Supporting Information

Perrin et al. 10.1073/pnas.1117206109

SI Text

Functional Connectivity: Data Analysis. The CHART method that we used for the functional connectivity analysis involves a number of processing steps of varying complexity. We first provide an overview of the general workflow before we describe each processing step in detail. As illustrated in Fig. S1, the CHART analysis comprised two stages. During stage 1, the average global connectivity, also known as weighted global connectivity (WGC), was calculated for each subject and each condition (i.e., pre- and post-ECT treatment), and a nonparametric statistical permutation test revealed a restricted cluster of 12 voxels in and around the left DLPFC region, where the average global functional connectivity was considerably decreased after ECT treatment (*Results* and Fig. 1). During stage 2, this significant cluster of voxels was then used as a seed region to determine the related functional brain networks [i.e., the brain areas to which the seed region was functionally connected before and after ECT treatment (*Methods*)], and the two resulting seed correlation maps were saved. Finally, the two seed correlation maps were thresholded using a correlation threshold of 0.5, which corresponded to an FWE-corrected probability of $P < 0.001$. The FWE-corrected probability was calculated in a Monte Carlo simulation from a set of 1,000 surrogate data with matching spatial smoothness and temporal autocorrelation. The corresponding workflow is illustrated in Fig. S2 and is described in detail in a separate section (*Functional Connectivity: Surrogate Data Analysis*). We now describe each processing step shown in Fig. S1 in detail. Unless otherwise stated, all routines were implemented in MATLAB (version 7.13; The MathWorks, Inc.).

Data. The statistical parametric mapping software package (SPM8; www.fil.ion.ucl.ac.uk/spm) was used for data preprocessing, which included realignment, slice timing correction, coregistration, normalization, reslicing to $4 \times 4 \times 4$ mm, and spatial smoothing with an 8-mm FWHM Gaussian isotropic kernel. Voxels that consisted of constant time courses as a result of the SPM8 realignment routine were masked before smoothing, and the corresponding time courses were replaced by Gaussian random noise after smoothing. (A small number of constant time courses typically occur at the surface of the 3D brain volume, because the SPM8 realignment routine automatically replaces time courses that comprise missing data points by an array of zero values.) A binary mask was created using the SPM8 tissue probabilistic map for gray matter (spm/spm8/apriori/gray.nii); all voxels with a gray matter probability greater than 0.2 were included. For all gray matter voxels (identified by the binary mask), the corresponding time courses were extracted from the imaging data and stored in an array of $N = 27{,}600$ time course vectors.

Low-pass filter. A 10th order low-pass Butterworth filter with a cutoff frequency of 0.125 Hz was applied to each time course vector. The filter coefficients were computed using the MATLAB "butter" function and applied using the "filtfilt" function (details are provided in MATLAB documentation).

Baseline correction. A linear baseline correction was applied to each time course vector using linear regression, and a second-order cosine basis set consisting of the two vectors with the components $x_i = \cos(i \cdot t/T \cdot \pi)$ with $t = 1, 2, \ldots, T$ and $i = 1, 2$; $T = 240$ is the number of time points. In addition to the two cosine basis vectors, the six realignment vectors generated by the SPM8 realignment routine during data preprocessing were included as covariates to correct for movement-related global signal changes simultaneously. The linear regression coefficients were calculated in MATLAB for each time course vector $\boldsymbol{v}$ according to $\boldsymbol{\beta} = (\boldsymbol{X}' \cdot \boldsymbol{X})^{-1} \cdot \boldsymbol{X}' \cdot \boldsymbol{v}$, where $\boldsymbol{X}$ is the $9 \times T$ design matrix that consisted of a constant column vector of ones followed by the two cosine basis vectors and the six realignment vectors. Finally, the residual (i.e., corrected) time course vector was calculated for each time course according to $\boldsymbol{v}_c = \boldsymbol{v} - \boldsymbol{X} \cdot \boldsymbol{\beta}$.

CHART analysis: Stage 1. The corrected time course vectors were normalized according to $\boldsymbol{v}_{cn} = (\boldsymbol{v}_c - \bar{v}_c)/\|\boldsymbol{v}_c - \bar{v}_c\|$, where $\bar{v}_c$ is the mean value of $\boldsymbol{v}_c$; the double bars represent the vector norm. (Because the design matrix used for the baseline correction contained a constant column vector of ones, the mean value $\bar{v}_c$ was 0 for all corrected time course vectors.) The normalized time course vectors $\boldsymbol{v}_{cn}$ were then stored in an $N \times T$ data matrix $\boldsymbol{V}$, where N is the number of gray matter voxels and T is the number of time points ($N = 27{,}600$, $T = 240$). The Pearson correlation coefficient between each pair of the normalized and corrected time course vectors was then calculated by multiplying $\boldsymbol{V}$ with its transpose $\boldsymbol{A} = \boldsymbol{V} \cdot \boldsymbol{V}'$; each matrix element a_{ij} of the resulting $N \times N$ correlation matrix $\boldsymbol{A}$ corresponds to the Pearson correlation coefficient between the ith and jth time course vectors. The matrix $\boldsymbol{A}$ is also known as the similarity matrix. (In principle, the Pearson correlation coefficient could have been calculated directly from the corrected time course vectors $\boldsymbol{v}_c$ using the MATLAB function "corr". However, because of the specific implementation of this routine, this approach is computationally inefficient for large values of N, which would have resulted in extensive processing times.) Maps of the WGC were then calculated from the corresponding correlation matrices $\boldsymbol{A}$ according to Eq. 1 for each subject and each condition (i.e., pre- and post-ECT treatment); this was implemented by using the MATLAB "mean" function to calculate the mean over all columns of $\boldsymbol{A}$ and by mapping the result back into 3D image space. A pair-wise nonparametric permutation test was used to test for differences between the pre- and post treatment WGC maps. The statistical nonparametric mapping software toolbox (SnPM8; http://www2.warwick.ac.uk/fac/sci/statistics/staff/academic-research/nichols/software/snpm) was used for the nonparametric permutation testing utilizing an FWE-corrected significance level of $P < 0.05$; further details about the nonparametric permutation framework and the FWE correction used in SnPM8 are described in refs. 1 and 2, respectively. The predefined SnPM8 design "MultiSub: Paired T test, 2 conditions, 1 scan/condition" was used. [This test essentially calculates the paired t value between pre- and posttreatment data on a voxel-by-voxel level; however, compared with a standard (parametric) t test, it does not use the t distribution to derive the level of significance. The corresponding null distribution is obtained by means of a nonparametric permutation approach.] The remaining SnPM8 parameters were as follows: "FWHM (mm) for variance smooth" = 0, "collect suprathreshold stats" = no, "no global normalization," "no grand mean scaling," "threshold masking" = none, and "analysis mask" = no. The SnPM8 analysis revealed a cluster of 12 voxels that exhibited a significant decrease in the WGC ($P < 0.05$, FWE-corrected) posttreatment (Figs. 1 and 2). In the following, we refer to this cluster of significant voxels as seed voxels.

CHART analysis: Stage 2. For all seed voxels, the normalized and corrected time course vectors $\boldsymbol{v}_{cn}$ were stored in an $M \times T$ data matrix $\boldsymbol{W}$. In analogy to the calculation of $\boldsymbol{A}$ (CHART analysis, stage 1), a seed correlation matrix was calculated according to $\boldsymbol{C} = \boldsymbol{W} \cdot \boldsymbol{V}'$; each matrix element c_{ij} of the resulting $M \times N$ correlation matrix $\boldsymbol{C}$ corresponds to the Pearson correlation coefficient between the ith seed voxel time course and jth time course vectors. The average correlation coefficient was then

calculated according to Eq. 2; note that for each seed voxel, the correlation of the corresponding time series with itself was excluded from the summation as described in the main text (*Methods*). Two seed correlation maps, showing each gray matter voxel's average connectivity with the seed region pre- and post-ECT treatment, were then created by mapping the corresponding average correlation values back into 3D image space.

Functional Connectivity: Surrogate Data Analysis. The aim of the surrogate data analysis was to calculate a statistical threshold for the pre- and posttreatment seed correlation maps (Fig. S1). The general workflow of the surrogate data analysis is illustrated in Fig. S2. Like in the previous section, we first provide a general overview before we describe each processing step in detail. A total of 1,000 random datasets were used in the surrogate data analysis. In analogy to the analysis of the experimental data, each random dataset consisted of nine (1 for each subject) arrays of $N = 27{,}600$ time course vectors. As shown in Fig. S2, the 1,000 random datasets were spatially smoothed and a temporal autocorrelation structure was generated that matched the temporal autocorrelation of the experimental data. Following low-pass filtering, the data were run through the second stage of CHART analysis (details are provided in the previous section) and the maximum statistic was calculated from the resulting seed correlation maps to control the FWE rate. Finally, a correlation value that corresponded to $P < 0.001$ (FWE-corrected) was derived from the maximum static distribution to give a correlation threshold of 0.50.

Random data. For each subject, an $N = 27{,}600$ array of random time course vectors (each consisting of $T = 240$ data points) was generated, which consisted of normally distributed random data (mean = 0, SD = 1) to match the size of the experimental data (details are provided in the previous section). Random numbers were created using the MATLAB function "randn". This process was repeated 1,000 times to generate 1,000 random datasets.

Spatial smoothing. To match the spatial smoothness of the experimental analysis, an 8-mm FWHM smoothing kernel was applied using the SPM8 spm_smooth function. This was implemented in the following way. First, the random time course vectors were mapped back into 3D image space (using the gray matter mask described in the previous section), the spatial smoothing was then applied using spm_smooth, and the spatially smoothed time course vectors were extracted (using the gray matter mask) and stored in the $N \times T$ matrix $\boldsymbol{V_r}$.

AR estimation. For each time course vector in the experimental data, the first-order temporal autocorrelation was estimated using a first-order autoregressive AR(1) model. The model parameters were estimated using the MATLAB function "aryule" and then applied to the corresponding smoothed random time course vector (row vector of matrix $\boldsymbol{V_r}$) using the MATLAB function "filter"; the results were stored in the matrix $\boldsymbol{V_{ra}}$.

Low-pass filter. To match the temporal smoothing of the experimental data, a low-pass filter identical to the one described in the previous section was applied to all the row vectors in the matrix $\boldsymbol{V_{ra}}$ and the result was stored in the matrix $\boldsymbol{V_{raf}}$.

CHART analysis: Stage 2. The time course vectors that corresponded to the seed voxels in the experimental data were extracted from the matrix $\boldsymbol{V_{raf}}$ and stored in a separate $M \times T$ seed data matrix $\boldsymbol{W_{raf}}$ ($M = 12$, $T = 240$). In analogy to experimental data analysis, a seed correlation matrix was calculated according to $\boldsymbol{C_{raf}} = \boldsymbol{W_{raf}} \cdot \boldsymbol{V_{raf}}'$, and a seed correlation map was calculated for each of the 1,000 surrogate datasets following the procedure described in the previous section.

Maximum statistic computation. For each surrogate dataset, the maximum value in the corresponding seed correlation map was calculated, resulting in an array of 1,000 correlation values, which constitutes the maximum statistic distribution. A correlation value that corresponded to $P < 0.001$ (FWE-corrected) was obtained from the upper 0.1 percentile of the maximum static distribution to give a correlation threshold of 0.50; further details regarding the underlying theoretical framework are described elsewhere (2).

fMRI Paradigm. The paradigm used was based on a modification of the ball-playing task described by Eisenberger et al. (3). This task involved three figures playing a game of catch, with two players controlled by the computer and one (the hand at the bottom of the screen) controlled by the participant in the scanner. The game was set up with varying blocks of "inclusion" (i.e., the ball being passed to the participant), with each block lasting ~35 s. The sequence of percent inclusion was as follows: 100, 75, 50, 25, 0, 25, 50, 75, and 100. This sequence was then repeated once. Once in the scanner, the participants were given two button boxes, one for each hand. The participants were told that they were going to be playing a simple game of catch with two other players and would be controlling the hand at the bottom of the screen. When the ball was thrown to them, they could pass the ball to the right or left by pressing the button in their right or left hand. They were told that it did not matter which direction they passed the ball.

ECT Treatment. ECT was administered using a brief-pulse, constant-current apparatus (Thymatron DGx, Somatics Inc.). All participants underwent the standard ECT protocol set by the Royal Cornhill Hospital and accredited by the Scottish Electroconvulsive Accreditation Network. At the first treatment, individual seizure thresholds were determined using standard stimulus dosing techniques, and subsequent treatments were given at twice the seizure threshold. Seizure threshold was defined as the electrical dose needed to produce a seizure of at least 15 s (measured by muscle contraction). To determine seizure threshold, patients received an initial dose of 50 mC followed by 100 mC and then 175 mC.

After placement of an i.v. cannula in the right or left arm, the patients were administered a minimally hypnotic dose of propofol starting at 0.75 mg/kg and increasing up to 2.5 mg/kg, followed by a limb isolation and the administration of the muscle relaxant suxamethonium (0.5–1 mg/kg). The electrodes were then placed bitemporally 4 cm above the midpoint between the lateral angle of the eye and the external auditory meatus, and electric current was then administered.

Patients received ECT twice a week, until the clinicians responsible for their care decided that recovery had taken place. Patients received an average of 8.3 treatments. Depression ratings were made before, during (after ECT treatment 4), and at the end of the course of treatment using the MADRS (*Methods*).

1. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum Brain Mapp* 15:1–25.
2. Nichols T, Hayasaka S (2003) Controlling the familywise error rate in functional neuroimaging: A comparative review. *Stat Methods Med Res* 12:419–446.
3. Eisenberger NI, Lieberman MD, Williams KD (2003) Does rejection hurt? An FMRI study of social exclusion. *Science* 302:290–292.

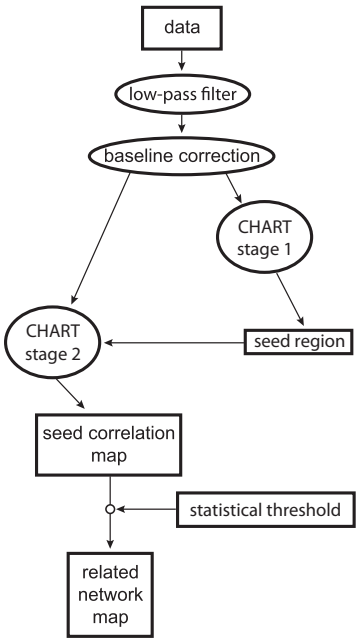

Fig. S1. Flow diagram of functional connectivity analysis. The CHART method consists of two stages. At stage 1, a seed region is derived based on significant differences in the WGC before and after ECT treatment. At stage 2, the related functional brain networks are calculated (i.e., the brain areas to which the seed region was functionally connected before and after ECT treatment; details are provided in *SI Text*).

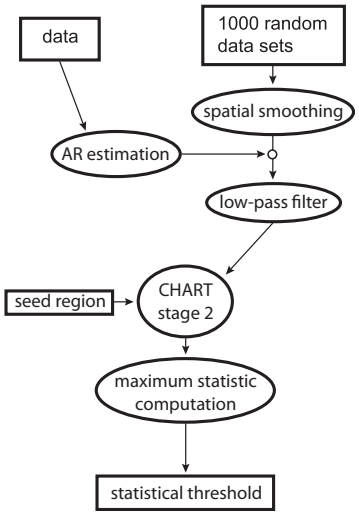

Fig. S2. Flow diagram of surrogate data analysis. A set of 1,000 surrogate datasets with matching spatial smoothness and temporal autocorrelation was used to calculate an FWE-corrected probability threshold (details are provided in *SI Text*).