Supplementary Material

S1. Comparison of within- and between-subject similarity for Datasets A and B.

The top panels of Figure S1 display the comparison of the within-subject and between-subject similarities (WSS and BSS, respectively) for the covariances FC0 and FC1. They are very similar to those for corrFC in Figure 2B (main text). The bottom left panel shows the same distributions for corrFC/SC, which are the elements of the usual correlation-based FC (corrFC) that correspond to existing connections, as determined by the structural connectivity (SC). The distributions are more compact than with corrFC (as well as FC0 and FC), but the overlap is similar, as indicated by the standard deviation represented by the error bars. In summary, all FC-based connectivity measures exhibit similar overlap between WSS and BSS, which is larger than that for EC in Figure 2B (main text).

The bottom right panel shows the overlap for Σ, the local excitability estimated by the model. The overlap there is larger and, in general, the similarity for the Σ profiles is much higher than that for the connectivity measures.

The results for Dataset B are qualitatively similar to those for Dataset A in Figures 2 and S1. A quantification of the overlap using the Kolmogorov-Smirnoff distance is given in Table 2 (main text).

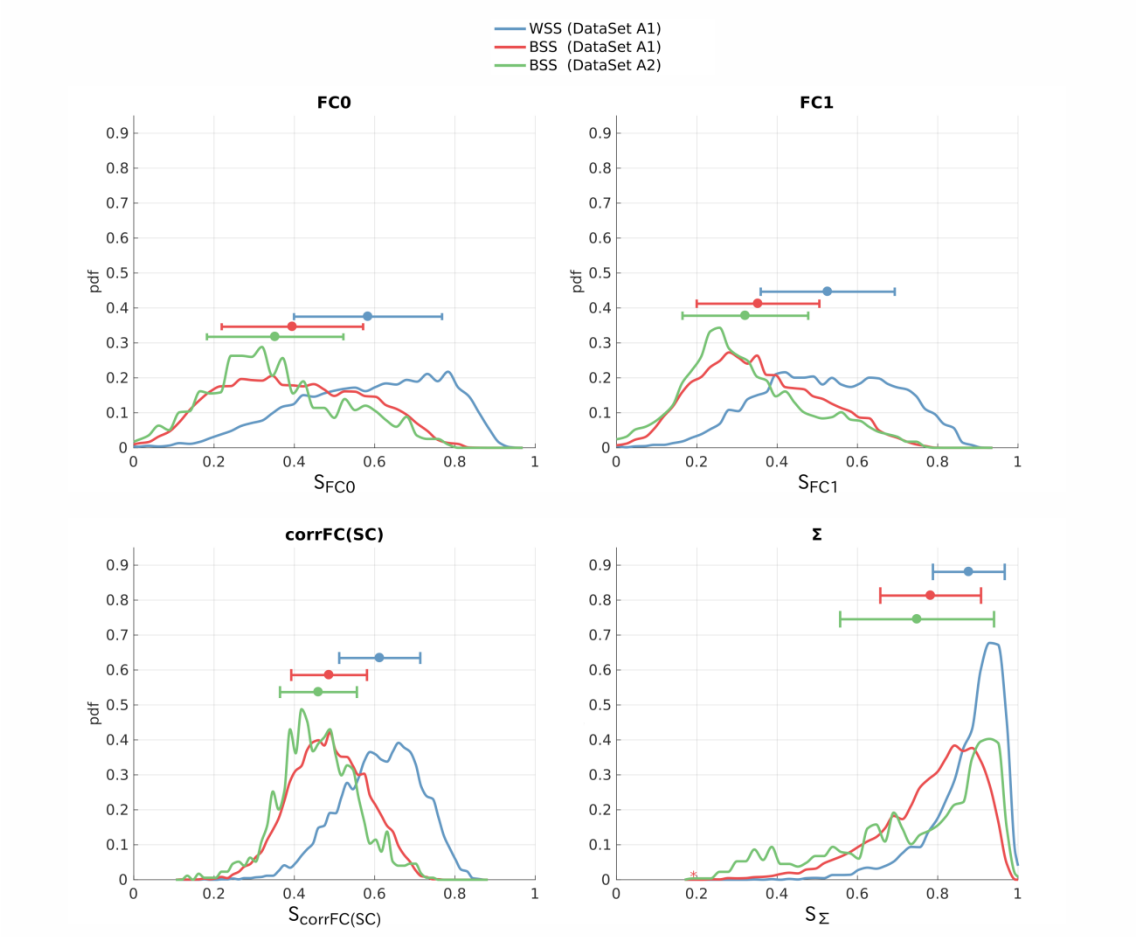

Figure S1. Similarity distributions of FC0, FC1, corrFC/SC and Σ calculated form Dataset A. Distributions of WSS and BSS similarity values for Datasets A1 (blue and red), as well as the BSS values for Dataset A2 (green). The plot is similar to Figure 2B: similarity is calculated using the Pearson correlation coefficient (PCC) between every pair of sessions from the same subjects or from different subjects respectively.

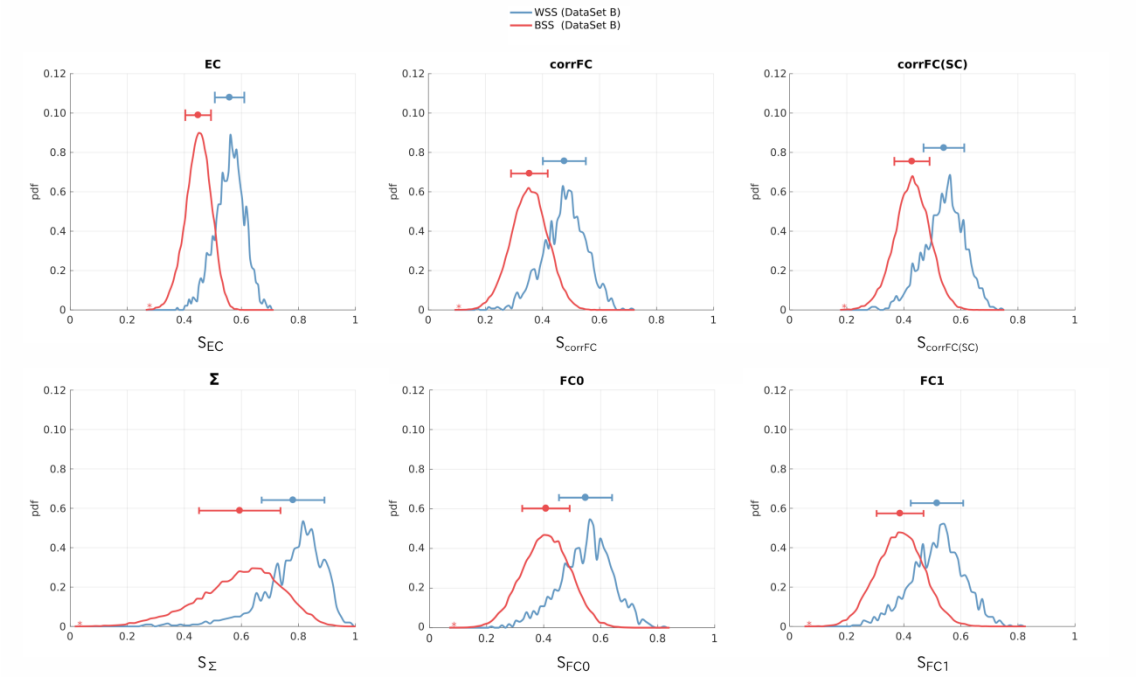

Figure S2. Similarity distributions for Dataset B: comparison of EC, corrFC, FC0, FC1, corrFC/SC, and Σ. Same as Figure 2B and S1 for Datasets B with WSS and BSS similarity values (blue and red).

S2. Degree of clustering in the space of principal components (PCs)

When applying principal component analysis (PCA), the silhouette coefficients become significantly larger (indicating a better separability), as can be seen in Figure 2E (main text). However, this depends on the choice of PCs used to project the data (6 in Figure 2E with Dataset A1). The curves in Figure S3 show the evolution of the mean of the distributions over all sessions and subjects (similar to those in the right panels of Figure 2E) when varying the number of first PCs (indicated on the x-axis). With both Datasets A1 and B, the mean silhouette coefficient for EC increases until the number of PCs reaches the number of subjects (6 and 30, respectively), then decreases. This means that the largest variability of EC estimates captures by the first PCs conveys individual information about the subjects. In contrast, the mean silhouette coefficient for FC is much lower; it also requires more PCs to reach its maximum, suggesting that the largest part of variability with corrFC (first PCs) is not very informative for subject identification.

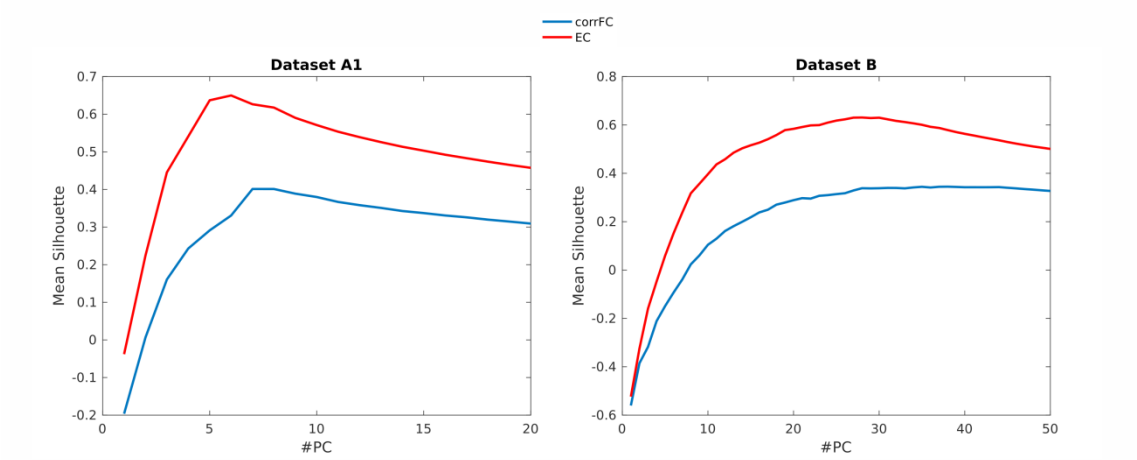

Figure S3. Average silhouette coefficient as a function of the number of first PCs (x-axis) used to calculate the similarity. PCA was applied on the whole set of subjects for both Datasets A1 and B. The silhouette coefficients are calculated for all sessions after projecting the original data in the space of the first PCs (from 1 to 20 for Dataset A1; from 1 to 50 for Dataset B). The curves correspond to the mean silhouette for EC (in red) and corrFC (in blue), with large value indicating the good quality of clustering.

S3. Classification of subjects using a 1NN classifier

Figure S4 depicts the method used in previous studies to identify subjects from resting-state sessions (Finn 2015; Kaufman 2017).

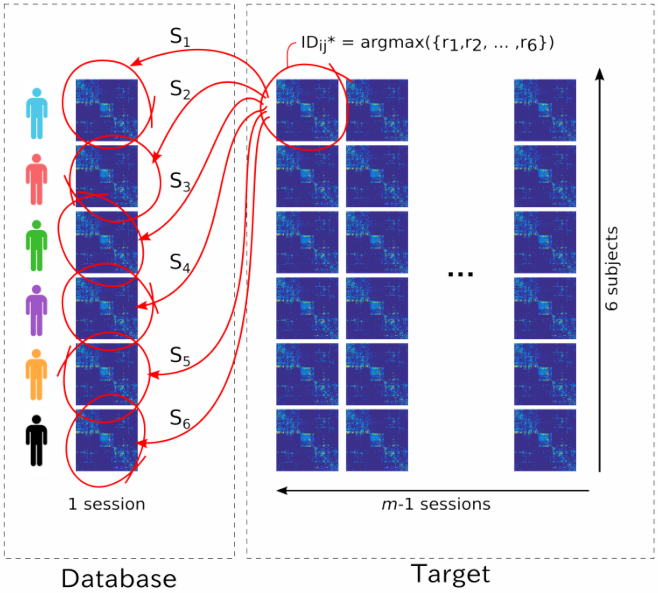

Figure S4. Principle of the 1NN classifier. For each target session to classify, we calculate the similarity values (as measured by PCC, same as in Figure S1 and S2) with each session in the database, giving S_1 to S_6 with Dataset A1 (6 subjects, 40-50 sessions per subject) as shown here. The session is attributed to the subject with largest similarity.

S4. MLR and 1NN in the PCA space

As mentioned in the main text about the silhouette coefficients, PCA may improve the classification by reducing the high dimensionality of the data. However, an important issue is whether the variability captured by first PCs – which account for the largest part of the total variability – carries information about subjects' identities or if it is rather related to the session-to-session variability. If the latter is true, the classification performance might even be worsened by an erroneous choice of PCs.

To firstly examine the influence of PCA on the classification, we optimized the 1NN and MLR while varying the number of training sessions with Dataset A1. Recall that PCA was calculated using the training sessions only (see Figure 3A). This hardly changed the performance for the MLR (less than 1% for FC, almost zero for EC), whereas the performance of the 1NN increased by a few % (but remained far below that of the MLR). This suggests that the 1NN benefited more from the data denoising. The small improvements by PCA for the MLR may come from the fact that the classification accuracy for MLR is already very high (~95% with EC and 1 training session); recall that EC is much less noisy than corrFC (see Figure 2B). Similar results were obtained for Dataset B.

PCA also allowed for the investigation of the distribution of the subject-specific information between PCs supporting the classification. By choosing the number of PCs to train the classifier with Dataset B, fewer PCs are sufficient to obtain a close-to-perfect classification for EC than corrFC, as shown in Figure S6. When the accuracy stabilizes close to 1, this means that the remaining PCs carry redundant information for the subject identification.

This is further illustrated for Dataset A1 in Figure S7, where the left panel shows that the first 6 PCs convey relevant information for subject identification, especially PC3 to PC6. The very first PCs actually are less informative than the following ones in Figure S8: PC2 for Dataset A1 and PC1 to PC6 (without PC4) for Dataset BThis suggests that the main variability (first PCs) of EC is common to all subjects; subsequent PCs with lower variability are actually more informative about the subjects' identity. Note also that PC10, PC11, PC13 and PC17 are as discriminative as PC4 and PC7 are (and much less than other PCs), meaning that there is not clear relationship between the order of PCs and subject-specific information. This supports the use of proper machine learning tools to extract this distributed information, because the PCs may mix subject-specific traits differently when the number of subjects increases. A proper extraction of discriminative features in the EC links is thus necessary, which is exactly what the MLR algorithm does.

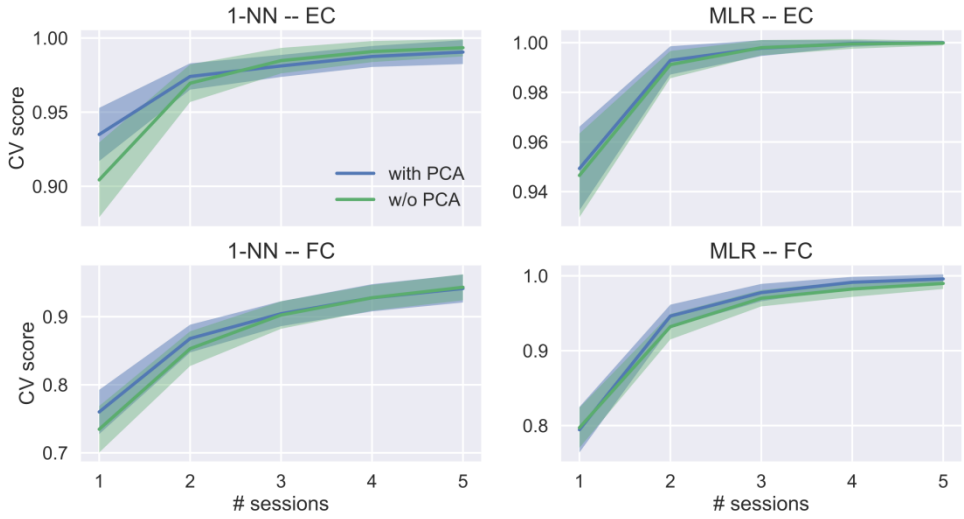

Figure S5. Marginal improvement of the classification performance by PCA. The panelscorrespond to the classification of sessions using 1NN (left column) and MLR (right column) with EC (top row) and corrFC (bottom row). The performance is plotted as a function of the number of training sessions (x-axis) using Dataset A1.

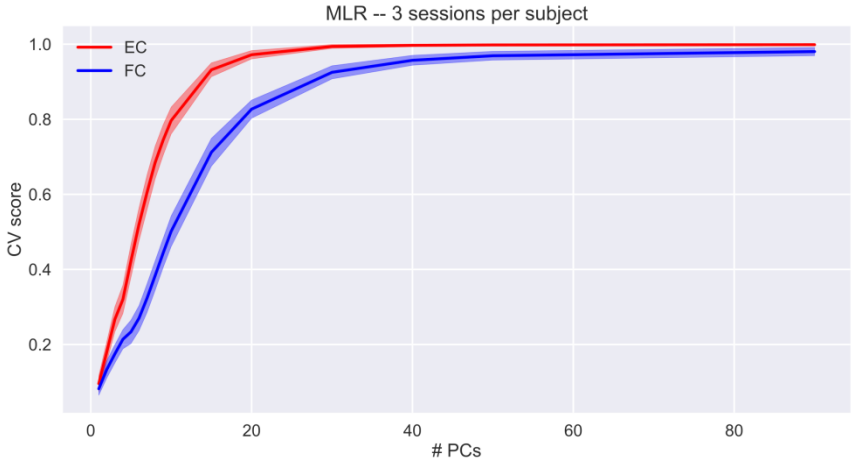

Figure S6. Performance of MLR classifier for EC and corrFC sessions when increasing the number of PCs. Dataset B with 30 subjects and 3 sessions per subject was used for training. PCA was calculated using the training sessions only. The information relevant for the classification is distributed across the first 20 PCs for EC, as indicated by the steep increase. More PCs (at least 40) are necessary with corrFC for the classifier to reach a similarly quasi-perfect performance.

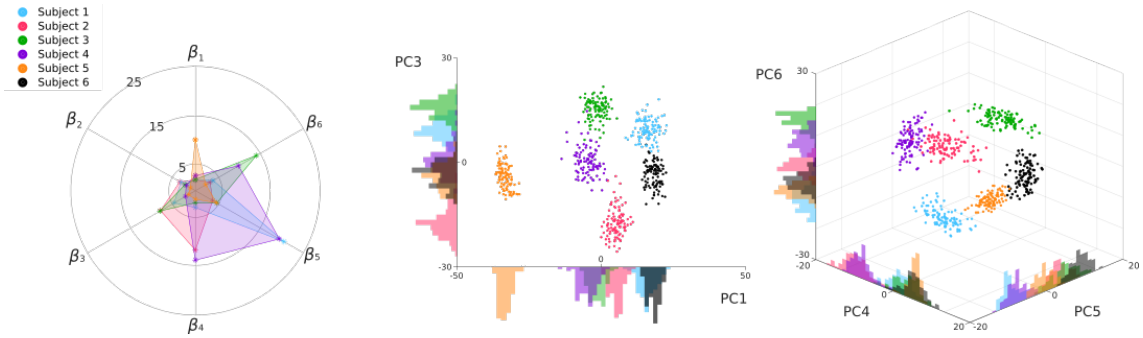

Figure S7. Decomposition of the contributions along each principal component (PC) of the MLR+PCA classifier. The MLR was trained using EC from Dataset A1 with 20 sessions per subject; these many training sessions capture the overall session-to-session variability of the whole DatasetA1. The left plot shows for subjects 1 to 5 (in color, corresponding each to a regressor of the classifier against the 6th subject) the weight for each PC (beta 1 to 6). Large values indicate that PC strongly contributes to the discrimination of the subject. The middle and right panels display a visualization of many averages of 20 sessions (distinct from the training ones, represented each by a colored dot) in the PC space; PC2 is omitted.

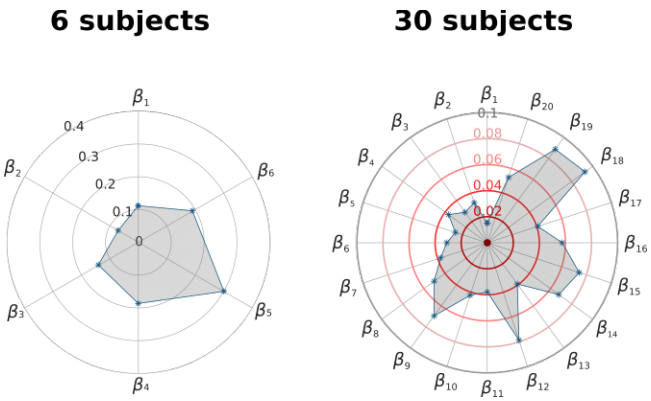

Fig S8. Comparison of the contributions of all MLR+PCA regressors along the PCs for Datasets A1 and B. After training the MLR with PCA on EC, we calculate for each PC a single value that measures the importance of that PC in the global classification; it simply collapses the values in the left panel of Figure S6 by summing them for each PC.

S5. Evaluation of the number of EC links necessary for quasi perfect classification with MLR.

This is the classification performance of subjects after sorting the most discriminant EC links selected by recursive feature extraction (RFE, see Methods for details). Here 90% of all available sessions were used for training and the remaining 10% for testing, in order to account for the session-so-session variability in the train set. The matrix of EC links corresponding to the saturation of the curves close to perfect classification is depicted in Figure 3D: 18 links for Dataset A1 and 44 for Dataset B.

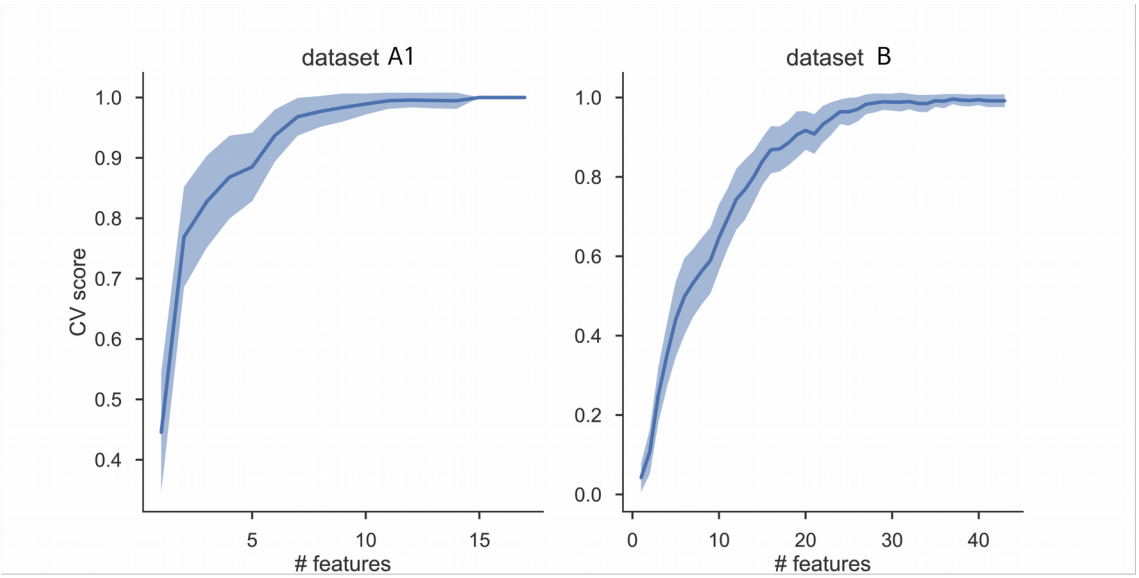

Figure S9: Test set accuracy of MLR classifier varying number of EC links (features, x-axis). The order of the EC links is determined by RFE, based on their individual contribution to the correct subject identification.

Supplementary Table S1
List of ROIs for the AAL parcellation with groups for the matrix in Figures 3d (main text) and S10.

ROI index	Brain region	Subsystem
1, 2	Calcarine fissure	occipital
3, 4	Cuneus	
5, 6	Lingual gyrus	
7, 8	Superior occipital gyrus	
9, 10	Middle occipital gyrus	
11, 12	Inferior occipital gyrus	
13, 14	Fusiform gyrus	
15, 16	Heschl gyrus	temporal
17, 18	Superior temporal gyrus	
19, 20	Superior temporal pole	
21, 22	Middle temporal gyrus	
23, 24	Middle temporal pole	
25, 26	Inferior temporal gyrus	
27, 28	Precentral gyrus	central
29, 30	Postcentral gyrus	
31, 32	Paracentral lobule	
33, 34	Superior parietal gyrus	parietal
35, 36	Inferior parietal gyrus	
37, 38	Supramarginal gyrus	
39, 40	Angular gyrus	
41, 42	Precuneus	
43, 44	Superior frontal gyrus, dorsolateral	frontal
45, 46	Superior frontal gyrus, orbital part	
47, 48	Middle frontal gyrus	
49, 50	Middle frontal gyrus, orbital part	
51, 52	Inferior frontal gyrus, opercular part	
53, 54	Inferior frontal gyrus, triangular part	
55, 56	Inferior frontal gyrus, orbital part	
57, 58	Rolandic operculum	
59, 60	Supplementary motor area	
61, 62	Olfactory cortex	
63, 64	Superior frontal gyrus, medial	
65, 66	Superior frontal gyrus, medial orbital	
67, 68	Gyrus rectus	
69, 70	Insula	cingulate
71, 72	Anterior cingulate	
73, 74	Median cingulate	
75, 76	Posterior cingulate	
77, 78	Hippocampus	
79, 80	Parahippocampal gyrus	
81, 82	Amygdala	
83, 84	Caudate nucleus	
85, 86	Putamen	
87, 88	Pallidum	
89, 90	Thalamus	
91, 92	Cerebellum crus 1	cerebellum
93, 94	Cerebellum crus 2	
95, 96	Cerebellum 3	
97, 98	Cerebellum 4-5	
99, 100	Cerebellum 6	
101, 102	Cerebellum 7	
103, 104	Cerebellum 8	
105, 106	Cerebellum 9	
107, 108	Cerebellum 10	
109, 110	Vermis 1-2	
111, 112	Vermis 4-5	
113, 114	Vermis 7	
115, 116	Vermis 9	

Supplementary Table S2
List of ROIs for the Hagmann parcellation with groups for the matrix in Figure 4c.

ROI index	Abbreviations	Brain region	Subsystem
1, 2	CUN	Cuneus	occipital
3, 4	PCAL	Pericalcarine cortex	
5, 6	LING	Lingual gyrus	
7, 8	LOCC	Lateral occipital cortex	
9, 10	FUS	Fusiform gyrus	
11, 12	ST	Superior temporal cortex	temporal
13, 14	TT	Transverse temporal cortex	
15, 16	MT	Middle temporal cortex	
17, 18	IT	Inferior temporal cortex	
19, 20	PREC	Precentral gyrus	central
21, 22	PSTC	Postcentral gyrus	
23, 24	PARC	Paracentral lobule	
25, 26	SP	Superior parietal cortex	parietal
27, 28	IP	Inferior parietal cortex	
29, 30	TP	Temporal pole	
31, 32	SMAR	Supramarginal gyrus	
33, 34	BSTS	Bank of the superior temporal sulcus	
35, 36	PCUN	Precuneus	
37, 38	FP	Frontal pole	frontal
39, 40	CMF	Caudal middle frontal cortex	
41, 42	RMF	Rostral middle frontal cortex	
43, 44	PTRI	Pars triangularis	
45, 46	PORB	Pars orbitalis	
47, 48	POPE	Pars opercularis	
49, 50	SF	Superior frontal cortex	
51, 52	LOF	Lateral orbitofrontal cortex	
53, 54	MOF	Medial orbitofrontal cortex	
55, 56	ENT	Entorhinal cortex	cingulate
57, 58	PARH	Parahippocampal cortex	
59, 60	CAC	Caudal anterior cingulate cortex	
61, 62	RAC	Rostral anterior cingulate cortex	
63, 64	PC	Posterior cingulate cortex	
65, 66	ISTC	Isthmus of the cingulate cortex	