Online supplement for Maron-Katz et al., Individual Patterns of Abnormality in Resting-State Functional Connectivity Reveal Two Data-Driven PTSD Subgroups. Am J Psychiatry (doi: 10.1176/appi.ajp.2019.19010060)

Supplementary Methods

Neurocognitive test battery

Neurocognitive functioning was tested using a standardized computer-based neurocognitive test battery which assesses multiple cognitive domains, including memory, attention, control, processing speed, executive functions and emotion processing**.** A detailed description of the tests and tasks can be found elsewhere [1]. Importantly, impairments in some of these domains are common in PTSD patients [2]. The specific cognitive measures extracted and used for analysis are specified in supplemental Table S3.

MRI data acquisition and preprocessing

General scan parameters.

All resting-state fMRI scans were eight minutes in length. Participants were scanned either on the Stanford GE 750 scanner or at NYU on a Siemens 3T Skyra scanner. Both sites acquired 32 axial slices with 3.5mm thickness using an echo-planar gradient-echo T2-weighted pulse sequence (repetition time, 2000ms; echo time, 29 ms; flip angle, 90 degrees; slice spacing, 0; field of view, 20cm; matrix size, 64x64). A high-resolution T1-weighted structural scan was acquired as follows: three-dimensional MPRAGE in the sagittal place with the following parameters: inversion time = 450 ms, TR = 8.21 ms, TE = 3.22 ms, flip angle = 15°, field of view = 24 cm, 184 slices, matrix = 256x256, acquired resolution = 0.9375 x 0.9375 x 1.0 mm. The quality of fMRI scans was monitored by MRI center staff weekly with scans of a Functional Biomedical Informatics Research Network (fBIRN) agar phantom, as previously described [3] [4].

The following text is reproduced from [5] . As this was a two-site study, prior to study initiation, we harmonized image acquisition sequences across the two scanners. This involved both assessing image quality and signal to noise ratio (SNR) of images acquired using different parameters at each site, as well as acquisition of the same sequences on several traveling non-study control participants. Scanning of traveling non-study controls was repeated at roughly mid-study. Though the same acquisition sequences were used at both sites, differences between scanners are expected. For example, the echo-planar resting-state scans typically had more ventral prefrontal and temporal lobe susceptibility-artifact dropout at the Stanford site than the NYU site. These differences in acquisition site were accounted for using a site variable in all statistical models.

During the progress of the study, assessment of signal quality and stability was done as follows. For each scan, quality was assessed by quantitative and qualitative factors, with results regularly reported to MRI center staff and principle investigators at weekly meetings. Quantitative factors include: scan parameters (check for correctness), slice-based signal to noise ratio, and total root mean square head motion as well as framewise displacement. We also monitored scanner performance by tracking reference voltage, imaging frequency, and bias field correction over time. Reference voltage (aka RF transmit reference voltage) determines the amplitude of the RF pulses. Imaging frequency variations can indicate scanner problems. The scanner's central frequency is typically set to the resonance frequency of water photons. Measurements of this are proportional to the field strength, and imaging frequency is a common calibration parameter and can be found in every image's DICOM header. Variations that exceed reference values can indicate magnet drift or RF instability. Qualitative factors were assessed visually by a trained image quality assessor. These included field of view clipping, wrapping, dropout, ringing/ striping, blurring, ghosting, RF problems (noise, spikes, leakage), and inhomogeneity.

fMRI preprocessing

The first 5 acquired volumes (10 seconds) were dropped, the data were then motion corrected using

FSL's mcFLIRT (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MCFLIRT). The non-linear registration to standard space was performed using FSL's FNIRT, registration from functional to T1-weighted structural images was estimated using FSL's implementation of boundary-based registration (fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT_BBR). The functional time series was residualized with respect to six motion parameters (estimated using mcFLIRT) and with respect to the mean white matter (WM) and cerebral spinal fluid (CSF) signal that was estimated using an MNI-space defined WM/CSF mask transformed to the native functional space. The data were then spatially smoothed with full-width half-maximum (FWHM) Gaussian of 6mm, consistent with our prior work and how our connectivity atlas was previously used [6]. Next, ICA AROMA was applied on the data in its non-aggressive mode and volumes with framewise-displacement of >0.5 mm were removed from the data to ensure the removal of any remaining motion artifacts. Lastly, a bandpass filter was applied to data using cut-off frequencies of 0.008Hz – 0.1Hz. Only subjects with maximum root mean square motion <2mm and fewer than 20 scrubbed volumes were included in analyses. The resulting NIFTIs were transformed to the MNI standard space for downstream analysis.

Parcel-level time series extraction from fMRI data

The mean time series was extracted for 133 brain regions based on a recently-published cortical parcellation into 100 parcels derived from an independent resting-state fMRI cohort [7] combined with a set of 33 subcortical parcels that include 14 cerebellar parcels [8], 13 striatal parcels [9] right and left Amygdala, right and left hippocampus [10] and right and left thalamus [11]. The mean time series for each parcel was estimated by calculating a weighted average of the BOLD signals of all voxels in a parcel. To control for local differences in signal quality on a per subject basis, and further improve signal quality, only voxels for which the signal to noise ratio (SNR) was higher than 100 were used as suggested in [12].

SNR was estimated by dividing the mean BOLD across time with its standard deviation. Weighted average was calculated using equation S1 below. Next, parcel-level BOLD signals were centered and scaled using the z-scoring procedure described in [13].

Equation S1: Weighted average BOLD signal $Aw = \left(\frac{\sum_{i=1}^{n} Vi * Wi}{\sum_{i=1}^{n} Wi}\right)$

Where Vi is the BOLD signal of voxel i, and Wi is the weight of the i-th voxel, which was defined as the maximum of 0 and the Pearson correlation coefficient between Vi and the unweighted average BOLD signal across all the voxels in the parcel.

EEG data acquisition, preprocessing and source localization

EEG recordings were acquired with a BrainAmp DC amplifier (sampling rate: 5 kHz; measurement range: ± 16.384 mv; cut-off frequencies of the analog high-pass and low-pass filters: 0 and 1 kHz) and the Easy EEG cap with 64 extra-flat, freely rotatable, sintered Ag-AgCl electrodes (Brain Products GmbH, Germany). The electrode montage followed an equidistant arrangement extending from below the cheekbone back to below the inion. Electrode impedances were kept below 5 kΩ. An electrode attached to the tip of the nose was used as the reference. Participants were seated on a comfortable reclining chair and were instructed to remain awake and let their mind wander in the 3-minute eyes closed paradigm and fixate a given point for 3 minutes in the eyes open paradigm. Recordings were immediately assessed for quality using a custom MATLAB script and rerun if necessary.

Offline preprocessing of recordings was conducted using a custom MATLAB pipeline calling selected functions from the EEGLAB toolbox. [14] A detailed step-by-step description of our preprocessing pipeline can be found in [15]. However, in more general terms, preprocessing proceeded as follows: EEG data were downsampled to 250 Hz, notch filtered to remove 60 Hz line noise and its harmonic, and bandpass filtered between 1 Hz and 50 Hz using a zero-phase finite impulse response filter. Bad channel

(mean 0.9, SD 2.3 rejected channels) and paroxysmal segment (mean 96.5, SD 6.8, percent time points retained) rejections were performed in a semi-automatic manner, assessing extreme or flatline amplitudes, channel correlations, and robust paroxysmal spike measures with custom pipelines and EEG expert validation. Rejected channels were spherically interpolated and paroxysmal segments excised. To remove remaining artifacts including ocular, heartbeat, and high-frequency persistent muscle artifact, the extended information maximization (Infomax) algorithm, widely used for independent components analysis (ICA), was then performed on the data reduced in dimensionality by principal components analysis (PCA). The number of the dimensions was determined to be the least number of principal components that account for more than 99.9% of the total variance. Artefactual independent components were likewise identified by combination of custom pipelines identifying extreme skew in component topographical weights, stereotypical power spectral density profiles, deviant robust statistics, etc., confirmed by EEG expert, and then rejected.

Source localization to the cortical surface was performed using the Brainstorm toolbox [16]. A three-layer symmetric boundary element model of the head was computed with the OpenMEEG [17] plugin based on a Montreal Neurological Institute (MNI) brain template. Rotating dipoles at 5003 vertices were generated on the cortical surface. The lead-field matrix was obtained by projecting the standard electrode positions onto the scalp. For each subject, an imaging kernel that maps from the channel space EEG to the source space current density was then estimated by the minimum norm estimation approach with depth weighting and regularization. In the method validation study, current densities obtained from native and MNI template head models were compared and MNI template head models were found to be sufficient. For each vertex, the current density time series were reduced from their 3 orthogonal axes to a single principal direction by PCA, then bandpass filtered into canonical frequency bands associated with different neurophysiological processes: theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (31-50 Hz). The current density time series at each canonical frequency band were

Hilbert transformed to yield source space analytical time series. Welch's periodogram was calculated in source space to generate the power spectral density for comparison with the orthogonalization method. In order to calculate resting EEG power envelope connectivity, the analytical time series at each vertex of source space was orthogonalized with respect to all other vertices and power envelopes were then calculated from each of the orthogonalized analytical time series [15]. Pearson's correlation coefficient between the power envelopes was calculated for each vertex pair. The connectivity matrix was next averaged with its transpose and corrected for underestimation inherent to orthogonalization [18].

EEG power envelope connectivity feature extraction

Resting EEG power envelope connectivity was calculated among 31 cortical regions that were previously defined and used for this type of analysis [15] [19]. For each pair of regions, connectivity was calculated by averaging over all corresponding vertex pairs. As a result, 465 unique region-region connectivity values were computed in each of the 4 frequency bands, in each of two resting paradigms (eyes open and closed).

Using tolerance intervals to identify abnormal resting state fMRI connectivity features in PTSD cases

We used a tolerance interval (TI)-based method to examine neurobiological heterogeneity within PTSD patients. A tolerance interval is a statistical interval designed to measure the proportion of the population expected to be in a given interval with confidence. TIs have been used for mapping individual abnormalities with respect to a reference population [20] [21] [22] [23] [24]. One typically constructs confidence intervals for the sample average. Analogously, tolerance intervals are like confidence intervals but for the estimates of percentage of the population instead of the sample average. Importantly, tolerance intervals take the inherent sampling variability in the sample mean and sample standard deviation of a distribution into account when making statements of the percentage of

the population. In contrast, a confidence interval for the sample mean only takes the uncertainty in the sample mean into account.

We used the "tolerance" R package (https://cran.r-project.org/web/packages/tolerance/tolerance.pdf) for non-parametric estimation of 2-sided TIs for each node-to-network connectivity measure to obtain a range within which 90% (proportion=0.9) of the TEHC population is expected to fall with a 10% confidence level (alpha=0.1). Parameters were chosen while accounting for the required sample size using the ssnorm function provided in the tolerance package to best suit the number of subjects.

CLICK clustering analysis

The CLICK algorithm does not make any prior assumptions on the number or the structure of the clusters. It treats each sample in the data as a node in a weighted graph and uses a recursive minimum - cut approach to partition the graph into components. Probabilistic meaning is assigned to edge weights and the stopping criterion, which gives the algorithm high accuracy in identifying tight groups of elements that are likely to belong to the same true cluster, while allowing non-clustered elements (singletons) [25]. We used the existing CLICK implementation available in http://acgt.cs.tau.ac.il/expander/expander.html, with Pearson correlation coefficient as a measure of similarity between each two subjects.

Cluster stability estimation

Cluster stability was estimated based on the consistency of cluster co-occurrence for each pair of patients across the 500 subsampling iterations in the following manner: for each clustering solution, a binary co-occurrence matrix was generated, which for every pair of patients, contains the value 1 if they were assigned to the same cluster and 0 if not. Next, a stability matrix (S) of size nXn (n being the

number of clustered samples, i.e. patients) was generated by averaging these pairwise co-occurrence matrix across subsamples, and then summarized into a single stability index using equation S2 below.

Equation S2: Stability index$= \frac{4}{n*n-1} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} |Sij * 0.5|$

To estimate the significance of the resulting stability index, we applied the entire process of features selection and clustering on a set of 100 random permutations of the original case/control labels. This allowed us to generate a null distribution of stability index values. Due to runtime limitation we settled for 100 random case-control label permutations and in each permutation, we used 100 subsamples of 90% of the data instead of 500 that were used in the above process. For each permutation we recorded the average stability of cluster assignment across subsamples.

Statistical analysis

In order to control for false discovery rate when examining clinical and cognitive differences between clusters we used the Benjamini-Hochberg false discovery rate estimation procedure [26].

Significance of classification accuracy was assessed using 1000 random permutation of cluster labels.

Examining feature and cluster utility in case/control classification

We used the k-nearest neighbors algorithm (k-NN) with k=5 to train a classifier on our set of node-net rsfMRI functional connectivity within the same data, to distinguish between TEHCs and PTSD patients. K-NN was selected due to its simplicity, intuitiveness as well as known robustness to noisy training data and competitive performance in many domains [27]. K was set to 5, but a range of k=3-9 was tested and found to yield similar performance. The classification performance was assessed using a 10X10-fold-cross-validation analysis. In addition, in order to account for uneven cluster sizes, we used a

bootstrapping (random sampling with replacement) up-sampling technique to increase the size of the smaller class. This up-sampling procedure was done only on the training set to avoid biasing the validation sets or having samples of training data in the validation sets. This procedure was repeated for each cluster separately to test whether treating each cluster as a separate class/label would improve classification performance. Accuracy significance was assessed using 1000 random label permutations.

Cluster Validation using rsEEG-based classification

EEG power envelope connectivity features were calculated at each of the four frequency bands for 67 TEHCs and 60 PTSD cases (35 from cluster1, 20 from cluster2 and 5 unclustered) for whom both good quality rsfMRI and rsEEG data were collected. These data were compared between the two rsfMRI-identified clusters using machine learning, as they represent a single, but multi-dimensional, data modality (unlike the clinical and behavioral measures that were tested). For machine learning, we used K-NN with k=5 to train a classifier with simultaneous feature selection to distinguish between the two clusters. K-NN was selected due to its simplicity, intuitiveness as well as known robustness to noisy training data and competitive performance in many domains [27]. The classification performance was assessed using 10X10-fold-cross-validation. In addition, in order to account for uneven cluster sizes, we used bootstrapping (random sampling with replacement) to increase the size of the smaller cluster. This up-sampling procedure was done only on the training set to avoid biasing validation sets or having samples of training data in the validation sets. This procedure was repeated with each of the 2 paradigms and each of the 4 frequency bands. Accuracy significance was assessed using 1000 random label permutations.

Supplementary Results

Utility of feature and clusters in case/control classification

We used K-NN to train a classifier on our set of node-net rsfMRI functional connectivity within the same data, to distinguish between TEHCs and PTSD patients. Classification performance was assessed using 10X10-fold-cross-validation. This procedure was repeated for each cluster separately to test whether treating each cluster as a separate class/label would improve classification performance. Performance significance was assessed using 1000 random label permutations. Results are shown in Figure S4. The identified subgroups improve case/control discriminability leading to significant accuracy using both TI-based features and independently selected features. However, while using independent feature selection leads to significant sensitivity and specificity, using the set of consistently selected TI-based features lead to very low (below chance level) sensitivity but superior (highly significant) specificity. Notably, classification accuracy does not go above chance level when treating all PTSD patients as one group regardless of the features used.

Classification of patients into clusters using resting state EEG data

To further validate that the identified clusters indeed reflect PTSD subtypes with distinct neurobiological correlates, we trained a k-NN model to distinguish between the two clusters using resting state EEG power envelope connectivity features that were available for 60 of the patients (35 from cluster1, 20 from cluster2 and 5 unclustered) and for 67 of the TEHCs. Figure S5 shows the results of the classification analysis. Our classifier was able to distinguish between the two patient clusters with an accuracy of 68% ($p<0.005$), with a sensitivity of 73% in detecting cluster2 and 62% in detecting cluster1 (Figure S5a). This was done using features extracted from the eyes open beta frequency condition. The connectivity features that contributed most to the classification are shown in Figure S5b

and involve regions of the frontoparietal and visual networks as well as regions of the default mode and ventral attention networks.

Supplementary Tables and Figures

Table S1A. Demographic Characteristics of PTSD and TEHC Groups

Abbreviations: CAPS, Clinician Administered PTSD Scale for DSM-5; BDI, Beck Depression Inventory

Characteristic	TEHCs (n = 105)	PTSD (n = 87)	Stat value	P value
Age, mean (SD), y	32 (7.7)	32.7 (6.9)	0.66 (t)	0.51
Males, No. (%)	93 (88.6)	76 (87.4)	0.07 (χ^2)	0.8
Educational attainment, mean (SD), y	15.5 (2.1)	15.3 (2.1)	-0.58 (t)	0.56
Site distribution, No. Stanford (%)	34 (32.4)	23 (26.4)	0.8 (χ^2)	0.37
Clinical Assessments, mean (SD)				
CAPS total score	3 (4.1)	24.6 (10.6)	16.8 (t)	< .00001
CAPS subscale B score	0.5 (1.1)	5.9 (3.4)	13.6 (t)	< .00001
CAPS subscale C score	0.2 (0.9)	3.2 (1.9)	12.9 (t)	< .00001
CAPS subscale D score	0.6 (1.5)	7.6 (5.2)	11.3 (t)	< .00001
CAPS subscale E score	1.7 (2.4)	7.9 (3.6)	12.5 (t)	< .00001
BDI total score	3.9 (6.1)	17.3 (11)	9.1 (t)	< .00001
Comorbidity, No. subjects (%)				
Major Depressive Disorder	0 (0)	24 (27.6)		
Traumatic Brain Injury	0 (0)	32 (36.8)		
Both	0 (0)	22 (25.3)		

Table S1B. Demographic and clinical characteristics of the two sites

Characteristic	Stanford (n = 57)	NYU (n = 135)	Stat value	P value
PTSD cases, No. (%)	23 (40.4)	64 (34)		
Age, mean (SD), y	32.1 (7.3)	32.4 (7.4)	-0.27 (t)	0.79
Males, No. (%)	51 (89.5)	118 (87.4)	0.16 (χ^2)	0.69
Educational attainment, mean (SD), y	15.1 (1.8)	15.5 (2.2)	-1.19 (t)	0.24
Clinical Assessments within cases,				

mean (SD)				
CAPS total score	29 (10.3)	23.2 (10.3)	-2.02 (z)	0.04
CAPS subscale B score	6.4 (4.1)	5.8 (3.1)	-0.24 (z)	0.81
CAPS subscale C score	3.65 (2)	3.09 (1.8)	-1.49 (z)	0.14
CAPS subscale D score	9.4 (4.7)	7 (5.3)	-1.9 (z)	0.052
CAPS subscale E score	9.6 (3.6)	7.3 (3.5)	-2.1 (z)	0.033
BDI total score	24.5 (7.5)	15.4 (11.1)	-2.99 (z)	0.003
Comorbidity within cases, No. subjects (%)				
Major Depressive Disorder	11 (19.3)	13 (9.6)	6.4 (χ^2)	0.01
Traumatic Brain Injury	8 (14)	24 (17.8)	0.05 (χ^2)	0.82
Both	6 (10.5)	8 (5.9)	2.3 (χ^2)	0.13

Table S2. Demographic Characteristics of PTSD clusters

Abbreviations: CAPS, Clinician Administered PTSD Scale for DSM-5; BDI, Beck Depression Inventory

Characteristic	**Cluster 1 (n = 49)**	**Cluster 2 (n = 30)**	**Stat value**	**P value**
Age, mean (SD), y	32.9 (7.4)	32.2 (6.9)	0.41 (t)	0.68
Males, No. (%)	41 (83.7)	28 (93.3)	1.57(χ^2)	0.21
Educational attainment mean (SD), y	15.2 (2.2)	15.4 (2.2)	-0.46(t)	0.64
Site distribution, No. Stanford (%)	15 (30.6)	6 (20)	1.07(χ^2)	0.3
Subject level motion (max abs displacement), mean (SD)	0.59 (0.32)	0.63 (0.45)	0.22 (z)	0.83
Clinical Assessments, mean (SD)				
CAPS total score	22 (8.4)	28.5 (11.9)	-2.17 (z)	0.03
CAPS subscale B score	4.8 (2.4)	7.6 (3.9)	-3.4 (z)	< .001
CAPS subscale C score	3 (1.8)	3.7 (1.8)	-1.5 (z)	0.14
CAPS subscale D score	6.8 (4.4)	8.7 (5.7)	-1.5 (z)	0. 14
CAPS subscale E score	7.4 (3.5)	85 (3.9)	-0.94(z)	0.35
BDI total score	16.1 (9.6)	19 (12.2)	-0.72 (z)	0.47
Comorbidity, No. subjects (%)				
Major Depressive Disorder	12 (24.5)	12 (40)	2.1 (χ^2)	0.15
Traumatic Brain Injury	14 (28.6)	17 (56.7)	6.16 (χ^2)	0.01
Both	7(14.3)	7 (23.3)	1.04 (χ^2)	0.3

Table S3: Clinical and cognitive measures:

Domain	Test	measures derived
Verbal Memory	Word list learning task	Total learning score, delayed recognition score
Sustained Attention	Continuous Performance Task	Percentage of correct responses, mean reaction time to correct responses
Working Memory	Digit Span forward	Maximum length of digits correctly entered in a correct order, number of correct trials
Inhibitory Control	Color-word Stroop interference task	Difference between mean reaction time when responding to font color and responding to color word (for correct responses only)
Response Inhibition	Go/NoGo task	Percentage of correct responses, mean reaction time to "go" trials.
Processing Speed	Choice reaction time task	Average time taken to press the left or right arrow to correspond with which circle (left or right) is highlighted in each trial
Response Speed	Motor Tapping	Number of times subject was able to tap the spacebar in 30 seconds with their dominant hand
Information processing efficiency (switching)	Trail making test	Difference between total time to complete a whole sequence of digits and letters to the whole time to complete sequence of digits.
Executive functions	Maze	Total time to complete the maze correctly, number of trials taken to complete the maze correctly, number of errors (total, overruns)
Emotion Identification	Emotion recognition	Interference scores computed by subtracting the mean accuracy/mean reaction time for identifying neutral faces from the mean accuracy/reaction time to identify emotional faces.
Emotion bias	Delayed emotion recognition	Interference scores computed by subtracting the mean accuracy/mean reaction time for recognition of neutral faces from the mean accuracy/reaction time for recognition of emotional faces.
PTSD	Clinician administered PTSD questionnaire for DSM 5 past month (CAPS-5)	Structured clinical interview. Sum of scores for each cluster: Criterion B: Re-experiencing; Criterion C: Avoidance; Criterion D: Negative alterations in cognitions and mood; Criterion E: Alterations in arousal and reactivity; Past moth total score.
Quality of Life	WHO Quality of Life Brief questionnaire (WHO-QOL-BREF) [28]	Self-report questionnaire with 26 items. Scores are derived for 4 subscales: physical, psychological, social, environmental and a measure of overall quality of life.
Depression	Beck Depression Inventory – II (BDI-II) [29]	Self-report questionnaire with 21 items added up to one total score.
Trauma	Early Trauma Inventory– Self Report (ETI-SR) [30]	Self-report questionnaire with 62 items. Scores derived for 4 domains of childhood traumatic events: general trauma, physical, emotional, and sexual

		abuse, as well as a total score.
Affect	Positive and Negative Affect Schedule (PANAS) [31]	20 item self-report questionnaire deriving two subscale scores: positive affect (PA) and negative affect (NA)

Table S4: Brain regions used for resting EEG functional connectivity feature extractions (shown in Figure S5b)

Region acronym	Region	Associated brain network
LV1, RV1	Left and right primary visual cortex	Visual
LSMC, RSMC	Left and right sensorimotor cortex	Sensorimotor
LIFG, RIFG	Left and right inferior frontal gyrus	Dorsal attention
LIPS, RIPS	Left and right intraparietal sulcus	Dorsal attention
LFEF, RFEF	Left and right frontal eye fields	Dorsal attention
LSEF, RSEF	Left and right supplementary eye fields	Dorsal attention
PCC	Posterior cingulate cortex	Default mode
MPFC	Medial prefrontal cortex	Default mode
LANG, RANG	Left and right angular gyrus	Default mode
LPMFG, RPMFG	Left and right posterior middle frontal gyrus	Frontoparietal
LIPL, RIPL	Left and right inferior parietal lobule	Frontoparietal
LORB, RORB	Left and right orbital cortex	Frontoparietal
LMTG, RMTG	Left and right middle temporal gyrus	Frontoparietal
LAMFG, RAMFG	Left and right posterior middle frontal gyrus	Ventral attention
LINS, RINS	Left and right insula	Ventral attention
DACC	Dorsal anterior cingulate cortex	Ventral attention
LSUP, RSUP	Left and right supramarginal gyrus	Ventral attention

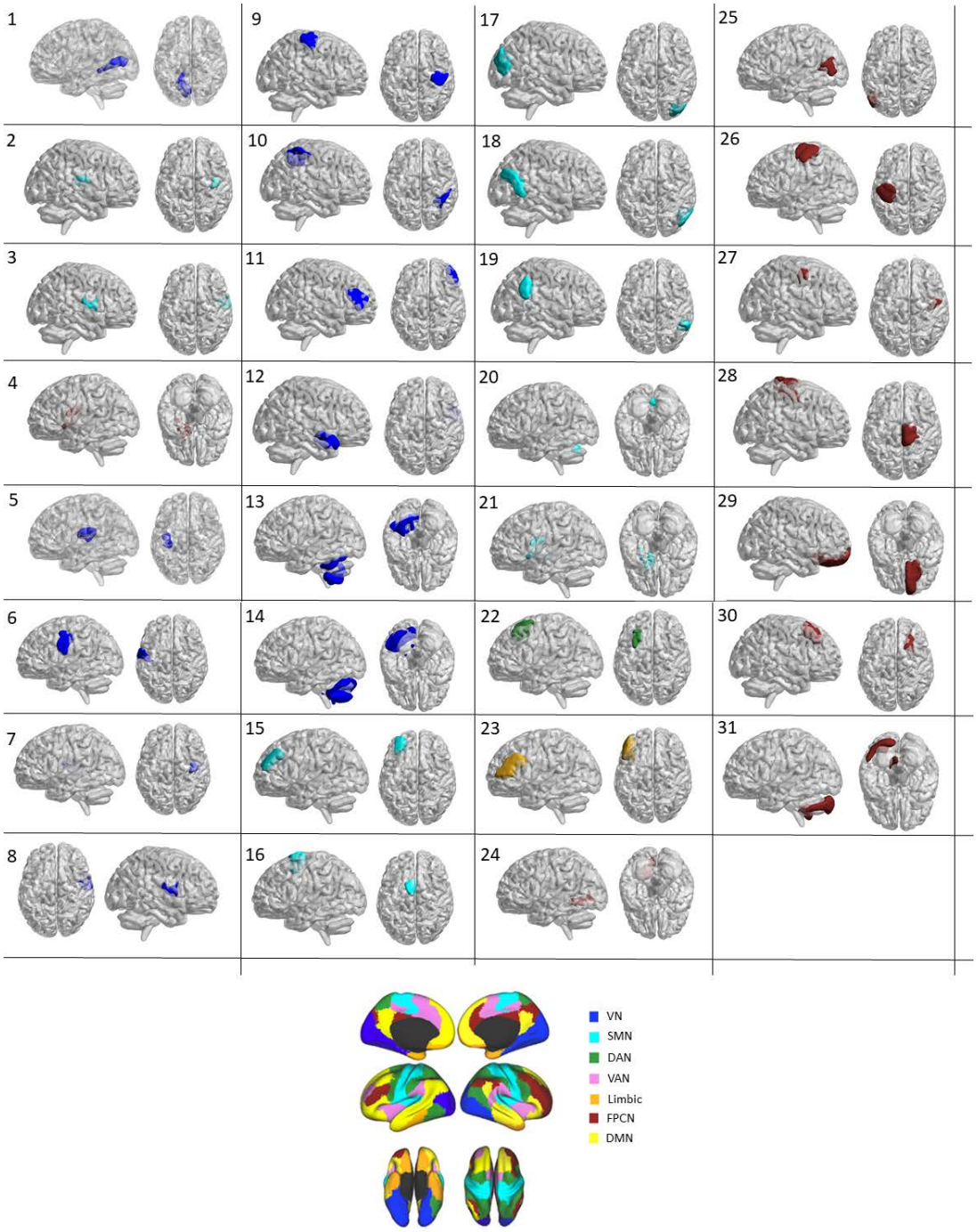

Figure S1: a visualization of all 31 node-network functional connectivity features that were consistently selected in over 70% of the subsamples. For each feature, the node is shown as overplay on the standard brain, while the network is indicated by the overlay color. Network colors are indicated in the network maps at the bottom of the figure. Visualization was created using the BrainNet viewer [32].

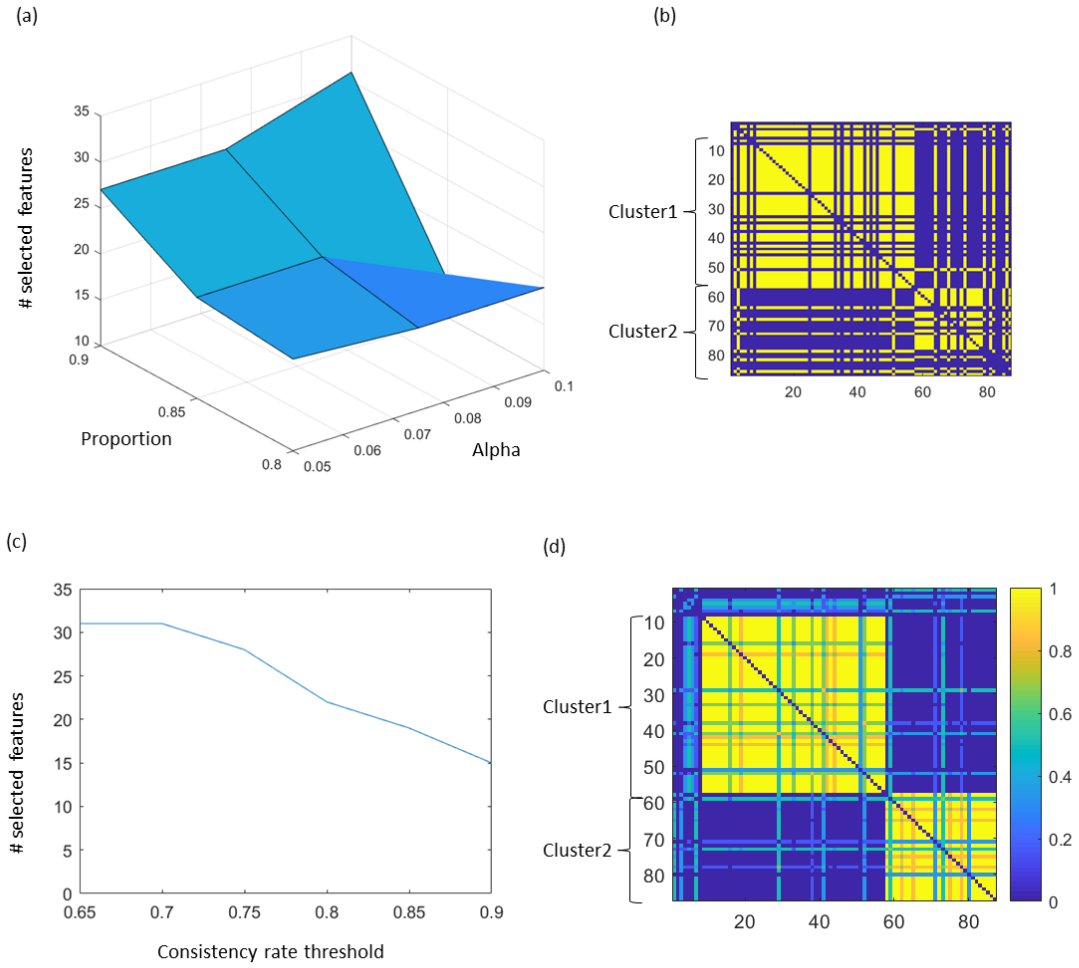

Figure S2: Results of CLICK clustering over features selected across a range of TI and consistency rate parameters. (a) The number of selected features as a function of alpha values and proportion values. (b) The co-occurrence of each pair of subjects in the cross-TI parameter clustering solution (1=same cluster (yellow), 0=not same cluster (blue)). The matrix is organized based on the order of the clustering solution reported in the manuscript. This matrix reflects the consistency between the two solutions. (c) The effect of used consistency rate threshold on the number of selected features (for TIs generated with alpha=0.1, proportion=90%) (d) The mean co-occurrence of each pair of subjects across clustering solution generated with different consistency rate thresholds (65-90%, in increments of 5%). The matrix is organized based on the order of the clustering solution reported in the manuscript.

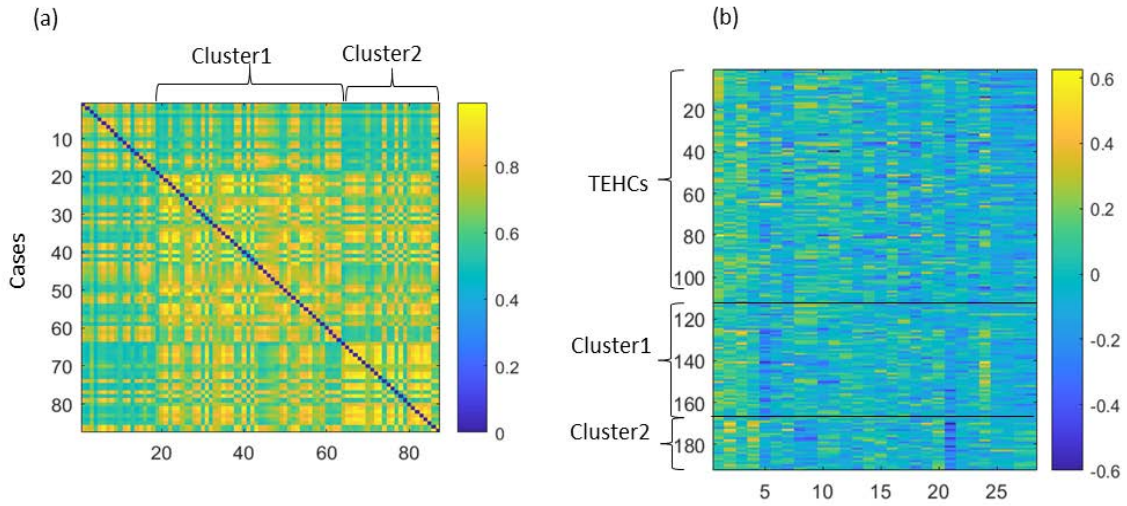

Figure S3: Identified clusters using the set of 28 connectivity features that were consistently selected using a t-test in > 70% of the 500 runs. The frequency at which each pair of patients were assigned to the same cluster across the 500 runs is shown in (a) and reflects cluster stability (matrix rows and columns are patients ordered according to the final clustering solution). Two case clusters were identified. Connectivity values for the 28 features organized by cluster order are shown in (d).

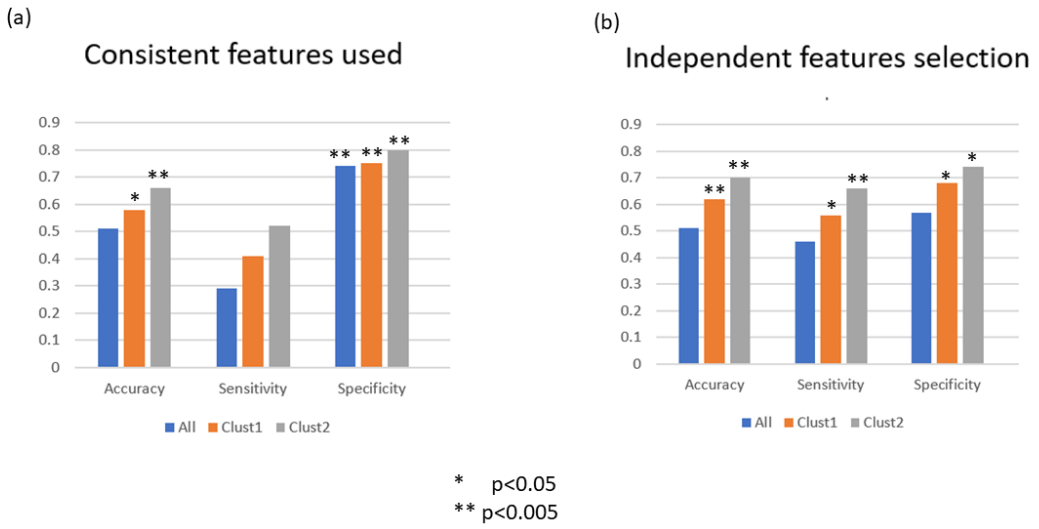

Figure S4: Case/control classification performance when training a k-NN model on the set of node-net fMRI connectivity features. Performance was estimated using a 10X10 cross validation framework and significance was assessed using 1000 random label permutations. (a)Classification performance using the set of 31 consistently selected features for controls vs. all cases (blue), controls vs. cluster1 (orange) and controls vs. cluster2 (gray). (b) Classification performance using independent within-fold feature selection (taking the 31 most distinctive features in each fold) for controls vs. all cases (blue), controls vs. cluster1 (orange) and controls vs. cluster2 (gray).

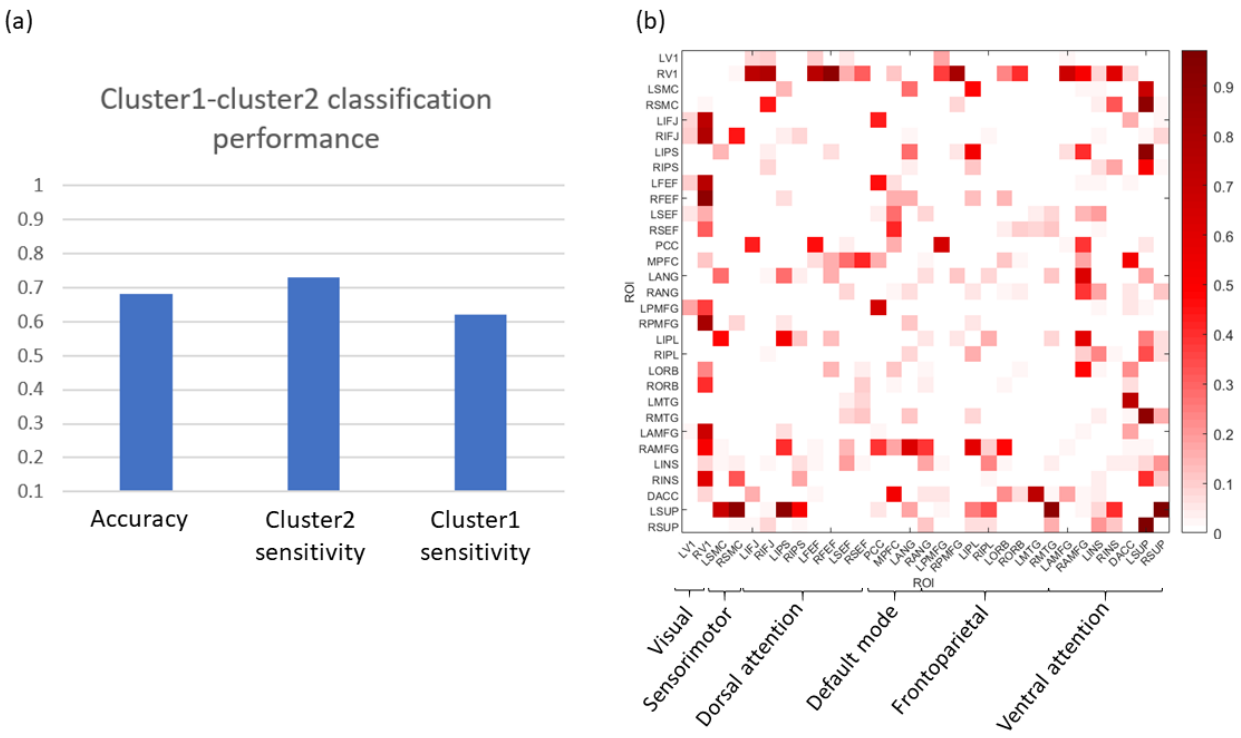

Figure S5: Distinction between the two rsfMRI abnormality-defined PTSD subgroups using resting-state EEG connectivity analyses and machine learning: Our classifier was able to distinguish between the two patient clusters with an accuracy of 68% ($p<0.005$), with a sensitivity of 73% in detecting cluster2 and 62% in detecting cluster1 (a). Feature selection frequencies are shown in (b) where values indicate the frequency at which every feature selection across the 10X10 models. Discriminative (most frequently selected) EEG features involved regions of the visual and ventral attention networks.

References

[1] S. M. Silverstein, S. Berten, P. Olson, R. Paul, L. M. Williams, N. Cooper and E. Gordon, "Development and validation of a World-Wide-Web-based neurocognitive assessment battery: WebNeuro," *Behavior Research Methods,* vol. 39, no. 4, pp. 940-949, 2007.

[2] J. C. Scott, G. E. Matt, K. M. Wrocklage, C. Crnich, J. Jordan, S. M. Southwick, J. H. Krystal and B. C. Schweinsburg, "A quantitative meta-analysis of neurocognitive functioning in posttraumatic stress disorder.," *Psychological Bulletin,* vol. 141, no. 1, p. 105, 2015.

[3] L. Friedman and G. H. Glover, "Report on a multicenter fMRI quality assurance protocol," *Journal of Magnetic Resonance Imaging,* vol. 23, no. 6, pp. 827-839, 2006.

[4] G. H. Glover, B. A. Mueller, J. A. Turner, T. G. M. Van Erp, T. T. Liu, D. N. Greve, J. T. Voyvodic, J. Rasmussen, G. G. Brown and D. B. Keator, "Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies," *Journal of Magnetic Resonance Imaging,* vol. 36, no. 1, pp. 39-54, 2012.

[5] A. Etkin, A. Maron-Katz, W. Wu, G. Fonzo, J. Huemer, P. E. Vertes, B. Patenaude, J. Richiardi, M. S. Goodkind, C. Keller, J. Ramos-Cejudo, Y. V. Zaiko, K. K. Peng, E. Shpigel, P. Longwell, R. T. Toll, A. Thompson, S. Zack, B. Gonzalez and R. O'Hara, Connectomics and Cognition Define a Treatment-Resistant Form of Post-Traumatic Stress Disorder, 2019.

[6] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cerebral cortex,* vol. 22, no. 1, pp. 158-165, 2012.

[7] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff and B. T. Yeo, "Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI," *Cerebral Cortex,* pp. 1-20, 2017.

[8] R. L. Buckner, F. M. Krienen, A. Castellanos, J. C. Diaz and B. T. T. Yeo, "The organization of the human cerebellum estimated by intrinsic functional connectivity," *Journal of Neurophysiology,* vol. 106, no. 5, pp. 2322-2345, 27 11 2011.

[9] E. Y. Choi, B. T. T. Yeo and R. L. Buckner, "The organization of the human striatum estimated by intrinsic functional connectivity.," *Journal of neurophysiology,* vol. 108, no. 8, pp. 2242-2263, 10 2012.

[10] A. C. Chen and A. Etkin, "Hippocampal network connectivity and activation differentiates post-traumatic stress disorder from generalized anxiety disorder," *Neuropsychopharmacology,* vol. 38, no. 10, p. 1889, 2013.

[11] T. E. J. Behrens, H. Johansen-Berg, M. W. Woolrich, S. M. Smith, C. A. M. Wheeler-Kingshott, P. A. Boulby, G. J. Barker, E. L. Sillery, K. Sheehan and O. Ciccarelli, "Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging," *Nature neuroscience,* vol. 6, no. 7, p. 750, 2003.

[12] A. T. Drysdale, L. Grosenick, J. Downar, K. Dunlop, F. Mansouri, Y. Meng, R. N. Fetcho, B. Zebley, D. J. Oathes, A. Etkin, A. F. Schatzberg, K. Sudheimer, J. Keller, H. S. Mayberg, F. M. Gunning, G. S. Alexopoulos, M. D. Fox, A. Pascual-Leone, H. U. Voss, B. J. Casey, M. J. Dubin and C. Liston, "Resting-state connectivity biomarkers define neurophysiological subtypes of depression," *Nature Medicine,* vol. 23, p. 28, 5 12 2016.

[13] R. A. Olshen and B. Rajaratnam, "Successive normalization of rectangular arrays," *Annals of Statistics,* vol. 38, no. 3, pp. 1638-1664, 6 2010.

[14] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods,* vol. 134, no. 1, pp. 9-21, 2004.

[15] R. T. Toll, W. Wu, S. Naparstek, M. Narayan, B. Patenaude, C. D. L. Angeles, K. Sarhadi, N. Anicetti, P. Longwell, E. Shpigel, R. Wright, J. Newman, B. Gonzalez, R. Hart, S. Mann, D. Abu-amara, K. Sarhadi, C. Cornelssen and C. Marmar, "An Electroencephalography Connectomic Profile of Post-Traumatic Stress Disorder," *American Journal of Psychiatry,* no. Under review, p. 7, 2019.

[16] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis and R. M. Leahy, "Brainstorm: a user-friendly application for MEG/EEG analysis," *Computational intelligence and neuroscience,* vol. 2011, p. 8, 2011.

[17] A. Gramfort, T. Papadopoulo, E. Olivi and M. Clerc, "OpenMEEG: opensource software for quasistatic bioelectromagnetics," *Biomedical engineering online,* vol. 9, no. 1, p. 45, 2010.

[18] J. F. Hipp, D. J. Hawellek, M. Corbetta, M. Siegel and A. K. Engel, "Large-scale cortical correlation structure of spontaneous oscillatory activity," *Nature neuroscience,* vol. 15, no. 6, p. 884, 2012.

[19] C. Imperatori, B. Farina, M. I. Quintiliani, A. Onofri, P. C. Gattinara, M. Lepore, V. Gnoni, E. Mazzucchi, A. Contardi and G. Della Marca, "Aberrant EEG functional connectivity and EEG power spectra in resting state post-traumatic stress disorder: A sLORETA study," *Biological psychology,* vol. 102, pp. 10-17, 2014.

[20] D. S. Young, "Computing tolerance intervals and regions using R," in *Handbook of Statistics*, vol. 32, Elsevier, 2014, pp. 309-338.

[21] D. S. Young and T. Mathew, "Improved nonparametric tolerance intervals based on interpolated and extrapolated order statistics," *Journal of Nonparametric Statistics,* vol. 26, no. 3, pp. 415-432, 2014.

[22] J. Zhang, W. Li, L. K. Roskos and H. Yang, "Immunogenicity assay cut point determination using nonparametric tolerance limit," *Journal of immunological methods,* vol. 442, pp. 29-34, 2017.

[23] G. A. Carlesimo, C. Caltagirone, G. Gainotti, L. Fadda, R. Gallassi, S. Lorusso, G. Marfia, C. Marra, U. Nocentini and L. Parnetti, "The Mental Deterioration Battery: Normative Data, Diagnostic Reliability and Qualitative Analyses of Cognitive Impairment," *European Neurology,* vol. 36, no. 6, pp. 378-384, 1996.

[24] N. Altman and M. Krzywinski, "Predicting with confidence and tolerance," vol. 15, no. November, pp. 843-845, 2018.

[25] R. Sharan, A. Maron-Katz and R. Shamir, "CLICK and EXPANDER: A system for clustering and visualizing gene expression data," *Bioinformatics,* vol. 19, no. 14, 2003.

[26] "benjamini_hochberg1995".

[27] M. Bolandraftar, S. Bafandeh and I. And, "Application of K-nearest neighbor (KNN) approach for predicting economic events theoretical background Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background".

[28] S. M. Skevington, M. Lotfy and K. A. (. O'Connell, "The World Health Organization's WHOQOL-BREF

quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group," *Quality of life Research,* vol. 13, no. 2, pp. 299-310, 2004.

[29] A. T. Beck, R. A. Steer and G. K. Brown, "Beck depression inventory-II," *San Antonio,* vol. 78, no. 2, pp. 490-498, 1996.

[30] J. D. Bremner, R. Bolus and E. A. Mayer, "Psychometric properties of the early trauma inventory–self report," *The Journal of nervous and mental disease,* vol. 195, no. 3, p. 211, 2007.

[31] D. Watson, L. A. Clark and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the PANAS scales.," *Journal of personality and social psychology,* vol. 54, no. 6, p. 1063, 1988.

[32] M. Xia, J. Wang and Y. He, "BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics," *PLoS ONE,* vol. 8, no. 7, 4 7 2013.