Supporting Information

Hawellek et al. 10.1073/pnas.1110024108

SI Materials and Methods

Study Participants. Sixteen patients (13 female and 3 male, 23–46 y old, mean age: 35.3) and 16 healthy controls (13 female, 3 male, 23–49 y old, mean age: 33.6) participated in the experiments. Inclusion criteria for the patients were the diagnosis of either a clinically isolated syndrome suggestive of multiple sclerosis (MS) or definite MS (1), an Expanded Disability Status Scale (EDSS) (2) score <4 (excluding motor disabilities), and <4 y of disease duration. All patients were relapse-free and had ceased relapse-related cortisone treatment at least 3 mo before the start of the experiments. All control participants were matched to the patients individually for sex, age, and school education. Table S1 provides a list with further demographic and clinical information for all patients.

Neuropsychological Examination. All neuropsychological examinations were carefully conducted according to published guidelines, using an identical setup and procedure for each participant in a soundproofed room. Each examination began with questionnaires and self-evaluation measures: Demographic Information Questionnaire, Edinburgh Handedness Inventory (3), Modified Fatigue Impact Scale (4), and Hospital Anxiety and Depression Scale (HADS) (5). Subsequently, each participant performed a set of neuropsychological tests, probing a range of different cognitive modalities: Paced Auditory Serial Addition Test (6, 7), Symbol Digit Modalities Test (8), Trail Making Test (9), Digitspan (10), Verbal Intelligence Test "Mehrfachwortschatztest-B" (11), Controlled Oral Word Association Test with the letters B, A, S, N, and the word category "supermarket." Additionally, each participant performed on the following subtests of the test battery of attentional performance (12): alertness, covert shifts of attention, cross modal integration, flexibility and incompatibility. These tests represent reaction time-based measurements of cognitive functioning, including several classic task-paradigm components, such as spatial attention cueing, set shifting, and the Simon effect.

MRI. All scans were done on a 3T Siemens MAGNETOM Trio Scanner. Functional images were acquired using an echo planar imaging sequence in the axial plane [repetition time (TR) = 2 s, echo time (TE) = 25 ms, flip angle (FA) = 80°, voxel size = $4 \times 4 \times 4$ mm^3, matrix = 64×52, field of view = 256×208 mm^2, 36 slices for whole-brain coverage]. These resting-state runs included 606 frames corresponding to ~20 min in which the participants were told to silently fixate their view on a visually presented cross and stay awake. The following additional sequences were recorded: (*i*) T1-weighted image, using a coronal magnetization-prepared rapid gradient echo sequence (MPRAGE; TR = 2.3 s, TE = 2.98 ms, FA = 9°, inversion time (TI) = 1,100 ms, voxel size = $1 \times 1 \times 1$ mm^3); (*ii*) T2-weighted image, using a fast-spin echo sequence in the axial plane (TR = 6 s, TE = 91 ms, FA = 120°, slice thickness = 4 mm, field of view = 220 mm, 0.7×0.7 mm in-plane resolution, 25 slices); (*iii*) diffusion-weighted echo planar imaging sequence (TR = 17,200 ms, TE = 115 ms, FA= 90°, TI = 2,400 ms, b = 1,000 mm^2/s, 24 noncolinear directions, three averages, 45 slices, voxel size = $2 \times 2 \times 2$ mm^3).

Software and Visualization. If not indicated otherwise, data analyses were done in MATLAB (MathWorks) using custom-written software. The diffusion tensor parameters were displayed on mosaic slices of the atlas brain (MNI152) with the mask used for statistical analyses as a black underlay. For visualizing the fMRI results, we projected the data on the inflated surface of the Population-Average, Landmark-, and Surface-Based (PALS) atlas (13).

Behavioral Analysis. The analysis of the behavioral data was confined to a subset of 10 neuropsychological test measures, including the cognitively more challenging subpart of each test. The abbreviations used to index the test names in the figures refer to the following test metrics: Trail Making, Trail Making Test Part B; Pac. Aud. Ser. Add, Paced Auditory Serial Addition Test in the 2-s version; verbal fluency, average performance on the letters B, A, S, and N of the Controlled Oral Word Association Test; flexibility, reaction time on a set shifting task [flexibility, subtest of the Test Battery of Attentional Performance (TAP)]; alertness, reaction time on a target detection task (alertness, subtest of the TAP); Symb. Dig. Mod., Symbol Digit Modalities Test; incompatibility, reaction time in a Simon task on incompatible trials (incompatibility, subtest of the TAP); spat. att., reaction time in a Posner paradigm on invalid cue trials (covert shifts of attention, subtest of the TAP); dig. span. fwd., longest digit span forward; dig. span bwd., longest digit span backward.

Before further analysis, the raw test results were transformed into deviations from the corresponding age and education norm of each test for each participant. The behavioral parameter cognitive efficiency was then estimated as the first principle component of the test performance covariance matrix. To control for possible overfitting effects within the small data set, we made several analyses using a leave-one-out cross-validation (LOOCV) procedure (Fig. S1 *B* and *C*). Here, the principal component analysis (PCA) was repeated once for each participant. In each round, one dataset (test data set) was left out of the analysis, deriving the components' eigenvector (weights) only from the remaining data (training data set). Afterward, the left-out data were projected onto the derived weights. This procedure resulted in one estimate of the components' explained variance per participant, without the specific data of any participant directly contributing to the components' structure. Fig. S1 *B* and *C* report the average LOOCV estimates of the components' values across participants. Sensitivity and specificity were evaluated for a binary classification into patients and controls based on the level of cognitive efficiency (Fig. S1*D*). Sensitivity reflects the true positive rate (fraction of correctly identified patients), whereas specificity reflects the true negative rate (fraction of correctly identified healthy controls).

fMRI Preprocessing. The functional data were realigned within scanning runs to correct for head motion using an eight-parameter (rigid body plus in-plane stretch) cross-modal registration. Differences in the acquisition time of each slice within a frame were compensated for by sync interpolation. A whole-brain normalization factor was applied to correct for changes in signal intensity between runs (mode of 1,000). For each subject, an atlas transformation was computed on the basis of the first frame of each functional run, the T2-weighted and MPRAGE structural images to the atlas representative target using a 12-parameter general affine transformation. Functional data were interpolated to 3-mm^3 voxels in atlas space. The atlas representative MPRAGE target brain (711-2C) was produced by mutual coregistration (12-parameter affine transformations) of images obtained in six young and six older subjects. In preparation for functional connectivity analysis, data were passed through several additional preprocessing steps: (*i*) spatial smoothing (6-mm FWHM Gaussian

blur); (*ii*) temporal filtering retaining frequencies in the 0.009- to 0.08-Hz band; and (*iii*) removal of several sources of spurious variance unlikely to reflect spatially specific functional correlations through linear regression: (*a*) six parameters obtained by rigid-body correction of head motion, (*b*) the whole-brain signal averaged over a fixed region in atlas space, (*c*) the signal from a ventricular region of interest, (*d*) the signal from a region centered in the white matter ; and (*e*) the first derivatives of all regressors. These additional preprocessing steps help to separate true from spurious sources of signal variance and have been shown to result in a substantial improvement in the accessibility of functional networks in resting-state fMRI data (14, 15). All preprocessing steps were performed using in-house software. Before the actual analysis, the first five frames of the functional data were removed to allow settlement of the signal. For one control subject (C03), only the first 330 frames (11 min) of the functional data could be used for analysis due to excessive movement artifacts toward the end of the recording.

Control Analyses. To assess whether the preprocessing of the functional data or movement levels might have had confounding effects on the results, we carried out several control analyses. The removal of a global signal influences the presence of anti-correlations in the data (15) and calls into question the physiological origin of the across networks effect. We thus repeated the connectivity analysis without regression of the global signal and found all connectivity effects to still persist (correlation with cognitive efficiency, DMN connectivity: $r = -0.65$, $P = 4 \times 10^{-5}$; CN connectivity: $r = -0.66$, $P = 4 \times 10^{-5}$; across-network connectivity: $r = 0.58$, 4.9×10^{-4}). To further assess whether the removal of any of the noise regressors' variance might have had a confounding effect for our analyses, we tested whether the amount of variance that each regressor removed from the BOLD data of each participant was associated with either cognitive efficiency or the connectivity across subjects. There were no significant correlations (Pearson's r, all $P > 0.1$).

We assessed the levels of head movement during the resting-state recordings by calculating the peak-to-peak excursion (PE) and rms for each of six head-movement parameters (three translational, three rotational) (16). Before calculating the metrics, the rotational head-movement levels were transformed from degrees into the most conservative worst-case millimeter rotational movements by multiplying with 72.6 mm, the geometric mean of the maximal extents of the standard space's brain mask. The maximal movement components across the six parameters within each subject are reported here. The movement levels were 2.03 ± 1.03 mm (PE) and 0.45 ± 0.24 mm (rms) for the patients and 1.72 ± 0.92 mm (PE) and 0.47 ± 0.31 mm (rms) for the healthy controls, not significantly differing between the groups (t tests, PE: $P = 0.38$, rms: $P = 0.85$). None of the movement parameters were correlated with cognitive efficiency (all $P > 0.17$). Thus, neither the preprocessing nor differences in the levels of gross head movement had confounding effects on the results.

Modulation of Functional Connectivity. We restricted the analyses to voxels within a standard-space brain mask, which resided within 6 mm Euclidian distance of the standard-space target brain's cortical surface (13). In a first step, the global connectivity matrix of these voxels was calculated for each participant. In this matrix each entry represents the correlation of two voxels' BOLD time series. Before passing these matrices to further analysis, all values were Fisher's z-transformed. To quantify the modulation of connectivity by the level of cognitive efficiency, the group-level correlation of the connectivity strength between any two voxels with cognitive efficiency was calculated. This calculation resulted in a matrix of global connectivity modulations with each row representing the modulation profile (increases and decreases in functional connectivity) of a given voxel. We derived the raw modulation for each voxel as the sum across all columns (or rows) of this modulation matrix after statistical thresholding. All results were obtained with a threshold that corresponded to an uncorrected α-level of $P = 0.01$. Changing this threshold within reasonable ranges ($P = 0.05$ to $P = 0.001$) yielded highly similar results. We then performed permutation statistics on the raw modulation by repeating the entire procedure 100× while randomly permuting the behavioral parameter on each round. A normal distribution was fitted to these resamples to derive an empirical distribution for the null hypothesis of no connectivity modulation with cognitive efficiency. This distribution was used to derive z-scores (subtraction of mean and division by the SD) and P values.

We extracted the dominant pattern of the connectivity modulations (Fig. 4*B*) as the first principal component of all modulation profiles identified in the procedure described above. To map the actual range of connectivity that was indicated by the modulations (Fig. 4 *D–F* and Fig. S4*C*), we constructed for each participant a connectivity graph based on these PCA results. The graph contained the voxels belonging to the top 5% of negative (DMN) or positive values (CN) of the dominant modulation pattern (Fig. S4 *A* and *B*) as well as those voxels whose loadings of this pattern ranked within the top 50% of negative or positive values. Thus, the resulting graph contained four voxel groups: the two networks of the dominant spatial pattern of modulations (DMN, CN) and the two voxel clusters, which exhibited this pattern of modulation in an inverted way (Fig. 4*C*). We excluded any voxel that occurred in more than one of these groups. First, the graphs were Fisher's z-transformed, then we derived the within- and across-network connectivity as the average connectivity between the corresponding voxel groups for each participant.

1. McDonald WI, et al. (2001) Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann Neurol* 50(1):121–127.
2. Kurtzke JF (1983) Rating neurologic impairment in multiple sclerosis: An Expanded Disability Status Scale (EDSS). *Neurology* 33:1444–1452.
3. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia* 9(1):97–113.
4. Zimmermann C, Hohlfeld R (1999) "Fatigue" in multiple sclerosis (Translated from German). *Nervenarzt* 70:566–574.
5. Snaith P, Zigmond AS (1988) Anxiety and depression in general medical settings. *BMJ* 297:1544.
6. Rao SM, Leo GJ, Haughton VM, St Aubin-Faubert P, Bernardin L (1989) Correlation of magnetic resonance imaging with neuropsychological testing in multiple sclerosis. *Neurology* 39(2 Pt 1):161–166.
7. Gronwall DM (1977) Paced auditory serial-addition task: A measure of recovery from concussion. *Percept Mot Skills* 44:367–373.
8. Parmenter BA, Weinstock-Guttman B, Garg N, Munschauer F, Benedict RHB (2007) Screening for cognitive impairment in multiple sclerosis using the Symbol Digit Modalities Test. *Mult Scler* 13(1):52–57.
9. Tombaugh TN (2004) Trail Making Test A and B: Normative data stratified by age and education. *Arch Clin Neuropsychol* 19:203–214.
10. Kaufman AS, Lichtenberger EO (2006) *Assessing Adolescent and Adult Intelligence* (Wiley, New York).
11. Lehrl S (1991) *Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B)* (Straube, Erlangen, Germany), 2nd Ed.
12. Zimmermann P, Fimm B (2002) A test battery for attentional performance (Chapter 4). *Applied Neuropsychology of Attention: Theory, Diagnosis and Rehabilitation*, eds Leclercq M, Zimmermann P (Psychology Press, London), pp 110–151.
13. Van Essen DC (2005) A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28:635–662.
14. Fox MD, et al. (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
15. Fox MD, Zhang D, Snyder AZ, Raichle ME (2009) The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol* 101:3270–3283.
16. Thomason ME, Glover GH (2008) Controlled inspiration depth reduces variance in breath-holding-induced BOLD signal. *Neuroimage* 39:206–214.

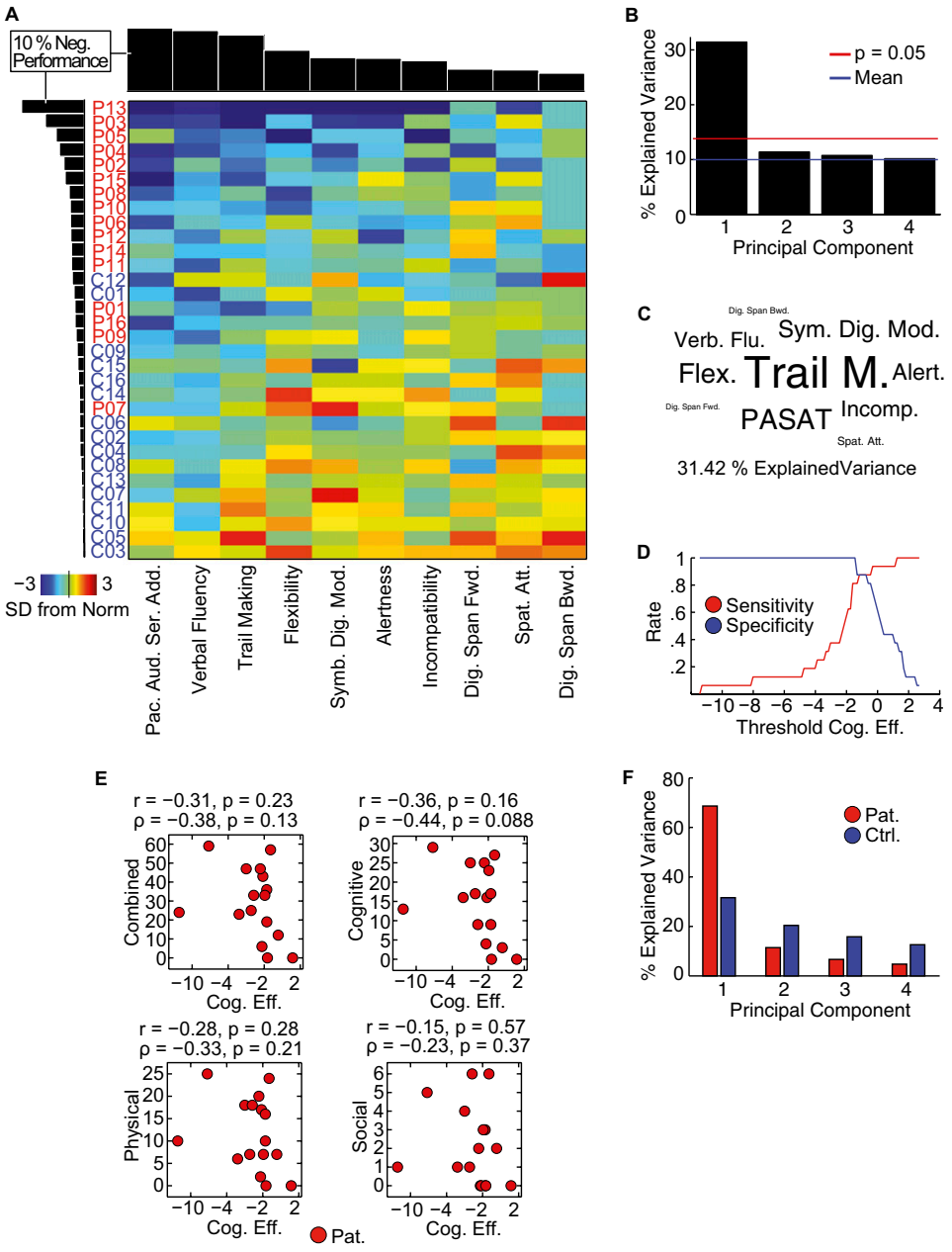

Fig. S1. (*A*) Performance matrix as shown in Fig. 1*A* for the combined patient and control dataset. (*B*) Average explained variance of the first four principal components of the combined patient and control dataset, using a LOOCV procedure (*SI Materials and Methods*). The blue line indicates the average percent explained variance of Gaussian noise weightings, obtained by projecting the data onto random eigenvectors with unit length 1,000×. The red line corresponds to the 95th percentile of this distribution. Crossing this line indicates the component's explained variance to be significantly different from Gaussian noise at an α-level of $P = 0.05$. These results suggest that only the first behavioral component contains meaningful predictive power for further analyses ($P < 10^{-16}$). (*C*) Test name size scaled by the loading of the first principle component derived from the combined dataset using a LOOCV procedure. These loadings correspond to the first component shown in *C* (cf. Fig. 1*B*). Spatial test name positions are arbitrary. (*D*) Sensitivity and specificity for a binary classification between patients and controls based on cognitive efficiency. (*E*) The relation of the level of cognitive efficiency to fatigue. Each scatter plot shows the correlation of cognitive efficiency with a subscale of the Modified Fatigue Impact Scale. (*Left* to *Right*) All subscales combined, cognitive subscale, physical subscale, and social subscale. The results of correlation analysis using Pearson's *r* as well as Spearman's ρ are given above the panels. (*F*) Explained variance of the first four principal components calculated for each study group separately. The control group exhibited a richer pattern of behavioral variability, which led to a wider distribution of components.

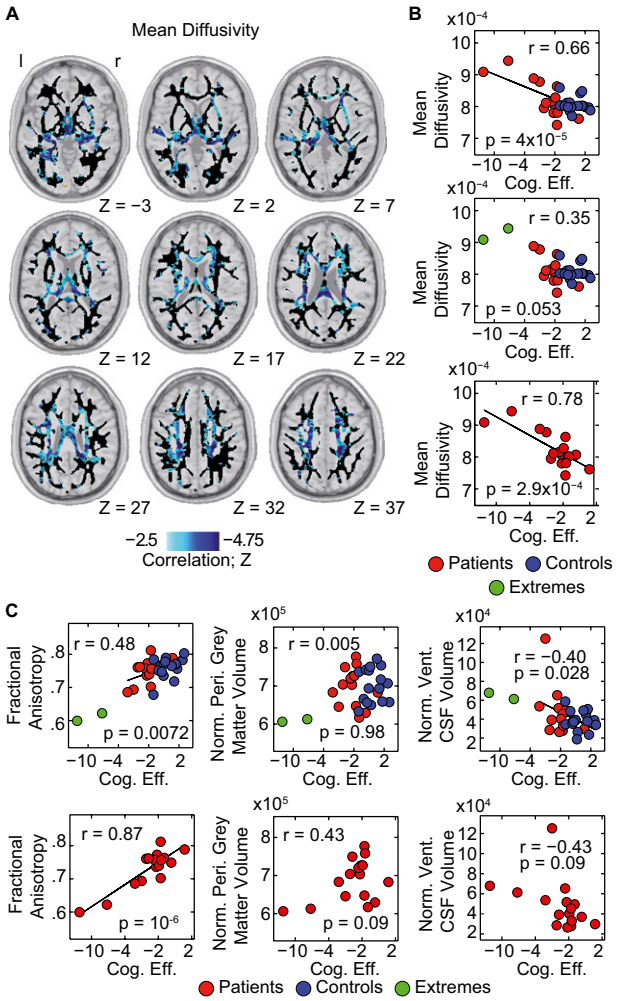

Fig. S2. (*A*) Correlation of cognitive efficiency with voxel-wise mean diffusivity. The analysis was performed within a standard-space white matter mask shown as a black underlay. Correlations are shown as corresponding z-scores and have been thresholded at a false discovery rate-corrected α-level of $P = 0.05$. No positive correlations passing this threshold were observed. (*B*) The relation of the average mean diffusivity within a corpus callosum mask to the level of cognitive efficiency for different compositions of the study participants. (*Top*) All participants; (*Middle*) excluding the two most extreme patients; (*Bottom*) patients only. (*C*) Corresponding figures to *B* for the three anatomical parameters shown in Fig. 2.

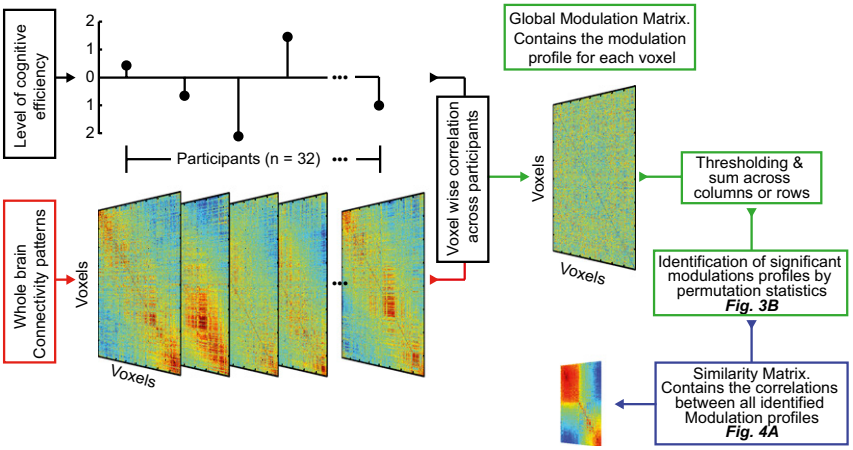

Fig. S3. Schematic depiction of the analysis stream. The colored frames correspond to the three levels of analysis shown in Fig. 3*A*. For each participant, we calculated a whole-brain connectivity graph in a first step. Each entry in this matrix represents the correlation (Pearson's *r*) of two voxels' BOLD time series. To quantify the change of connectivity between two voxels across participants, we calculated the correlation between the connectivity of each voxel pair and the level of cognitive efficiency, which resulted in the global connectivity modulation matrix. The amount of connectivity modulation for each voxel was subsequently calculated as the sum across all columns (or rows) after thresholding. We derived z-scores and *P* values by subtracting the mean and dividing by the SD of an empirical null hypothesis distribution of connectivity modulations at each voxel. This distribution was obtained by randomly permuting the behavioral parameter and repeating the analysis 100×. Based on these results, we constructed the similarity matrix, containing the correlations between all identified modulation profiles.

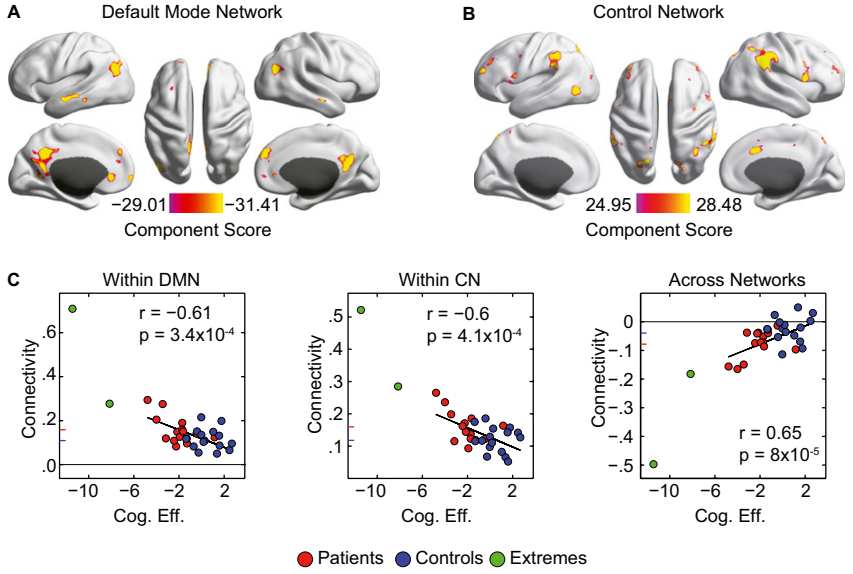

Fig. S4. (*A*) The top 5% of negative values of the dominant spatial pattern of connectivity modulations (shown in Fig. 4*B*). These voxels were used to derive the connectivity to the default mode network. (*B*) The top 5% of positive values of the dominant spatial pattern of connectivity modulations. These voxels were used to derive the connectivity to the control network. (*C*) Relation of the identified connectivity to cognitive efficiency effects when omitting the two most extreme cases of the patient group. The correlation of connectivity with behavior is given in the corresponding panels. The group means of the connectivity are indicated as colored bars at the ordinate. All connectivity measures significantly differed between the groups (*t* test, DMN: $P = 0.028$, CN: $P = 0.012$, across networks: $P = 0.018$).

Table S1. Clinical and demographic information of all patients

Patient no.	Age	Sex	School education, y	Diagnosis	Disease duration, y/mo	EDSS	Medication	HADS depression status	HADS anxiety status
P01	23	Female	10	RRMS	1.75/21	0	Copaxone	Normal	Normal
P02	30	Female	13	CIS	0.5/6	2	—	Normal	Positive
P03	34	Male	10	MS	1.25/15	0	—	Normal	Normal
P04	41	Female	10	RRMS	1.83/22	3	—	Normal	Ambiguous
P05	36	Male	10	RRMS	1.5/18	1.5	—	Normal	Ambiguous
P06	29	Female	10	RRMS	3/36	2	Copaxone	Normal	Normal
P07	45	Female	13	RRMS	2.08/25	2	—	Normal	Normal
P08	36	Female	12	RRMS	2.67/32	1	Copaxone	Normal	Normal
P09	36	Female	13	RRMS	2.17/26	3	Rebif	Normal	Ambiguous
P10	25	Female	10	RRMS	2.33/28	0	Mitoxanthrone	Ambiguous	Positive
P11	39	Female	10	RRMS	3.66/44	2	Copaxone	Normal	Positive
P12	46	Female	12	CIS	2.17/26	2	Copaxone	Normal	Normal
P13	41	Male	10	RRMS	1.16/14	2.5	Avonex	Ambiguous	Normal
P14	43	Female	13	RRMS	1.75/21	3.5	—	Normal	Normal
P15	29	Female	12	MS	3.66/44	0	—	Normal	Normal
P16	32	Female	13	RRMS	1.75/21	0	—	Normal	Ambiguous
Mean	35.3		11.31		2.03/24.93	1.53			
Median	36		11		1.95/23.5	2			
SD	6.9		1.40		0.86/10.29	1.21			

CIS, clinically isolated syndrome; RRMS, relapsing-remitting MS.

Table S2. Average coordinates of the 10 largest clusters exhibiting significant ($P < 0.05$, false discovery rate-corrected) modulation

Cluster size, voxels	Anatomical label	*X*	*Y*	*Z*
53	BA 23	5.83	−53.97	16.44
50	BA 39	−44.1	−70.68	18.72
37	BA 10	−3.4	60.2	0.12
21	BA 39	40.28	−67.5	19.92
18	BA 39	50	−63.83	9.66
18	BA 10	−27.16	57.66	−1
14	BA 40	51.21	−44.57	30.64
13	BA 31	−10.84	−44.88	27.11
13	BA 24	−3.46	29.19	6.8
12	BA 45	38.25	27.75	5.5

Data from Fig. 3*B*.