurosci* 13: 1–9.

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

28. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, *et al.* (2004): Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23: S208–S219.

29. Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, *et al.* (2009): Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45: S173–S186.

30. Jenkinson M, Bannister P, Brady M, Smith S (2002): Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *Neuroimage* 17: 825–841.

31. Smith SM (2002): Fast robust automated brain extraction. *Hum Brain Mapp* 17: 143–155.

32. Moran P, Leese M, Lee T, Walters P, Thornicroft G, Mann A (2003): Standardised Assessment of Personality Abbreviated Scale (SAPAS): Preliminary validation of a brief screen for personality disorder. *Br J Psychiatry* 183: 228–232.

33. Morey LC (1991): Personality Assessment Inventory. American Psychological Association (APA). https://doi.org/10.1037/t03903-000

34. RAVEN JC (1941): STANDARDIZATION OF PROGRESSIVE MATRICES, 1938. *Br J Med Psychol* 19: 137–150.

35. Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, *et al.* (2016): The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb Cortex* 26: 3508–3526.

36. Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT (2011): The organization of the human cerebellum estimated by intrinsic functional connectivity. *J Neurophysiol* 106: 2322–2345.

37. Murphy K, Fox MD (2017): Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154: 169–173.

38. Gallagher HL, Happé F, Brunswick N, Fletcher PC, Frith U, Frith CD (2000): Reading the mind in cartoons and stories: an fMRI study of `theory of mind' in verbal and nonverbal tasks. *Neuropsychologia* 38: 11–21.

39. Lahnakoski JM, Glerean E, Salmi J, Jääskeläinen IP, Sams M, Hari R, Nummenmaa L (2012): Naturalistic fMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Front Hum Neurosci*.

40. Dichter GS, Gibbs D, Smoski MJ (2015): A systematic review of relations between resting-state functional-MRI and treatment response in major depressive disorder. *J Affect Disord* 172: 8–17.

41. Plitt M, Barnes KA, Wallace GL, Kenworthy L, Martin A (2015): Resting-state functional connectivity predicts longitudinal change in autistic traits and adaptive functioning in autism. *Proc Natl Acad Sci* 112: E6699–E6706.

42. Whitfield-Gabrieli S, Ghosh SS, Nieto-Castanon A, Saygin Z, Doehrmann O, Chai XJ, *et al.* (2015): Brain connectomics predict response to treatment in social anxiety disorder. *Mol Psychiatry* 21: 680–685.

43. Mancke F, Schmitt R, Winter D, Niedtfeld I, Herpertz SC, Schmahl C (2018): Assessing the marks of change: how psychotherapy alters the brain structure in women with borderline personality

disorder. *J Psychiatry Neurosci* 43: 171–181.

44. Schilbach L (2019): Using interaction-based phenotyping to assess the behavioral and neural mechanisms of transdiagnostic social impairments in psychiatry. *Eur Arch Psychiatry Clin Neurosci* 269: 273–274.

45. Herbort MC, Soch J, Wüstenberg T, Krauel K, Pujara M, Koenigs M, *et al.* (2016): A negative relationship between ventral striatal loss anticipation response and impulsivity in borderline personality disorder. *NeuroImage Clin* 12: 724–736.

46. Whelan R, Garavan H (2014): When Optimism Hurts: Inflated Predictions in Psychiatric Neuroimaging. *Biol Psychiatry* 75: 746–748.

47. Müller VI, Cieslik EC, Serbanescu I, Laird AR, Fox PT, Eickhoff SB (2017): Altered Brain Activity in Unipolar Depression Revisited. *JAMA Psychiatry* 74: 47.

48. Yang GJ, Murray JD, Glasser M, Pearlson GD, Krystal JH, Schleifer C, *et al.* (2016): Altered Global Signal Topography in Schizophrenia. *Cereb Cortex*. https://doi.org/10.1093/cercor/bhw297

49. Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, Gilmore AW, *et al.* (2018): Functional Brain Networks Are Dominated by Stable Group and Individual Factors, Not Cognitive or Daily Variation. *Neuron* 98: 439-452.e5.

50. Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016): Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7. https://doi.org/10.1038/ncomms12141

51. Noble S, Spann MN, Tokoglu F, Shen X, Todd Constable R, Scheinost D (2017): Influences on the Test-Retest Reliability of Functional Connectivity MRI and its Relationship with Behavioral Utility. *Cereb Cortex* 27: 5415–5429.

52. Vanderwal T, Kelly C, Eilbott J, Mayes LC, Castellanos FX (2015): Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage* 122: 222–232.

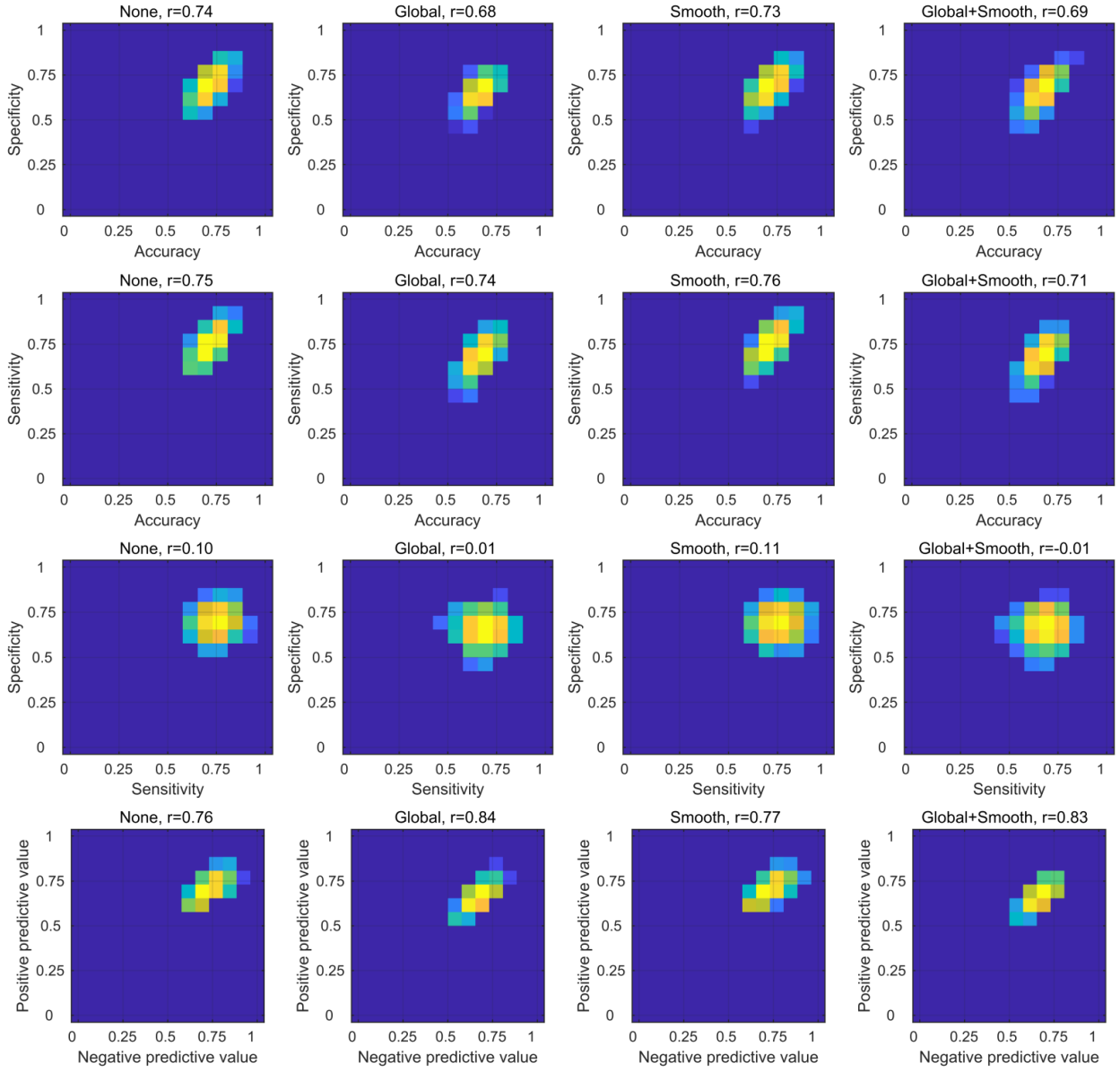
53. Finn ES, Glerean E, Khojandi AY, Nielson D, Molfese PJ, Handwerker DA, Bandettini PA (n.d.): Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. https://doi.org/10.31234/OSF.IO/YEU89

54. Elliott ML, Romer A, Knodt AR, Hariri AR (2018): A Connectome-wide Functional Signature of Transdiagnostic Risk for Mental Illness. *Biol Psychiatry* 84: 452–459.

55. Gluschkoff K, Jokela M, Rosenström T (2021): General psychopathology factor and borderline personality disorder: Evidence for substantial overlap from two nationally representative surveys of U.S. adults. *Personal Disord Theory, Res Treat* 12: 86–92.

56. Carcone D, Tokarz VL, Ruocco AC (2015): A systematic review on the reliability and validity of semistructured diagnostic interviews for borderline personality disorder. *Can Psychol Can* 56: 208–226.

Supplementary materials



Supplementary Figure 1: Dependence of specificity, sensitivity and accuracy over 5000 iterations of the linear classification based on global connectivity structure with different preprocessing strategies. *The distributions are depicted as the normalized log frequencies in 2-dimensional binned histograms. Specificity and sensitivity are both correlated with the overall accuracy. During most iterations specificity and sensitivity also both are near 70%, as evidenced by the peak location of the joint distribution, and the correlations between the two are low.*

Supplementary Table 1: Accuracies and MNI coordinates ROIs in seed-based ROI classification.

	MNI Coordinates			Classification accuracy		
Region label	**X**	**Y**	**Z**	**SCID II**	**Referral**	**PAI-BOR**
1 SFG_L_7_1	-2.11	18.38	56.99			
2 SFG_R_7_1	9.50	19.26	57.20			
3 SFG_L_7_2	-15.55	26.76	55.65	0.629		
4 SFG_R_7_2	24.48	28.45	54.59	0.688		
5 SFG_L_7_3	-8.58	52.25	42.38	0.631		
6 SFG_R_7_3	15.72	51.52	43.15	0.633	0.588	
7 SFG_L_7_4	-15.51	2.30	68.01	0.629		
8 SFG_R_7_4	22.91	7.09	67.18			
9 SFG_L_7_5	-3.30	-2.14	60.84			
10 SFG_R_7_5	10.08	-0.95	63.17			
11 SFG_L_7_6	-1.89	39.41	41.10			
12 SFG_R_7_6	8.83	41.47	38.07	0.619	0.603	
13 SFG_L_7_7	-4.62	59.63	18.00			
14 SFG_R_7_7	10.63	61.73	15.74	0.618		
15 MFG_L_7_1	-25.01	46.03	33.82			
16 MFG_R_7_1	33.14	40.30	38.74			
17 MFG_L_7_2	-39.28	16.36	39.22	0.591		
18 MFG_R_7_2	45.19	14.51	41.51	0.599		
19 MFG_L_7_3	-25.48	59.39	14.59		0.593	
20 MFG_R_7_3	30.66	58.56	19.56			0.607
21 MFG_L_7_4	-38.28	44.48	18.73			
22 MFG_R_7_4	44.91	47.71	16.24			
23 MFG_L_7_5	-30.52	26.24	48.21	0.605		
24 MFG_R_7_5	44.93	30.07	41.87			
25 MFG_L_7_6	-29.94	6.85	57.55	0.592		
26 MFG_R_7_6	36.59	10.89	57.60			
27 MFG_L_7_7	-22.53	64.10	-2.87			
28 MFG_R_7_7	28.58	64.83	-1.25			
29 IFG_L_6_1	-43.35	16.65	26.59			
30 IFG_R_6_1	48.78	19.82	27.99			
31 IFG_L_6_2	-45.23	35.41	16.30	0.615		
32 IFG_R_6_2	51.10	38.68	15.66			
33 IFG_L_6_3	-50.26	26.37	13.51	0.690	0.625	0.590
34 IFG_R_6_3	57.19	26.94	13.96			
35 IFG_L_6_4	-46.53	39.69	0.05			
36 IFG_R_6_4	54.28	39.62	2.11			
37 IFG_L_6_5	-36.94	25.64	6.54	0.588		
38 IFG_R_6_5	44.75	24.87	6.17	0.614		
39 IFG_L_6_6	-49.32	16.81	8.83			

40	IFG_R_6_6	56.45	16.71	13.15			
41	OrG_L_6_1	-3.59	57.31	-4.26	0.641		
42	OrG_R_6_1	9.16	50.58	-4.72			
43	OrG_L_6_2	-33.77	36.46	-13.36	0.642		
44	OrG_R_6_2	43.20	42.01	-11.82	0.599		
45	OrG_L_6_3	-19.81	40.89	-15.39			
46	OrG_R_6_3	26.26	39.35	-15.88			
47	OrG_L_6_4	-3.29	55.32	-17.20			
48	OrG_R_6_4	9.10	59.63	-14.58			
49	OrG_L_6_5	-7.49	20.50	-16.48			
50	OrG_R_6_5	11.99	23.32	-17.23			
51	OrG_L_6_6	-38.68	35.89	-6.79	0.604		
52	OrG_R_6_6	45.63	34.86	-6.41			
53	PrG_L_6_1	-46.49	-4.84	41.90			
54	PrG_R_6_1	57.69	0.39	35.93			
55	PrG_L_6_2	-29.19	-6.19	60.84			
56	PrG_R_6_2	36.17	-3.76	59.42	0.694	0.607	
57	PrG_L_6_3	-23.93	-22.03	65.78	0.609		
58	PrG_R_6_3	37.10	-15.92	61.80			
59	PrG_L_6_4	-10.93	-17.66	76.01	0.633		
60	PrG_R_6_4	18.17	-18.81	73.85	0.596		
61	PrG_L_6_5	-49.90	3.31	10.43	0.606		
62	PrG_R_6_5	56.62	6.95	11.74			
63	PrG_L_6_6	-46.74	7.56	33.44			
64	PrG_R_6_6	54.14	10.06	33.70			
65	PCL_L_2_1	-4.87	-35.45	61.74			
66	PCL_R_2_1	12.78	-31.67	57.02	0.629		0.638
67	PCL_L_2_2	-1.32	-19.93	63.75	0.624	0.593	
68	PCL_R_2_2	7.59	-18.07	64.40			
69	STG_L_6_1	-29.43	16.90	-32.34			
70	STG_R_6_1	34.33	18.34	-31.53			
71	STG_L_6_2	-51.64	-28.73	14.84			
72	STG_R_6_2	57.18	-20.69	13.35	0.597		
73	STG_L_6_3	-47.57	-8.10	4.15	0.627	0.612	
74	STG_R_6_3	53.88	-0.67	1.51	0.595		
75	STG_L_6_4	-59.99	-30.02	9.60	0.606		
76	STG_R_6_4	69.36	-17.08	8.66			
77	STG_L_6_5	-42.59	14.31	-17.25	0.638	0.598	
78	STG_R_6_5	50.12	15.64	-17.28			
79	STG_L_6_6	-52.53	-0.22	-7.91	0.624	0.625	0.591
80	STG_R_6_6	59.45	-9.00	-2.51			
81	MTG_L_4_1	-63.02	-26.95	-9.12	0.628		
82	MTG_R_4_1	68.01	-25.54	-10.87			

83 MTG_L_4_2	-50.88	5.22	-27.22	0.591	
84 MTG_R_4_2	54.32	8.71	-29.66	0.663	
85 MTG_L_4_3	-56.69	-54.86	7.01	0.605	0.602
86 MTG_R_4_3	63.27	-50.85	5.34	0.588	
87 MTG_L_4_4	-56.15	-16.60	-7.10	0.604	
88 MTG_R_4_4	61.13	-13.08	-7.35		
89 ITG_L_7_1	-43.05	-23.15	-24.95		
90 ITG_R_7_1	48.90	-11.21	-30.62		
91 ITG_L_7_2	-48.79	-54.61	-12.49	0.595	0.601
92 ITG_R_7_2	56.33	-50.06	-15.77	0.597	
93 ITG_L_7_3	-41.18	0.74	-38.98		
94 ITG_R_7_3	43.89	3.41	-40.55		
95 ITG_L_7_4	-53.56	-12.66	-25.02		
96 ITG_R_7_4	57.78	-8.04	-29.97		
97 ITG_L_7_5	-52.90	-57.51	-3.10	0.614	
98 ITG_R_7_5	57.03	-54.29	-5.67		
99 ITG_L_7_6	-57.02	-39.18	-13.41		
100 ITG_R_7_6	63.92	-37.21	-14.85		
101 ITG_L_7_7	-52.39	-28.09	-24.54		
102 ITG_R_7_7	56.57	-28.77	-23.38		
103 FuG_L_3_1	-30.33	-13.98	-29.41		
104 FuG_R_3_1	36.48	-12.05	-31.23		
105 FuG_L_3_2	-28.48	-62.37	-11.70	0.603	
106 FuG_R_3_2	34.06	-58.92	-11.08	0.624	
107 FuG_L_3_3	-39.76	-47.43	-15.06		
108 FuG_R_3_3	45.45	-46.55	-16.26		
109 PhG_L_6_1	-24.66	-4.25	-31.77		
110 PhG_R_6_1	30.81	-5.32	-30.98	0.662	
111 PhG_L_6_2	-22.12	-22.58	-23.47	0.596	
112 PhG_R_6_2	29.33	-20.40	-24.49	0.619	
113 PhG_L_6_3	-25.85	-28.67	-15.72	0.659	0.592
114 PhG_R_6_3	32.87	-26.73	-15.72		
115 PhG_L_6_4	-16.17	-9.45	-27.53		
116 PhG_R_6_4	21.95	-8.35	-26.92		
117 PhG_L_6_5	-21.04	4.97	-29.49		
118 PhG_R_6_5	24.70	4.00	-33.76		
119 PhG_L_6_6	-14.48	-36.74	-7.28		
120 PhG_R_6_6	21.95	-33.58	-8.90		
121 pSTS_L_2_1	-52.21	-37.20	6.75	0.619	
122 pSTS_R_2_1	55.74	-34.17	6.09	0.605	
123 pSTS_L_2_2	-50.06	-47.53	13.41	0.674	0.594
124 pSTS_R_2_2	59.87	-37.37	15.44		
125 SPL_L_5_1	-13.65	-57.26	65.59	0.679	

126	SPL_R_5_1	22.32	-54.18	67.92	0.601	0.596	
127	SPL_L_5_2	-12.73	-68.49	54.62	0.605		
128	SPL_R_5_2	21.52	-66.40	56.53	0.605		
129	SPL_L_5_3	-30.65	-44.48	52.56	0.587		
130	SPL_R_5_3	38.05	-39.06	56.97			
131	SPL_L_5_4	-19.95	-44.77	68.16	0.624		
132	SPL_R_5_4	26.09	-40.26	69.83	0.587		
133	SPL_L_5_5	-24.87	-56.58	56.70	0.633		
134	SPL_R_5_5	34.15	-51.36	56.13	0.613		
135	IPL_L_6_1	-31.70	-77.89	31.63	0.624		
136	IPL_R_6_1	48.42	-68.88	23.22	0.588		
137	IPL_L_6_2	-35.48	-58.52	49.33			
138	IPL_R_6_2	42.38	-62.65	46.37	0.617		
139	IPL_L_6_3	-48.76	-30.62	44.45			
140	IPL_R_6_3	50.30	-32.32	48.21			
141	IPL_L_6_4	-53.50	-46.73	40.17			
142	IPL_R_6_4	60.33	-40.89	41.03			
143	IPL_L_6_5	-44.18	-62.55	28.80	0.636		
144	IPL_R_6_5	56.08	-51.81	27.55	0.595		
145	IPL_L_6_6	-51.02	-28.45	25.28	0.676		
146	IPL_R_6_6	58.18	-23.21	28.25	0.612		
147	Pcun_L_4_1	-2.25	-60.95	54.00	0.644		
148	Pcun_R_4_1	8.94	-62.45	53.86	0.696	0.601	
149	Pcun_L_4_2	-5.14	-44.69	60.07			
150	Pcun_R_4_2	10.33	-44.47	61.55	0.606		0.602
151	Pcun_L_4_3	-9.19	-64.11	28.26			
152	Pcun_R_4_3	19.29	-61.10	27.22	0.636		
153	Pcun_L_4_4	-3.34	-52.07	36.48	0.614		0.623
154	Pcun_R_4_4	9.35	-51.31	37.46			0.591
155	PoG_L_4_1	-47.18	-13.82	46.23			
156	PoG_R_4_1	52.59	-11.72	47.35			
157	PoG_L_4_2	-53.22	-11.23	19.04			
158	PoG_R_4_2	58.90	-6.87	17.47			
159	PoG_L_4_3	-43.48	-27.02	52.62	0.612		
160	PoG_R_4_3	51.89	-20.68	50.57			
161	PoG_L_4_4	-19.03	-31.94	71.33			
162	PoG_R_4_4	23.04	-29.65	72.90	0.651		
163	INS_L_6_1	-33.95	-16.89	11.80			
164	INS_R_6_1	40.70	-14.67	10.42			
165	INS_L_6_2	-29.60	17.42	-10.27	0.631	0.591	
166	INS_R_6_2	35.98	17.78	-10.24			
167	INS_L_6_3	-31.93	20.27	3.82	0.605	0.620	
168	INS_R_6_3	39.74	21.04	3.07	0.644	0.588	

169	INS_L_6_4	-35.82	-1.27	-6.53			
170	INS_R_6_4	42.20	0.67	-6.60			
171	INS_L_6_5	-36.26	-4.79	10.93			
172	INS_R_6_5	42.22	-3.80	10.11			
173	INS_L_6_6	-35.87	8.24	6.42	0.591		
174	INS_R_6_6	41.19	8.41	7.45			
175	CG_L_7_1	-1.25	-36.64	33.88			0.590
176	CG_R_7_1	7.13	-33.65	34.27	0.615		
177	CG_L_7_2	-0.92	12.34	26.94			
178	CG_R_7_2	7.54	24.61	15.17			
179	CG_L_7_3	-2.84	37.60	23.17			
180	CG_R_7_3	8.21	30.74	30.29	0.597	0.588	
181	CG_L_7_4	-5.84	-44.73	12.02	0.660		
182	CG_R_7_4	11.47	-40.37	14.26	0.633	0.593	0.593
183	CG_L_7_5	-1.91	10.45	39.86	0.606		
184	CG_R_7_5	7.00	9.19	40.95			
185	CG_L_7_6	-4.79	-20.14	43.74	0.610		
186	CG_R_7_6	9.20	-17.03	43.31			
187	CG_L_7_7	-1.64	42.09	0.55	0.622		
188	CG_R_7_7	8.11	44.10	8.90			0.597
189	Cun_L_5_1	-8.05	-79.94	-8.25	0.588	0.613	
190	Cun_R_5_1	13.26	-83.08	-6.11			
191	Cun_L_5_2	-2.04	-78.53	12.89		0.608	
192	Cun_R_5_2	10.09	-73.28	13.81		0.654	
193	Cun_L_5_3	-2.84	-92.02	3.81	0.663		
194	Cun_R_5_3	11.60	-87.33	14.94	0.668	0.613	
195	Cun_L_5_4	-13.46	-58.05	-4.12			
196	Cun_R_5_4	20.78	-57.18	-4.47			
197	Cun_L_5_5	-10.39	-65.87	14.69	0.591		
198	Cun_R_5_5	17.28	-60.79	14.83			
199	OcG_L_4_1	-28.41	-86.73	13.47	0.597		
200	OcG_R_4_1	37.30	-83.59	13.37	0.709		
201	OcG_L_4_2	-43.53	-71.33	5.73			
202	OcG_R_4_2	50.94	-67.67	2.01			
203	OcG_L_4_3	-15.41	-97.07	5.22	0.603		
204	OcG_R_4_3	25.20	-94.87	7.12			
205	OcG_L_4_4	-27.69	-85.66	-9.83	0.662	0.590	
206	OcG_R_4_4	35.23	-82.16	-8.97	0.641		
207	sOcG_L_2_1	-8.16	-85.81	33.25			
208	sOcG_R_2_1	19.23	-82.78	37.54			
209	sOcG_L_2_2	-19.59	-74.78	39.06	0.588	0.596	
210	sOcG_R_2_2	31.73	-71.97	38.78	0.659		
211	Amyg_L_2_1	-15.88	0.64	-17.15	0.612		

212 Amyg_R_2_1	22.25	0.52	-16.68			
213 Amyg_L_2_2	-24.76	-1.16	-17.56	0.599		
214 Amyg_R_2_2	30.73	-0.47	-16.80			
215 Hipp_L_2_1	-19.28	-10.60	-16.66	0.653		
216 Hipp_R_2_1	24.68	-9.24	-17.99	0.603		
217 Hipp_L_2_2	-25.06	-27.30	-7.35	0.629		0.589
218 Hipp_R_2_2	32.39	-24.09	-7.95			
219 Str_L_6_1	-9.47	17.61	2.14	0.601		
220 Str_R_6_1	17.69	18.02	0.15			
221 Str_L_6_2	-20.28	0.87	6.44			
222 Str_R_6_2	24.81	0.77	5.73	0.667		
223 Str_L_6_3	-14.55	6.43	-6.70	0.600		
224 Str_R_6_3	18.07	9.10	-6.08			
225 Str_L_6_4	-20.17	9.78	-0.69	0.638		
226 Str_R_6_4	25.12	10.40	0.67			
227 Str_L_6_5	-12.61	4.63	18.72			
228 Str_R_6_5	16.89	8.40	17.43			
229 Str_L_6_6	-26.49	-2.51	4.02			
230 Str_R_6_6	32.90	-0.52	3.61	0.626	0.593	
231 Tha_L_8_1	-3.82	-9.20	7.67		0.588	0.594
232 Tha_R_8_1	9.47	-8.53	8.35		0.635	
233 Tha_L_8_2	-15.69	-10.36	5.52	0.606	0.597	
234 Tha_R_8_2	15.46	-10.57	3.76			
235 Tha_L_8_3	-14.64	-20.05	6.32	0.605		
236 Tha_R_8_3	20.68	-18.52	5.88	0.626		
237 Tha_L_8_4	-3.27	-10.55	9.65		0.605	
238 Tha_R_8_4	5.39	-10.40	8.38		0.602	
239 Tha_L_8_5	-13.41	-21.58	8.99	0.604		0.607
240 Tha_R_8_5	18.27	-22.51	8.78	0.600		0.603
241 Tha_L_8_6	-12.28	-24.88	6.73			
242 Tha_R_8_6	15.70	-24.31	10.44			
243 Tha_L_8_7	-8.64	-18.58	15.64			
244 Tha_R_8_7	12.44	-11.34	16.89	0.590		
245 Tha_L_8_8	-8.78	-11.84	5.18			
246 Tha_R_8_8	14.75	-12.79	8.98			0.614
247 Cer_L_9	-4.75	-42.65	-14.56			
248 Cer_R_9	11.32	-41.61	-15.59			
249 Cer_L_9	-10.60	-48.69	-16.16			0.597
250 Cer_R_9	15.82	-49.24	-16.48	0.622		
251 Cer_L_9	-20.51	-57.13	-22.53			
252 Cer_V_9	2.59	-68.46	-18.95			

253 Cer_R_9	26.29	-56.32	-22.76	0.606		
254 Cer_L_9	-33.71	-66.56	-29.68	0.697	0.615	
255 Cer_R_9	39.22	-65.67	-29.51	0.687	0.598	
256 Cer_L_9	-23.95	-73.39	-39.90			
257 Cer_V_9	2.57	-72.68	-29.06	0.591		
258 Cer_R_9	28.19	-73.92	-38.98			
259 Cer_L_9	-23.94	-64.66	-48.99			
260 Cer_V_9	1.22	-66.00	-29.11			
261 Cer_R_9	30.03	-64.10	-48.63			
262 Cer_L_9	-22.10	-55.81	-50.97			
263 Cer_V_9	2.27	-65.18	-35.96			
264 Cer_R_9	27.53	-56.44	-51.16			
265 Cer_L_9	-14.94	-48.56	-52.65			
266 Cer_V_9	1.90	-61.37	-39.51	0.595		
267 Cer_R_9	19.72	-49.25	-52.74			
268 Cer_L_9	-4.66	-51.67	-46.09	0.619		
269 Cer_V_9	1.90	-53.70	-35.76	0.646		
270 Cer_R_9	8.60	-51.75	-46.65			
271 Cer_L_9	-19.09	-34.44	-43.32	0.642		
272 Cer_V_9	2.58	-46.06	-33.23			
273 Cer_R_9	23.85	-35.06	-43.58			0.591

Supplementary Table 2: Mean accuracies for global classification at different levels of motion matching

Training sample=20/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.
1	r=0.657, p=0.0000*	r=0.626, p=0.0050*	r=0.697, p=0.0000*	r=0.616, p=0.0100*
2	r=0.610, p=0.0300*	r=0.650, p=0.0000*	r=0.656, p=0.0000*	r=0.633, p=0.0050*
3	r=0.616, p=0.0200*	r=0.633, p=0.0100*	r=0.663, p=0.0000*	r=0.633, p=0.0100*
4	r=0.621, p=0.0150*	r=0.620, p=0.0150*	r=0.669, p=0.0050*	r=0.615, p=0.0150*
5	r=0.622, p=0.0050*	r=0.622, p=0.0050*	r=0.645, p=0.0000*	r=0.624, p=0.0050*

Training sample=25/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.
1	r=0.675, p=0.0050*	r=0.628, p=0.0100*	r=0.701, p=0.0000*	r=0.622, p=0.0100*
2	r=0.614, p=0.0350*	r=0.663, p=0.0000*	r=0.667, p=0.0000*	r=0.657, p=0.0050*
3	r=0.623, p=0.0250*	r=0.643, p=0.0150*	r=0.683, p=0.0000*	r=0.655, p=0.0050*
4	r=0.633, p=0.0050*	r=0.633, p=0.0050*	r=0.672, p=0.0000*	r=0.631, p=0.0050*
5	r=0.631, p=0.0350*	r=0.636, p=0.0150*	r=0.639, p=0.0150*	r=0.638, p=0.0150*

Training sample=30/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.
1	r=0.685, p=0.0000*	r=0.620, p=0.0050*	r=0.699, p=0.0000*	r=0.614, p=0.0050*
2	r=0.615, p=0.0350*	r=0.684, p=0.0000*	r=0.667, p=0.0000*	r=0.679, p=0.0000*
3	r=0.640, p=0.0050*	r=0.657, p=0.0050*	r=0.698, p=0.0000*	r=0.666, p=0.0000*
4	r=0.641, p=0.0100*	r=0.638, p=0.0100*	r=0.680, p=0.0050*	r=0.641, p=0.0100*
5	r=0.646, p=0.0200*	r=0.648, p=0.0200*	r=0.651, p=0.0100*	r=0.649, p=0.0100*

Training sample=35/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.
1	r=0.688, p=0.0000*	r=0.614, p=0.0300*	r=0.704, p=0.0000*	r=0.603, p=0.0450*
2	r=0.625, p=0.0150*	r=0.700, p=0.0050*	r=0.662, p=0.0050*	r=0.695, p=0.0050*
3	r=0.645, p=0.0100*	r=0.660, p=0.0050*	r=0.699, p=0.0000*	r=0.670, p=0.0000*
4	r=0.654, p=0.0000*	r=0.639, p=0.0100*	r=0.683, p=0.0000*	r=0.645, p=0.0100*
5	r=0.656, p=0.0150*	r=0.653, p=0.0150*	r=0.646, p=0.0150*	r=0.658, p=0.0150*

Supplementary table 3: Motion vs. SVM score correlations at different levels of motion matching

Training sample=20/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.	Null
1	r=-0.224, p=0.1651	r=-0.230, p=0.1532	r=-0.113, p=0.4887	r=-0.209, p=0.1967	r=0.016, p=0.9230
2	r=-0.072, p=0.6573	r=-0.139, p=0.3939	r=0.002, p=0.9879	r=-0.090, p=0.5827	r=-0.002, p=0.9909
3	r=-0.021, p=0.8990	r=-0.085, p=0.6035	r=0.053, p=0.7441	r=-0.079, p=0.6287	r=-0.006, p=0.9720
4	r=0.078, p=0.6311	r=0.014, p=0.9338	r=0.119, p=0.4641	r=-0.033, p=0.8382	r=0.004, p=0.9814
5	r=0.002, p=0.9906	r=-0.027, p=0.8683	r=0.174, p=0.2839	r=-0.073, p=0.6538	r=0.012, p=0.9403

Training sample=25/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.	Null
1	r=-0.223, p=0.1671	r=-0.254, p=0.1131	r=-0.123, p=0.4493	r=-0.240, p=0.1361	r=-0.020, p=0.9048
2	r=-0.092, p=0.5713	r=-0.166, p=0.3066	r=0.015, p=0.9273	r=-0.107, p=0.5093	r=-0.005, p=0.9755
3	r=-0.028, p=0.8635	r=-0.099, p=0.5436	r=0.054, p=0.7405	r=-0.088, p=0.5888	r=0.009, p=0.9546
4	r=0.057, p=0.7268	r=0.020, p=0.9019	r=0.157, p=0.3323	r=-0.032, p=0.8443	r=0.008, p=0.9613
5	r=0.001, p=0.9953	r=-0.045, p=0.7812	r=0.178, p=0.2716	r=-0.089, p=0.5844	r=-0.007, p=0.9641

Training sample=30/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.	Null
1	r=-0.241, p=0.1339	r=-0.259, p=0.1071	r=-0.112, p=0.4909	r=-0.254, p=0.1136	r=0.001, p=0.9971
2	r=-0.098, p=0.5477	r=-0.175, p=0.2811	r=0.013, p=0.9377	r=-0.102, p=0.5315	r=0.015, p=0.9244
3	r=-0.036, p=0.8276	r=-0.117, p=0.4716	r=0.062, p=0.7022	r=-0.090, p=0.5818	r=-0.001, p=0.9968
4	r=0.070, p=0.6686	r=0.012, p=0.9402	r=0.141, p=0.3859	r=-0.044, p=0.7894	r=-0.000, p=0.9980
5	r=0.008, p=0.9594	r=-0.044, p=0.7896	r=0.191, p=0.2373	r=-0.106, p=0.5145	r=-0.014, p=0.9311

Training sample=35/group

Mot. levels	Inner	Outer	Inner reg.	Outer reg.	Null
1	r=-0.255, p=0.1119	r=-0.284, p=0.0759	r=-0.122, p=0.4536	r=-0.276, p=0.0845	r=-0.004, p=0.9789
2	r=-0.107, p=0.5112	r=-0.197, p=0.2239	r=0.028, p=0.8626	r=-0.103, p=0.5270	r=-0.003, p=0.9837
3	r=-0.028, p=0.8655	r=-0.119, p=0.4659	r=0.085, p=0.6012	r=-0.076, p=0.6427	r=0.012, p=0.9410
4	r=0.074, p=0.6483	r=0.024, p=0.8840	r=0.156, p=0.3356	r=-0.024, p=0.8851	r=-0.010, p=0.9532
5	r=0.001, p=0.9949	r=-0.063, p=0.6984	r=0.200, p=0.2149	r=-0.138, p=0.3956	r=0.019, p=0.9064